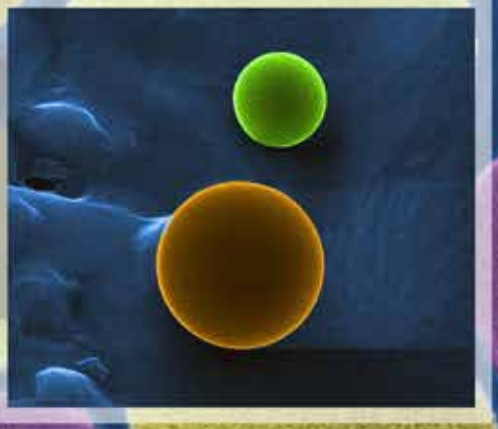
Laboratory Directed Research & Development
at Pacific Northwest National Laboratory



2013 Annual Report

Pacific Northwest
NATIONAL LABORATORY

Proudly Operated by **Battelle** *Since 1965*

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor Battelle Memorial Institute, nor any of their employees, makes **any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights.** Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or Battelle Memorial Institute. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

PACIFIC NORTHWEST NATIONAL LABORATORY

operated by

BATTELLE

for the

UNITED STATES DEPARTMENT OF ENERGY

under Contract DE-AC05-76RL01830

Printed in the United States of America

Available to DOE and DOE contractors from the
Office of Scientific and Technical Information,
P.O. Box 62, Oak Ridge, TN 37831-0062;
ph: (865) 576-8401
fax: (865) 576-5728
email: reports@adonis.osti.gov

Available to the public from the National Technical Information Service,
U.S. Department of Commerce, 5285 Port Royal Rd., Springfield, VA 22161
ph: (800) 553-6847
fax: (703) 605-6900
email: orders@ntis.fedworld.gov
online ordering: http://www.ntis.gov/ordering.htm

This document was printed on recycled paper.

PNNL-23265

Laboratory Directored Research and Development Annual Report

Fiscal Year 2013

March 2014

Prepared for
the U.S. Department of Energy
under Contract DE-AC05-76RL01830

Pacific Northwest National Laboratory
Richland, Washington 99352

Contents

Laboratory Director's Message . . . vii

Advanced Sensors and Instrumentation

Advanced Surface Mass Spectrometry for Characterization of Explosives. . . . 1
Advanced Switching Concepts applied to a Structure for Lossless Ion Manipulations (SLIM) . . . 3
Alpha Coincidence Techniques for Actinide Assay. . . . 5
Controlled Datasets to Inform the Study of Sensor Degradation for CBRNE Signature . . . 6
Isotope Enrichment using a Structure for Lossless Ion Manipulation (SLIM) Device . . . 7
Knowledge-based Automated DIDSON Data Extraction for the Assessment of Smolt Sized Salmon . . . 9
Laser Ablation Capillary Absorption Spectrometer for Trace Isotopic Sampling and Imaging . . . 10
Millimeter-wave Shoe Scanner. . . . 12
Miniaturization of Multi-Modal Regenerative Feedback Sensor . . . 13
Mössbauer Spectral Imaging. . . . 15
Optical Fiber Sensing System for Process and Health Monitoring (Stain and Strikes). . . . 16
Platform for High-Throughput Determination of Enzyme Kinetic Parameters for Hemicellulose Saccharification . . . 17
Prototype Fast Neutron Detector . . . 19
Reactor Materials Degradation During Plastic Deformation and Creep: High-resolution Experiments and Integrated Models of Microstructural Evolution and Materials Response . . . 21
Secondary Signatures for Provenance Attribution. . . . 23
Super Resolution Fluorescence Barcoding and High Throughput Microfluidics for Quantifying the Expression of Multiple Genes in Individual Intact Cells . . . 24
Ultra Low Noise Detectors: Building an Advanced Capability. . . . 25

Biological Sciences

Analysis of Functional Potential from Metagenome Data. . . . 28
Characterization of Signaling Networks in Single Cells . . . 29
Combining Proteomic Technologies to Create a Platform for Spatiotemporal Enzyme Activity Profiling . . . 30
Correlative High Resolution Imaging and Spectroscopy to Characterize the Structure and Biogeochemical Function of Microbial Biofilms . . . 31
Developing A Next Generation Biogeochemical Module for Earth System Models . . . 33
Development of Bromotyrosine Antibody Assay for blood and sputum samples. . . . 35
Directed Strain Evolution through Riboswitch-Controlled Regulatory Circuit . . . 36
Elucidating the Natural Spore Phenotype . . . 37
Engineering of Oleaginous Yeast for Production of Advanced Hydrocarbon Biofuels . . . 38
Exploring and Engineering Phototrophic-Heterotrophic Partnerships . . . 39
Fungal Siderophore Production for Recognition of Uranium Compounds . . . 40
Genome-Enabled Systems Approach to Predict Immobilization of Technetium in the Subsurface . . . 42
Imaging and Monitoring the Initial Stages of Biofilm Formation. . . . 43
Integrated Nano-Scale Imaging for Investigating Applications and Implications of Nanomaterials . . . 44
Mapping and Characterization of Organic Matter in Soil Aggregates using Laser-Ablation Sampling . . . 45
Microbial Processes Accompanying Deep Geologic CO_2 Sequestration . . . 47
Micro-Fluidic Models for Studying Microbial Communities—Integration of Micro-Fluidic Model Experimentation, Multimodal Imaging, and Modeling . . . 49
Multimodal Imaging of Cellulytic Microbial Activity in Oxygen-constrained Microenvironments . . . 50
Multiscale Simulation of Microbial Carbon Transformation in Soils: Connecting Intra- and Inter-Aggregate Scales. . . . 51
Optofluidics and Microfluidics for Exploring Biofuel Production at the Single Cell and Molecule Levels . . . 53
Predicting the Response of Complex Biological Systems . . . 54
Predictive Chemical Bond Information in Living Cells . . . 55

Proteomics Measurements of Functional Redundancy and Stability Testing of Cellulose Degrading Anaerobic Microbial Communities Within Engineered Bioreactors . . . 57
Pulmonary Injury from Acute Events Related to Nuclear Energy Production. . . . 59
Quantitative Modeling of the Metabolism of Technologically Important Fungal Systems . . . 60
Signatures of Environmental Perturbation - Microbial Community and Organic Matter Resilience . . . 62
Structure and Dynamics of Biological Systems . . . 63
Understanding the Processes that Govern Subsurface Microbial Communities. . . . 64

Chemistry

Bio-Inspired Actinides Recognition for Separation Science . . . 66
Characterization of Catalyst Materials in the Electron and Atom-Probe Microscopes . . . 68
Chemical Imaging Analysis of Environmental Particles . . . 70
Conversion of Biomass to Jet Fuels . . . 71
Design, Synthesis, and Activity Measurements of Monodispersed Multifunctional Catalysts for CO_2 Reduction . . . 73
Development of New Soft Ionization Mass Spectrometry Approaches for Spatial Imaging of Complex Chemical and Biological Systems . . . 74
Development of Preparative Mass Spectrometry for the Creation of Novel Catalyst Materials. . . . 75
Exploitation of Kinetic Processes in Gas Separations . . . 76
Fundamentals of Carbonate Formation: Interactions of Carbon Dioxide with Supported Metal Oxide Clusters. . . . 78
In Situ Molecular-Scale Investigations of Reactions between Supercritical CO_2 and Minerals Relevant to Geological Carbon Storage . . . 79
Increased Sensitivity and Improved Quantification of Th and U in Particles by SIMS . . . 80
Light Source Photocathode Performance and Development . . . 82
Localized Surface Plasmon Resonance Spectroscopy, Microscopy, and Sensing . . . 83
Molecular Structure and Interaction at Aqueous, Non-Aqueous Liquid Interfaces and Catalytic Solid Surfaces. . . . 84
Molecular-level Analysis and Microscopy Studies of Brown Carbon . . . 86
Monitoring Sorbed and Reacting Species of Lignin Fragments in Zeolites in an Aqueous Environment. . . . 87
NMR Spectroscopy of Pu-Containing Materials . . . 89
Novel Alloy Nanoparticle Materials for Catalysis and Energy Storage. . . . 90
Novel Inorganic Complexes for Tc Management in the Tank Waste . . . 92
Probing Composition and Structure of Polarizable Reaction Mixtures Inside the Pores of Supported Metal-oxo Catalysts . . . 94
Probing Structural Dynamics with High Spatial and Temporal Resolution . . . 95
Quantitative Imaging of Atomic Scale Chemistry Changes at Interfaces . . . 96
Simultaneous Electrochemical and Nuclear Magnetic Resonance Techniques for the Study of Electrochemically Active Biofilms . . . 98
The Statistical Mechanics of Complex Process in Bulk and Interfacial Environments . . . 100
Upgrading of Hydroprocessed Pyrolysis Bio-Oils by Selective Ring Opening of the Aromatics and Cycloparaffins to Paraffins for Diesel and Jet Markets . . . 101

Earth and Space Sciences

A New Non-aqueous Fracturing Technology for Eliminating Water Requirements for Fracturing Liquid-Sensitive and Non-Liquid-Sensitive Reservoirs . . . 104
Decision Support Research for Integrated Regional Earth System Modeling . . . 105
Developing a Dynamic LULCC Model for Spatially-explicit Future Realizations of Projected Land Use under an Integrated Regional Earth System Modeling Framework . . . 106
Development of Coupled Flow, Thermal and Geomechanical Capability for Carbon Sequestration . . . 108
Enhanced Sediment Geochronology Achieved Using Ultra-Low Background Materials and Ultra-Sensitive Detection Capability . . . 109
Enhancing EMSL Mass Spectrometry Capabilities for Characterization of Soil Organic Matter . . . 111
Exploration of Human and Environmental System Interactions due to Renewable Technology Penetration in the Midwest Pilot Region . . . 113

Measuring and Modeling the Climatic Effects of Brown Carbon Atmospheric Aerosols: Developing an Integrated Capability . . . 114
Microscale Reconstruction of Biogeochemical Substrates Using Combined X-ray Tomography and Scanning Electron Microscopy . . . 116
Numerically Robust Climate Simulation Through Improved Interaction between Model Components . . . 117
Predicting the Feasibility of Geologic Co-Sequestration of CO_2, SO_x and NO_x Under a Broad Range of Conditions . . . 118
Simultaneous ^{14}C and T Dating: A Case Study Using Soil Organic Matter . . . 119
Tank Residual Waste Stabilization to Reduce Contaminant Release . . . 120
The Integration of Water in iRESM . . . 121
Uncertainty Quantification and Risk Assessment Pipeline for Carbon Sequestration . . . 122

Energy Supply and Use

A Multi-Layer Data-Driven Advanced Reasoning Tool for Smart Grid Integrated Information Systems . . . 124
A Statistical State Prediction Methodology to Improve Reliability and Efficiency of Power System Operation . . . 125
Advanced Visual Analytic for the Power Grid . . . 126
Battery Manufacturing Cost Model Tool - Flow Battery . . . 127
Decision Support for Future Power Grid Organizations . . . 128
Development of the Capability for Assessing Benefits of the Smart Grid to Building Owners . . . 129
Electric System Intra-hour Operation Simulator . . . 131
Future Power Grid Control Paradigm . . . 132
GridOPTICS . . . 134
GridPACK: Grid Parallel Advanced Computational Kernels . . . 135
Hardware in the loop testing and power system simulation of high penetration levels of PV . . . 136
Hybrid Load Control Strategy for Demand Response System . . . 137
Market Design Analysis Tool . . . 138
Modeling of Distributed Energy Resources in the Smart Grid . . . 139
Multi-Zone Dynamic Commercial Load Modeling . . . 140
Numerical Simulator for the Utilization and Storage of CO_2 in Natural Gas and Petroleum Reservoirs . . . 141
Operations and Planning Fusion . . . 143
Optimal Sizing Tool for Energy Storage in Grid Applications . . . 144
Recyclable Methane Hydrate Inhibitors . . . 146

Engineering and Manufacturing Processes

Manufacturable Gel-based Membranes for Gas Separations . . . 148

Materials Science and Technology

Design and synthesis of peptoid-based functional materials . . . 150
Developing Next-Generation Multimodal Chemical Imaging Capability by Combining STEM/APT/STXM/HIM . . . 152
Development of Graphene/Ionic Liquid Hybrid Material for Ultracapacitors . . . 153
Development of Hard X-Ray Emission Spectroscopy Nanoprobe . . . 155
Directed Mesoscale Synthesis of Tunnel Structured Materials for Energy Applications . . . 156
Facet Specific Chemistry of Noble Metal Nanoparticles Using an Enhanced Scattering Infrared Scattering Near-Field Optical Microscope . . . 158
High Energy Density Non-aqueous Metal-Organic Redox Flow Battery . . . 159
Hybrid Electrodes for Next Generation High Energy Ultracapacitors . . . 160
Imaging the Nucleation and Growth of Nanoparticles in Solution . . . 162
Impedance Spectroscopy: Next Generation Tool for Waste Form Characterization . . . 163
Improving magnetoelectric coupling in novel single-phase multiferroic thin films of the $MTiO_3$ (M = Fe, Mn, Ni, . . .) family . . . 164
Improving the Performance of Li-Air and Li-S Batteries Using Polymeric and Metallic Nanomaterials . . . 166
Meso-scale Science and Technology: Manufacturing of Nanostructured Soft Magnetic Materials . . . 168
Metal Hydrides for Thermal Energy Storage . . . 170

Novel CO_2-Selective Polymer/Double Salt Composite Membranes for Continuous CO_2 Removal from Warm Syngas 171
Novel phase selective (drygel) synthesis of Metal-Organic Frameworks 172
Novel Window Coatings for Dynamic Thermal Control via Infrared Switching 173
Optical properties modification in complex oxide epitaxial films via alloy formation 174
Optically Stimulated Luminescence Data Storage 175
Perfecting Atom Counting: 100% Efficient Mass Spectrometry 176
Photocathode Development for Next-Generation Light Sources 178
Probing Structure-Property Relationship of Energy Storage Materials Using Ex Situ and In Situ Dynamic Microscopy and Spectroscopy with High Spatial and Fast Temporal Resolution 179
Rare Earth-Free Phosphors for Lighting Applications 181
Site Specific Atomic Resolution Probing of Structure-Property Relationship Under Dynamic and/or Operando Conditions Using In Situ and Ex Situ Chemical Imaging Based on Multi-Instrument Approach 183
Towards understanding interfacial chemistry at reactive solid/liquid interphases 185
Understanding and Control of SEI Layer (Solid -Electrolyte Interface and Interphases) in Multivalent Energy Storage Systems 186

Mathematics and Computing Sciences

A Compressive Sampling Framework for Interactive Visualization of Massive Datasets 189
A Distributed Systems Architecture for the Power Grid 191
A Multi-Modal Integration Framework for Chemical Imaging 192
A New Modeling Approach for Biology: Combining Natural Selection and Thermodynamics for Biodesign and Natural Systems 194
Actionable Visualization Tools For Power Grid Situation Awareness 196
Advanced Optimizations for Extreme-Scale Homogeneous Systems 197
Analytic Framework: Signature Discovery Workbench 199
Compressive Sensing for Threat Detection 200
Computational Optimization and Predictive Simulation for Synthetic Biology 201
Cyber Security Testbed and Dataset Generation 203
Data-Intensive Algorithms for Bioinformatics-Inspired Signal Detection 204
Developing Functionality and Performance Enhancements to the Global Array Toolkit 205
Development of Parallel Multi-Reference Coupled Cluster Capabilities 206
Discrete Mathematical Foundations for Cyber Systems Analysis 207
Enabling the Meaningful Exploitation of Integrated Regional Earth Systems Data 208
Encrypted CPU Instruction Stream 211
Establishing a Strategic Goods Testbed to Disrupt Illicit Nuclear Trafficking 212
Experts Inundated with Data: the Biomarker Problem 213
Fishing for Features: Discovering Signatures when the Underlying Phenomenon is Poorly Understood 214
Geological Sequestration Software Suite Framework 216
GRADIENT: Graph Analytic Approach for Discovering Irregular Events - Nascent and Temporal 217
Hierarchical Signature Detection in High-Throughput Environments 218
Immense Social Media Stream Analytics 219
Integrating Advanced Optimizations for Extreme Scale Systems 221
Intelligent Networked Sensors Capable of Autonomous, Adaptive Operations in the Electric Power System 222
Kritikos: Identifying Cyber Assets and Assessing Criticality in Terms of Business Processes 224
Linear Algebra Solvers and Associated Matrix-Vector Kernels for Power Grid Simulations 225
M&Ms4Graphs: A Multi-scale, Multi-Dimensional Graph Analytics Framework for Cyber Security 226
Manifold Learning for Accurate Search and Locate Tasks in Image Datasets 228
Mapping Molecular Dynamics Algorithmic Parallelism to Heterogeneous Architectures 230
Multimedia Analysis of Cyber Data 231
Multimedia Information Fusion 232
Multi-Resolution Data Model and Directed Data Reduction, Reconstruction and Aggregation 233

Multi-Source Signatures of Nuclear Programs . . . 235
Nanoscale-Macroscale Three-Dimensional Integration Using High Performance Computing . . . 236
Network Analysis and Modeling of Illicit Nuclear Trafficking . . . 238
Next Generation Network Simulations for Power System Applications . . . 240
Real-time High-Performance Computing Infrastructure for Next-Generation Power Grid Analysis . . . 242
Robustness . . . 243
Scalable Knowledge Extraction on Extreme Scale Scientific Data . . . 245
Scalable Sensor Data Management Middleware . . . 246
Scalable Solvers for Uncertainty Quantification of Large-scale Stochastic Partial Differential Systems . . . 247
Scire: Scientific Process for Validation and Verification . . . 248
Semantic Workflows for Signature Discovery . . . 249
Signature Discovery Analytic Framework . . . 250
Signature Quality Metrics . . . 251
Single Node Optimizations for Extreme Scale Systems . . . 252
Synergistic Integration of Feature Recognition and Analysis for Chemical Imaging Data . . . 254
Targeting Extreme Scale Computational Challenges with Heterogeneous Systems . . . 256
TASCEL: An Execution Model for Task-Based Optimizations . . . 257
Theory of Resilience . . . 258
Visualizing Uncertainty in Conceptual and Numerical Models for Geological Sequestration . . . 260

Nuclear Science and Engineering

Argon-39 Measurement . . . 262
Atomic Mass Separation for Enhanced Radiation Detection Measurements . . . 263
BazookaSPECT Neutron Imager . . . 264
Characterization of a 14 MeV Neutron Generator and Measurement of Fission Products Produced . . . 265
Exploiting Correlated Radiographic and Passive Signatures for Threat Detection in Cargo . . . 266
Three Dimensional Neutronics Analysis Capabilities for Nuclear Archeology Applications . . . 267
Ultra-Low-Background Gas Measurement: Building an Advanced Capability . . . 268
Uranium Enrichment Facility Signature Exploitation . . . 269
ZnS Scintillators . . . 270

Physics

Absolute Gas Counting Measurement Techniques . . . 272
Anthropogenic Uranium Detection with X-ray Microscopy . . . 273
Highly Efficient and Cost Effective Gamma Detection Arrays for FRIB . . . 274
Low Energy Threshold Germanium Detectors and Science . . . 275
Resolving the Reactor Neutrino Anomaly by Precision Beta Spectrometry . . . 276
Search for New Physics at the Intensity Frontier . . . 277
Super CDMS Initial Engagement . . . 278
Ultra-Precise Electron Spectroscopy to Measure the Neutrino Mass . . . 279

Laboratory Director's Message

The mission of Pacific Northwest National Laboratory is to transform the world through courageous discovery and innovation. Every day, scientists and engineers at PNNL set out to fulfill the promise of that mission by advancing scientific discovery and delivering critical solutions to the challenges that face our nation and the world.

The science and technology we work with, create, and improve upon inspires and enables the world to live prosperously, safely, and securely. Our investments in Laboratory Directed Research and Development are essential to our ability to advance our strategy and realize this mission. LDRD investments help us nurture science and technology capabilities while capitalizing on the breadth and depth of talent that our staff possess. This report describes how we conduct our LDRD program in compliance with DOE objectives and guidelines. We use rigorous internal and external peer review to maintain the scientific value and soundness of the research enabled by our LDRD program.

With great pride in our researchers' accomplishments, I present PNNL's Fiscal Year 2013 Laboratory Directed Research and Development Annual Report.

Michael Kluse
Director, PNNL

Advanced Sensors
and Instrumentation

Advanced Surface Mass Spectrometry for Characterization of Explosives

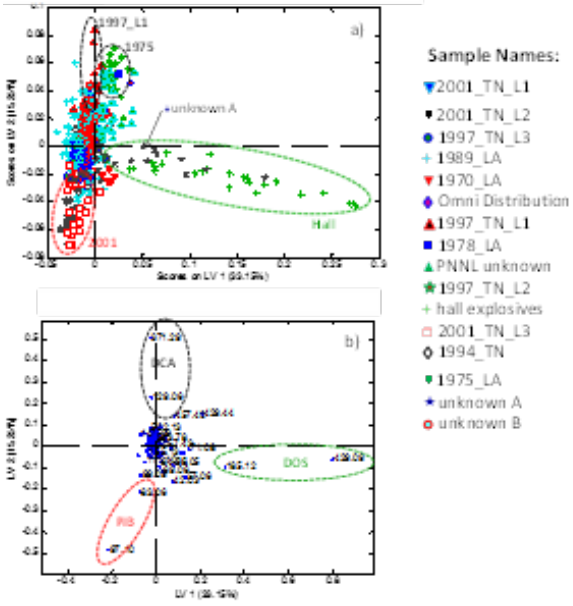
Christine Mahoney

The United States requires precise tools that can rapidly and accurately characterize explosive compositions. The Environmental Molecular Sciences Laboratory (EMSL) has state-of-the-art facilities that contain a unique, exclusive suite of analytical tools. Developing these facilities and tools for national security interests will help our government prevent and respond to terrorist acts in a timely manner.

We intend PNNL to be a center of excellence in addressing problems relevant to national security. As such, we are currently developing several novel methods available only through EMSL, a state-of-the-art PNNL user facility, for national security applications. Previous results indicated that C-4 sample signatures from around the world could be distinguished from each other. Specifically, this study highlighted secondary ion mass spectrometry (SIMS) as a powerful chemical imaging forensics tool for direct characterization and differentiation of explosive components (trace and bulk). An infrastructure has been developed for future collaborations, a relationship particularly beneficial for the integrity of our national security.

Currently, we are developing several novel instrumentations at EMSL and throughout PNNL for simultaneous elemental and molecular characterization of materials of interest to national security. As a starting point, we obtained a series of composition C-4 plastic explosive samples from across the country through collaborations with the Bureau of Alcohol, Tobacco, Firearms, and Explosives. These samples are commercially available and have been characterized extensively using gas chromatography-mass spectrometry in previous work. The goal was then to use our suite of advanced metrology tools for detailed characterization and differentiation of the C-4 samples based on a combination of unique organic signatures and trace elemental compositions.

Both bulk analysis and imaging microanalysis methods were developed for characterization of 17 different C-4 samples, where sample preparation methods were optimized for each method utilized. After preparation, the samples were distributed to various experts for method development. A full data set was obtained from all instrumentation utilized. During FY 2012, we focused on developing individual methods and identification of unique molecular and elemental signatures using multivariate statistical analysis approaches such as principal components analysis (PCA). For FY 2013, we developed a robust method of classification using partial least squares-discriminant analysis (PLS-DA) for each individual method with subsequent Bayesian integration of probabilities. We outlined the process using SIMS, one of the primary imaging methods used for signature discovery in the C-4 samples, as an example.



PLS-DA statistical classification process using positive ion SIMS. The top panel shows the scores plots for latent variables that capture 33% and 15% of the variance in the total dataset, respectively. The corresponding mass spectral loadings describing the LV variance are displayed in the bottom panel. Correlations between scores and loadings can be seen by the ellipses, drawn in for clarity. It is evident from this plot that the sample is highly correlated with hall explosives.

Positive and negative ion large area maps were obtained using SIMS from 17 different nanoimprinted C-4 samples. Mass spectral data were then averaged from 12 different regions in each C-4 residual map. Peak areas were selected and tabulated into a rectangular matrix, which was subsequently imported into Matlab for PLS-DA analysis, a multivariate statistical approach that uses the same projection methods PCA for optimization of a classification model. The results of PLS-DA are ideal, as they are presented as probabilities.

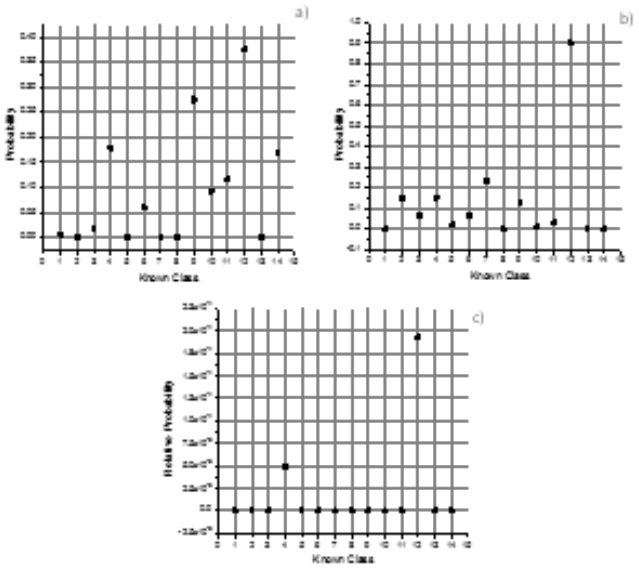
In this particular series of 17 samples, we identified 14 unique classes. Two "unknown" samples were run independently of the original dataset (e.g., several months later) for verification purposes. The samples selected were called "unknown A" (hall explosives) and "unknown B" (1994

PN12007/2408

C-4). These samples were distributed to all methods experts for analysis, and the results were input into the PLS-DA algorithm containing the original training datasets. The class was identified for each method independently using the discriminant analysis approach described earlier. The results are thus a series of probability matrices, plotting the likelihoods that each unknown sample belongs to one of the 14 classes. The classification is made based on which class contains the highest likelihood.

In the case of unknown A (hall explosives), only SIMS and electrospray ionization-FTICR were able to classify the samples correctly using this procedure. All of the other methods resulted in an incorrect classification. For unknown B (1994 C-4), the individual classification results were even worse. Only Aspex was able to correctly classify this sample, and with only a limited margin, as there was a lot of noise in the data.

What is important is that when integrating these data by multiplying the individual probability matrices, the noise drops out of the data, and what is left is a more accurate classification with a high level of certainty. For example, with unknown A, once the data was integrated from all of the methods, most of the probabilities for the incorrect classes dropped to zero (with the exception of a minor contribution from 1989 C-4). Even more shocking was that for unknown B after Bayesian integration, all probabilities other than for sample 1994 (class 5) dropped to zero, making sample 1994 the only viable class. This was particularly surprising given the poor quality signatures that the 1994 sample provided. It was expected that this sample would be representative of a failed case. We conclude from these results that this type of robust characterization can be achieved only through the integration of the data, and that one method alone is not powerful enough for classification purposes.



Second step of the PLS-DA classification process through discriminant analysis. Panel a) shows the resulting probability matrix from the unknown sample using positive ion SIMS data. Panel b) shows similar results using positive ion ESI-FTICR data. Both methods predict the unknown as belonging to class 12, though there is a significant amount of noise in the data. Panel c) shows the integrated data from positive and negative ion SIMS methods, positive ion ESI-FTICR, Aspex data, and laser ablation ICP-MS. The noise is significantly reduced, and we can be more confidently state that this "unknown A" sample is from class 12.

One of the more disappointing aspects of this work was that there was a surprising amount of error during the verification process. For the SIMS data, additional signatures were observed in the unknown data that were not observed in the original training dataset. These signatures are attributed to contamination from the environment. For the ESI-FTICR negative ion data, a completely different ionization process (likely due to the use of a different bottle of solvent) was observed in the unknown dataset with respect to the training set, resulting in the formation of a completely different mass spectrum. Hence, while the peaks were still informative, the model could not possibly be followed as the assumption of similar data was not able to be made. The nano-DESI (both positive and negative ions) data had similar problems as did the ESI data. Finally, the Aspex data was found to be particularly sensitive to background particle contaminants.

Despite this small setback, the integrated data could still identify correctly the compounds with a high confidence level, proving that this is indeed a robust data analysis procedure. It also shows that these methods will be particularly useful for analysis of "dirty" samples collected in the field.

During FY 2014, we will continue multi-technique data analysis from the previous year and apply advanced statistical analysis approaches developed in 2013 for the determination of more robust signatures. After obtaining relevant biological samples, we will characterize and image molecular distributions and content as well as perform trace elemental mapping in selected microbial samples.

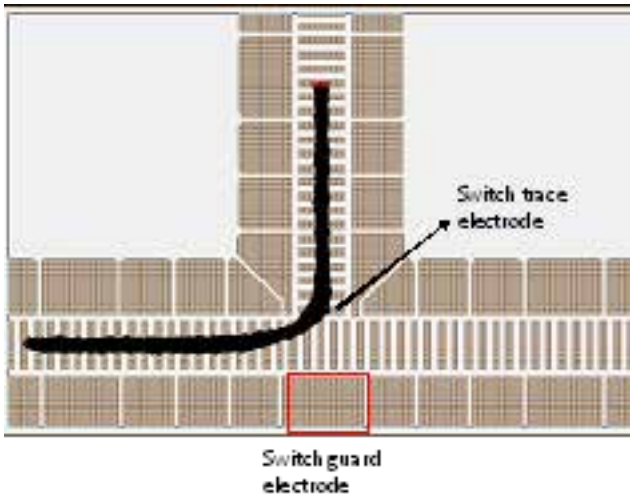
Advanced Switching Concepts Applied to a Structure for Lossless Ion Manipulations (SLIM)

Richard D. Smith

This project seeks to develop and initially demonstrate an advanced ion switching concept to a structure for lossless ion manipulations (SLIM) device.

The roles of mass spectrometry and other instrumental approaches based on ion measurements are continuously expanding in many fields of science. New opportunities to impact these fields significantly are becoming limited by the conventional approaches currently used for ion manipulations, in particular the expensive fabrication as well as highly "lossy" nature. Possible applications include use in conjunction with massively parallel mass spectrometry analyzers, a development that would provide orders of magnitude gains in 'omics applications. As ion manipulations become more sophisticated, conventional instrument designs and ionoptical approaches become increasingly impractical, expensive, and inefficient.

The primary objective of this project is to simulate and test a switching concept in SLIM device. The switch is a cornerstone in developing the SLIM device and will enable complex ion manipulation. The SLIM devices we envision make use of near lossless RF confinement to focus, transport, and trap ions in low pressure regions and are advantageous to ion mobility separations and gas phase reactions. Because SLIM devices will allow electric fields and timing to be controlled and enable sophisticated ion manipulations to be realized, essentially allowing complex integrated circuits to be constructed in the gas phase in a completely novel fashion.



Switch configuration in SLIM device

Aim 1: Model ion motion inside the SLIM device. Ion simulations software was used to calculate ion trajectories through the different switch configurations. In addition, a switch configured according to SLIM design concept was constructed. By applying appropriate RF and DC potentials, field gradients can be applied that can switch ions into an orthogonal board from its initial location. The key challenges for optimum switch performance are the timing needed to switch the ions into the next section (dependent on ion mobility) and achieving the switch with minimum or no loss of time resolution.

At operating pressures of 4 torr, ions of standard mobility values (~1 cm^2/V.s) need a voltage bias of 100 V applied for at least > 50 µs to the switching guard electrode. If the voltage is lower or switch time is shorter, the ions do not see the switching gradient, and efficiency is not 100%. Once ions switch into the orthogonal SLIM board, the critical step is identifying resolution loss while ions execute a turn. Ions at different positions and an ion packet executing a turn travel different path lengths (commonly known as the racetrack effect). This leads to loss of time resolving power (i.e., the ion packet becomes more diffuse in time after the turn). The loss must be minimized while ions execute turns in order for a complex device like ion cyclotron with multiple switches and turns to be effective in high resolution separation or for other complex ion manipulations.

To achieve the best switch design, different configurations were evaluated for performance. The time resolution in a T-section while ions travel in the straight path was calculated at 13, which served as baseline. When ions executed the switch and turned, the resolution was calculated as 12. The slight loss of resolution is from ions executing a wide turn. By increasing the field penetration into the straight section from the fields in the orthogonal T portion, there is potential for making ions execute sharper turns and avoid loss of resolution. A second design was evaluated on which the length of the switch trace was increased to allow ions to feel the switch gradient and thus make a more sharpe turn. A third design with orthogonal electrode inserted inside the straight portion can also provide better field penetration. SIMION simulations help us to see the relative benefits and caveats in each of these concepts.

The resolution with a long switching electrode was calculated as 13. The improvement in resolution is marked; however, there is increase in ion loss during switch. This is presumably due to the fact that there are several adjacent electrodes with same RF phase due to the orthogonal positioning of the switch electrode. This leads to drop in the

PN13090/2571

pseudo potential well confining the ions. Increasing the length of switch electrode is detrimental due to poorer pseudo-potential field though the relative field strength of DC focusing is higher resulting in better resolution. This loss in signal becomes more predominant when an insert electrode is added to further achieve field penetration as tested in the third design. This increased loss leads to loss of resolution, even though hypothetically an increase in resolution is anticipated.

Aim 2: Fabricate and test a SLIM device with advanced switching. We will design and fabricate a SLIM device based on results obtained from Aim 1. We will also evaluate the performance of the SLIM device in terms of sensitivity and efficiency of ion transport during the switching. In addition, the evaluation will include finding the optimum experimental parameters to maximize ion mobility separation.

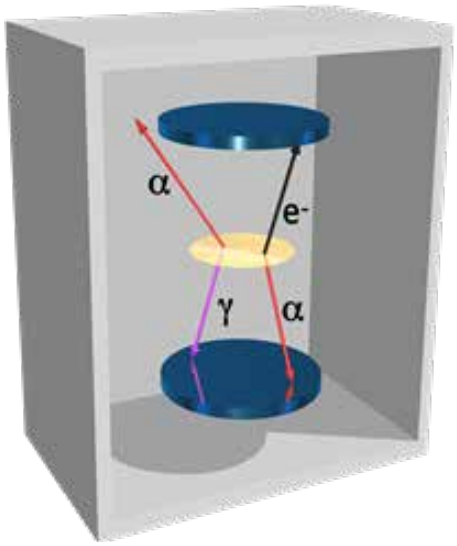
The electronic circuitry for two switch configurations based on SIMION simulation from Aim 1 was designed using Eagle software and was sent for fabrication. The vacuum housing that will be used to evaluate the switch was also fabricated. Efforts are currently underway to evaluate and test the switches. We anticipate completing Aim 2 in FY 2014.

Alpha Coincidence Techniques for Actinide Assay

Glen A. Warren

The effectiveness of conventional measurement techniques for environmental monitoring is limited by background and other interferences. This project will develop and demonstrate a new concept in radiation detection that disentangles these interferences and reduce background to improve our ability dramatically to assay environmental samples.

Examples abound of how actinides have been used in environmentally related studies. By measuring the ratio of ^{240}Pu to ^{239}Pu for example, one can determine whether a source is consistent with local or global fallout. Transuranics are strongly associated with particulates in water so that they can be used to study scavenging, the removal from water by attaching to particulates. The activity of plutonium as a function of depth beneath the sea floor surface has been used to measure bioturbatory processes that mix the surface into the rest of the sea floor. Likewise, other elements have been used to date fossil corals or carbonate rocks and study seasonal variations in river and sediment flow. Sample preparation requires time, resources, and people to complete. While the effort required to complete a few sample preparations may not be particularly arduous, there are many studies in which thousands of samples are collected, as some environmental science questions require easily scalable measurement techniques. Thus, minimizing sample preparation is an important aspect of the overall environmental monitoring effort.



Conceptual drawing of coincidence detection system, in which coincident α's and electrons are measured in separate silicon-based detector.

Any alpha-decaying radioisotopes have significant interferences when using current alpha-spectroscopy capabilities. By measuring other signatures that are generated in coincidence with the alpha particle, one can increase the probability of removing interferences between different radioisotopes. A good example of an isotope that may benefit from this approach is ^{238}Pu. Plutonium-238 is not typically measured with mass spectroscopy techniques because of interferences with much larger quantities of ^{238}U, while it cannot be measured using alpha-spectroscopy without chemical separations because of interferences with alphas from ^{241}Am. A system able to measure particles in coincidence with alphas from ^{238}Pu may eliminate the need for chemical separation and enable large-scale assaying of such samples. Thus, the project goal is to develop a radiometric assay system for alpha-decaying radioisotopes that focuses on coincidence signatures to increase specificity of the system.

To assay actinides in environmental samples simultaneously and rapidly, we are developing a detector design capable of observing coincidences between conversion electrons and alpha particles to conduct isotopic-specific assay on complex samples. The first year of the project is an initial scoping study to evaluate the type of detection methods that could be utilized. In the second year, a more thorough examination of the impact of the sample, both in terms of preparation and background contributions, will be completed. In addition, initial empirical measurements will be conducted. For the third year, a demonstration of the detection concept on a prototype detector using an application-relevant sample will be completed.

The objective of this project is to demonstrate a new detection modality that could significantly enhance radiometric assay capabilities for a variety of applications requiring assay of actinides in complex samples. This new approach has the potential to significantly reduce assay time, provide new sample diagnostic capability and be a fieldable technology. In FY 2013, we completed an initial scoping study to understand various aspects of designing the detector system. These aspects include an improved understanding of the signatures, an understanding of the substrate impact that holds the sample, and a comparison of two basic detector approaches: dual-gas proportional counters and double silicon-based detectors. The most important finding was that in the absence of background, the approach should achieve reasonable measurement uncertainties in a few hours for relevant samples.

The effort for FY 2014 will focus on evaluating the feasibility of our measurement approach for a realistic sample. Understanding the strength of the coincidence signal in a sea of background is a universal issue for detector design. Even though observing coincidence signatures reduces the background, accidental coincidences formed by two unrelated decays creates a background above which the system must be able to observe the true signal. We will evaluate the feasibility of the measurement approach by considering the impact of the background.

PN13019/2500

Controlled Datasets to Inform the Study of Sensor Degradation for CBRNE Signature

Marvin G. Warner

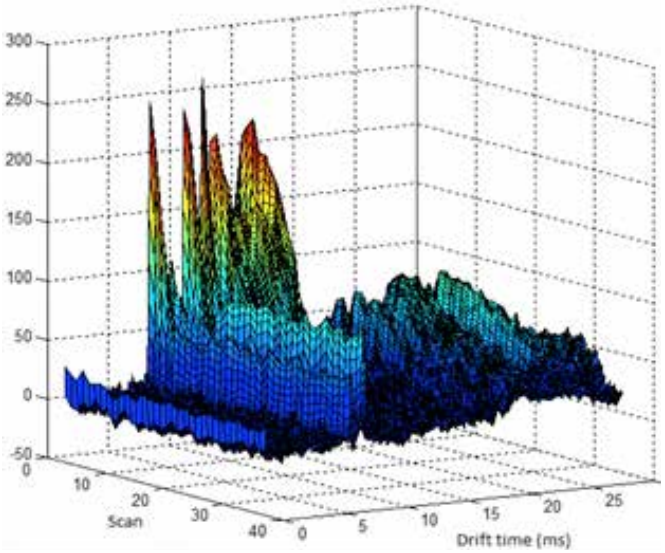
This project is developing a process for the study of sensor degradation under a specific set of operational conditions. If successful, we will have developed a streamlined, formalized process to study the degradation and drift of signatures arising from physical sensor measurements. This will be broadly applicable across many application domains and will have impact on our ability to determine the operational reliability of a given platform and the signature it measures.

Measurements taken by a physical sensor are a composite of features derived from the target phenomenon, operating environment, and the sensor itself. Signature detection entails interpretation of the measurement data and requires a fundamental understanding of how each of these impacts the observed measurements. This project seeks to utilize statistical design of experiments to avoid the explosion of experiments that arises from a full factorial exploration of a large number of variables. To that end, we are parameterizing the possible environmental and sensor effects on measurements and plugging them in to a variety of different experimental design models prior to performing the experiments in the laboratory.

The combination of subject matter expert (SME) driven input, statistical design of experiments, and focused laboratory experimentation allows us to assess the relative impact each parameter has on the observed measurement data, and determine how an understanding of measurement effects can be used to improve signature detection making and decision making capabilities. To accomplish this goal, a suite of measurement systems representing fundamental classes of physical measurement (e.g., mass loading, electro-optical, electrochemical) are used to examine how the relative impact of different parameters remains the same (generalizable) or differs across sensor classes.

In FY 2012, the bulk of our efforts focused on investigating the effects of signature drift and degradation on our ability to detect explosives and chemical agents using ion mobility spectrometry (IMS). IMS is a ubiquitous analytical instrument deployed in numerous operational environments (e.g., airports and forward deployed military bases) and is used primarily to screen for the presence of explosive chemicals and precursors. Specifically, IMS was chosen as a platform where it was relatively straightforward to identify both the relevant operational conditions as well as potential confounders that could make measurement of signature difficult. In the first year of this effort, we chose to examine the effects of analyte concentration, humidity, and a single confounder, diesel fuel.

Once the appropriate degradation parameters and their relatively magnitudes and ranges were determined through exhaustive SME polling, a three-factor experiment was designed in FY 2013 to investigate the interplay of each of these variables and to determine what the operational envelope – the range of conditions where signature detection occurs with near 100% confidence – was of the IMS system. One such experiment was conducted to determine the effect of the relative humidity on our ability to detect and classify a signature of a chemical agent surrogate (dimethyl methylphosphonate, DMMP). While it is known that humidity can cause shifts in peaks in an IMS spectrum, using the experimental design described above, we determined not only the effects of humidity but also how humidity and other degradation parameters (i.e., signature concentration and the presence of confounders) interact to cause signature drift and degradation. Using this method, we found many interdependencies among the degradation parameters. This leads to a powerful experimental design process that can greatly assist in determining how well a sensor such as IMS will operate under a variety of environmental conditions.



The relative effect of humidity on the IMS signal over time for a chemical agent surrogate (DMMP), whose peaks appear as the intense (red/yellow) peaks but begin to diminish over time and drift as effect of the relative humidity. The y-axis is an arbitrary signal intensity.

In FY 2014, we will continue to develop the experimental design process through the application of lessons learned in the IMS example given above to a wide range of different signature measurement challenges. In addition, we will focus some of our efforts on effective ways to capture SME input to streamline the experimental design process.

PN13021/2502

Isotope Enrichment using a Structure for Lossless Ion Manipulation (SLIM) Device

Richard D. Smith

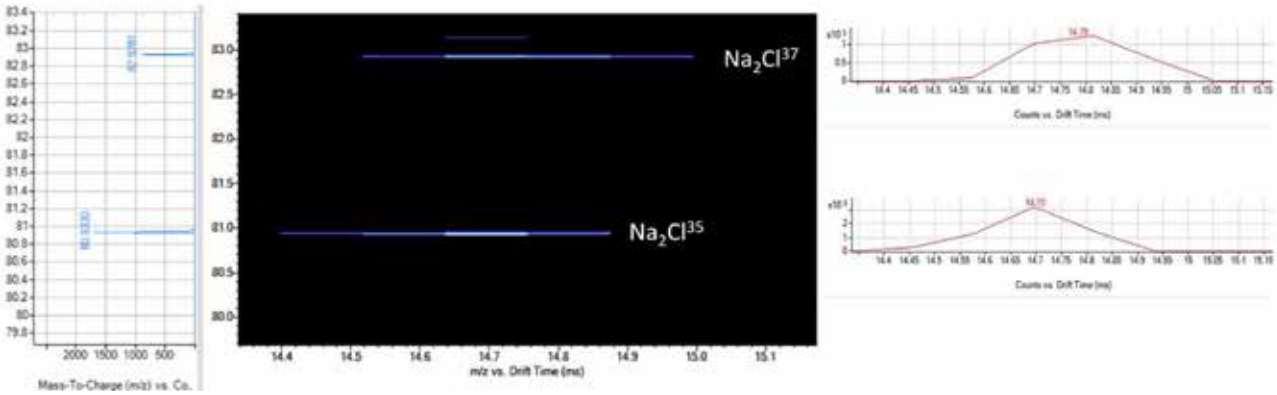
This project seeks to develop and initially demonstrate the potential of novel structure for lossless ion manipulations (SLIM) technology for isotope enrichment.

The primary objective of this project is to explore and initially demonstrate the potential for novel SLIM devices, very recently conceived at PNNL, as alternative platforms for isotopic enrichment. The new SLIM devices separate ions based on ion mobilities rather than mass to charge ratios. They offer the attractions of not requiring high vacuum, needing only low levels of energy for ionization and operation, and being inexpensive to construct. SLIM devices use RF and DC potentials applied to patterns of electrodes on two parallel planer surfaces to control ion motion and manipulate ions. The basic manipulations include controlling the ions path, switching, etc. as well as trapping and accumulating ions, and providing the potential for construction of much more complex devices and more sophisticated manipulations due to their lossless nature. Of particular note is that these devices can be fabricated using conventional printed circuit board (PCB) technologies.

This initial LDRD effort has supported the initial reduction to practice of the SLIM concept, which we believe has broad potential utility. However, the overall objective of this project is to initially explore the application of SLIM technology to the area of isotope enrichment. The research aims to address two key questions that will determine the applicability of the SLIM technology for isotope enrichment: the ability to separate isotopes of interest with the ion mobility separation power of the SLIM device, and determine if sufficient ion currents are achievable without degrading the necessary ion mobility separation power. The project has three key aims designed to provide an evaluation of SLIM technology for isotopic enrichment. The progress under each aim is described below.

Aim 1: Experimentally determine the extent of isotopic separation obtainable using surrogate species (e.g., potassium isotopes). To determine the feasibility of separating isotopes with ion mobility spectrometry (IMS), sodium containing ions were analyzed using an existing IMS-MS platform with a 1-m drift cell and resolving power of ~50. The mass spectrum of a sodium chloride solution showed peaks for both $Na_2{}^{35}Cl^+$ and $Na_2{}^{37}Cl^+$ species, allowing the evaluation of how a 2 Da difference in mass due to the chlorine isotopes changed the drift times of the Na_2Cl^+ ions, providing a potential basis for separation. The measurement of IMS drift times for each species indicated a shift (i.e., separation) between the two isotopes in line with expectations from theory, a mobility difference of ~1.5%. Although resolution was insufficient to separate the two species completely, partial separation shows that enrichment is feasible. This indicates that full isotopic separations should be obtainable with the SLIM cyclotron concept we have developed. Currently, high resolving power IMS separations (>100) have been limited by low sensitivity and by requiring high pressures and low temperature. The small size and lossless nature of SLIM devices offer a promising avenue for increasing IMS resolving power and providing separations that are presently unachievable.

Aim 2: Conduct simulations to determine if ion currents achievable are sufficient to support utility of SLIM type devices for isotopic enrichment based upon ion mobility separations and if separation power can be maintained in presence of the applicable space charge. We conducted modeling and simulations calculations to explore



The mass spectrum (left) and drift time plots (right) of $Na_2{}^{35}Cl^+$ and $Na_2{}^{37}Cl^+$ species illustrating the shift in drift time between the isotopes.

PN13078/2559

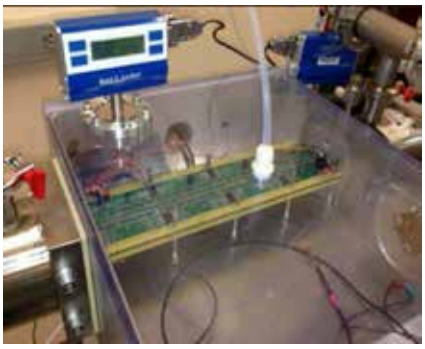
the achievable ion currents. Our experience has shown that the primary limitation on the IMS separation resolving power will arise due to coulombic interactions of ions. Coulombic repulsion leads to the ion cloud expansion that will contribute an additional effect to the diffusional broadening that will degrade separations if excessively high currents are utilized. Thus, such calculations provide a fast route to optimizing the combination of ion current and separation power needed for a specific separation. To achieve maximized ion currents while preserving a resolving power sufficient for the isotope separation, it is necessary to configure the SLIM device to limit coulombic effects.

A trial SLIM configuration assumed an IMS separation region 1 m long orthogonal to the ion injection direction. The device is filled with nitrogen at 4 Torr (the condition studied experimentally with the prototype SLIM device under Aim 3). The electric field used for IMS separation is 100 V/cm, below the breakdown voltage limit for pressure conditions. Ion mobility K = 500 cm^2/V s (calculated for m/z 40, z=1, p=4 Torr N_2). The coulomb expansion of the cylindrical ion packet calculated for such a configuration is r=8.3 mm, for the IMS resolving power of 60. The ion current maximum was ~10 nA, corresponding to a total ion count of separated ions 6 $x10^{10}$ ions/sec. To increase the ion current further, we can utilize a stack of such devices.

We also used simulations to explore if SLIM devices can provide sufficient resolving power to completely separate isotopic species based upon their mobilities. SIMION was used to study ion transmission in a simple SLIM "straight section." The total length of the section modeled (10 mm trace width with 5 mm board gap) was 76 mm and assumed a drift field for the ions in the device of 13V/cm (close to the experimental value used under Aim 1). Ions with nominal masses 50 and 52 were used to study the isotopic separation in the SLIM section. The estimated ion mobilities based on hard sphere model were 0.20 cm^2/V-s and 0.19 cm^2/V-s, respectively. Initially, the ion confinement and transmission efficiency were studied for the low m/z ions by varying the trapping RF frequency at 150V_{p-p} RF amplitude assumed. At conventional RF frequencies of ~10^6 Hz, low m/z ions are not confined and are lost to the electrodes. At 5×10^6 Hz, ions are effectively confined and transmitted. The corresponding arrival time distribution was calculated for the SLIM mobility cyclotron conditions and showed that more than adequate resolution (~960) for baseline ion separation was achievable.

Aim 3: Obtain experimental verification that ion current and resolution performance from Aims 1 and 2 can be achieved. The key effort this FY involved the implementation and evaluation of an actual SLIM device. In the initial work, a vacuum housing was designed and fabricated to accommodate a small SLIM device (which we note used smaller dimensions than the system modeled under Aim 1, and thus constrained to lower maximum currents that can be estimated to be a few nanoAmps). The fabricated housing was successfully tested for its vacuum seal. This SLIM device was integrated with an ion funnel trap for connection to the electrospray ion source and a "rear" funnel and quadrupole for coupling to an Agilent TOF mass spectrometer (MS). The SLIM building blocks were designed and fabricated by printed circuit-board technology (PCB) technology. The initial device consisted of elements the provided a short 30 cm straight section. System-control software was written in C# to allow the control of all voltages. The software also enabled acquiring data from the MS into an analog-to-digital acquisition card.

First SLIM experiments involved passage of ions from the ion source through the SLIM device and to the rear funnel to evaluate the transmission characteristics of the device and ion mobility separation. The initial experiments utilized a Faraday cup for ion current measurements and which revealed successful transmission of ~500 pA ion current through the SLIM device; this is a level of ion current higher than that of ion current from our PNNL-developed ion mobility separation (IMS) platform. We believe that refinements to the interfaces with both the ion source and the MS will enable higher currents to be transmitted. The SLIM device was then successfully coupled to the MS to characterize the ions and the sensitivity of this initial device. Initial experiments revealed a broad range of ions of different m/z can be successfully transmitted through the device in a lossless fashion (after some initial losses at the ion source interface as expected due to a mismatch in this initial work for the effective m/z or mobility ranges confined by the ion funnel and the SLIM device). Ion mobility separations that could be achieved in this short SLIM device were also evaluated. Initial results indicated a resolving power of ~18-20 for the 30 cm long straight SLIM device, which is ~65% of the maximum theoretically achievable resolving power. Experience indicates that the less than theoretical resolution is likely due to the limitations of the initial design of the interface between the SLIM device and MS, and that a resolving power > 90% of the theoretical resolving power should be achievable.



Initial SLIM device fabricated 30 (cm drift cell) shown coupled with conventional electrospray ion source (right) an Agilent TOF MS (left).

Our initial efforts have been highly supportive of the potential for SLIM devices and have indicated potential utility for isotopic enrichment that will be explored in more detail under the planned Year 2 efforts.

Knowledge-based Automated DIDSON Data Extraction for the Assessment of Smolt Sized Salmon

Kenneth D. Ham

Automated data extraction for fish behavior from DIDSON multibeam imaging sonar creates new possibilities for studying how fish interact with dam structures. The tools we created help make it practical to evaluate actions that are intended to make dams more fish-friendly.

Multibeam imaging sonar provides a way to peer into an otherwise murky underwater environment. Few other options exist to observe fish behavior under field conditions where water clarity can be low, and adding artificial lighting might influence behavior. Merely observing behavior can be useful, but the factor limiting rigorous studies of fish behavior is the extensive amount of human observation currently required to process the multibeam imaging data. Analyzing behavior recorded with the imaging device requires a prohibitive amount of time to evaluate using human observation.

The goal of this project is to speed and automate processing of fish behavior to reduce the need for human intervention, allowing for extensive observation times to be rapidly reduced to useful information. We created software tools that allowed fish behavioral information to be automatically extracted from DIDSON multi-beam imaging data files. The software parses the raw files and extracts header information that describes how the instrument was configured to properly orient the echo data in space and time. Display of the raw file as video is an option, but this project implemented two key approaches that greatly enhance the ability to observe the behavior of fish after processing is completed. A cross-talk filter removes internal interference among the beams of the device to improve the ability to locate and observe objects. In the absence of this filter, the echo of an object sometimes bleeds across multiple beams. After the filter is applied, only the strongest echo location remains.

The original noise level image (left) to one modified by the software developed in this project (right), including some automated tracks ending at the current frame.

The second approach to improving signal-to-noise ratio was peak detection. By computing a temporal distribution of echo strengths for each range-beam cell within the sampled area, it was possible to create a threshold for echo strength below which lesser echoes could be ignored. Actively moving fish, however, create echoes with strengths well above the threshold unless there is a large stationary object in the field of view. Peaks are retained if their echo strength falls more than three standard deviations above the background level.

In addition to improving the ability of humans to observe fish, removing cross talk and identifying peaks facilitates the automated identification of fish behavior. To capture behavior, it is necessary to identify the peaks created by a single organism across multiple frames. To associate the echoes of each individual through time, an algorithm tracks strong reflectors (peaks) in the sonar data. The Kalman filter was chosen for the algorithm because it is an established method for tracking changes in state (in this case, position) of a random process (e.g., a fish swimming). The openCV library implementation of the Kalman filter was initialized with a definition of the process state vector, measurement matrix, and values for process and measurement noise. The filter uses inputs to predict the next position of each active track in a new frame and the uncertainty around that position. Detection closest to the track's predicted position was assigned if the difference between the detected and predicted positions was less than 15 pixels. This threshold was chosen based on trial and error; a future improvement would be to calculate a threshold based on expected size of the fish, process noise value, and data resolution. In addition, each track that accrues a minimum number of echoes is stored along with statistics that include origin, direction, speed, variability, and mean echo strength. These data allow further evaluation about whether individual tracks are pertinent and if behavior is consistent with expectations.

Further development is warranted to incorporate additional multi-beam imaging sonar device file types into the list of those able to be processed. Software application to real-world situations will likely reveal additional opportunities for improvement.

PN13060/2541

Laser Ablation Capillary Absorption Spectrometer for Trace Isotopic Sampling and Imaging

M. Lizabeth Alexander

The goal of this project is to advance laser ablation-capillary absorption spectrometer (LA-CAS) system development to perform stable isotope measurements at a level of precision and accuracy required for detection and imaging in key DOE research areas such as biogeochemistry, atmospheric sciences, and genomic sciences.

Precise measurements of stable isotope ratios are well established as a valuable tool for basic areas of chemistry, biology and geology as well as forensics and medicine applications. The current workhorse for stable isotope measurements of lighter elements (hydrogen, carbon, nitrogen, and sulfur) is the isotope ratio mass spectrometer (IRMS) coupled with traditional analytical techniques for sample preparation and introduction such as a gas chromatograph (GC), elemental analyzer (EA), or offline methods. In recent work, we reported the use of LA-IRMS as a method for spatially resolved isotope analysis of samples. These measurements showed precision and accuracy consistent with IRMS measurement but required a much smaller (~1000-fold) sample size than traditional EA-IRMS. Most recently, we used a prototype LA-CAS that combines previous LA-IRMS work and carbon isotope measurements by IR spectroscopy into a single instrument.

The development of the LA-CAS was motivated by the need to track nutrients in microbial communities at the scale of one to tens of organisms, or near 1 µ in size. Carbon isotope measurements provide one method of tracking nutrients from sources with differing $^{13}C/^{12}C$ ratios such as C3 vs. C4 plants. The amount of carbon in this small area is less than a picomole (pmol), far below the detection limits of IRMS (5–10 nanomoles [nmol]) or current IRIS systems (100–200 nmol). Our results further demonstrate that there is no loss in sensitivity when coupling LA to the CAS in a flow-through system ideal for the imaging application intended for LA-CAS. Previous results of CAS with the gaseous CO_2 standards system demonstrated molar detection of 100 pmol at levels of precision sufficient to distinguish isotope ratios from different nutrient sources without the need for expensive, time-consuming isotopic labeling. Further improvements lowered detection to 2 pmol, and improvements are underway that are expected to measure carbon isotopes at the 100–200 femotmolar level.

Spatially resolved isotopic measurements on this scale can find a wide range of applications. For instance, this approach could be utilized for performing detailed tracking of labeled substrates through a stratified microbial mat, probing the stable isotopes of microfossils embedded in a rock matrix, assist in performing individual analysis of prokaryote and eukaryote communities such as in the rhizosphere or microbiome. A key advantage of this system over previous spatially resolved work is the improved sensitivity over an IRMS-based method. The enhanced sensitivity of the LA-CAS system over LA-IRMS will permit coupling to smaller sample spot sizes, thus enhancing the spatial clarity of such measurements and permitting probing of smaller systems.

The real-time compensation for thermal drift and the software development for rapid summing and peak integration were successfully completed with the exception of auto-

Peak integration and isotope ratio calculation

PN13097/2578

mating the cursor control for peak selection and the data storage process. Results were achieved using a National Instruments system crate and software to average absorption signals of small sets of sweeps and then using software to realign any peak drift to the known line frequency center, followed by obtaining an average for the entire period.

Hardware improvements to reduce thermal and mechanical drift are now half complete. The increased flexibility of a new fiber material allowed it to be coiled around a form and thermally insulated. The remaining part of this task is to add active temperature control. An additional benefit realized with the fiber (which also has an improved tapered design) is a significant reduction in optical feedback, allowing a much smaller heat source to be used in the optical isolation stage and reducing problems with thermal stability. We additionally demonstrated the conductance limited interface that will allow imaging down to 25 µ. Ultimately, we anticipate that adding a cryogenic interface would achieve a resolution of 2 µ.

The identification and purchase of an appropriate QC laser became necessary when the existing Alpes Lasair laser for CO_2 became unstable. The frequency slew rate was nonlinear and not constant, resulting in large apparent shifts in frequency. This resulted is a dependence of output power vs. frequency, which translated into a large slope in the baseline that dictated the use of 2f detection resulting in bipolar peak shapes.

A new QC laser was purchased from Adtech that eliminated these problems and demonstrated much improved linearity in slew and reliability as well as access to a frequency region with higher CO_2 absorption strength that will translate to higher sensitivity. The significantly reduced slope in baseline allows 1f detection, simplifying data collection and improving sensitivity. An additional benefit is that Adtech is willing to make QC lasers for other frequencies needed to measure other light isotopes. This was a task projected for FY 2014 and is now 50% complete.

Millimeter-wave Shoe Scanner

Justin Fernandes

To eliminate the need for personnel to remove shoes before being screened in airport security, we are integrating an array into the floor of the current millimeter-wave advanced imaging technology portal.

According to the International Air Transport Association in a survey of 142 airports prior to September 11, 2001, about 350 people passed through security checkpoints every hour. Clothing divestment, especially footwear, is a significant impediment to efficient airport screening. Reduced throughput in security checkpoints increases inconvenience to passengers, ultimately leading to lost revenue for airports and lost productivity for the general economy. A footwear scanning solution would increase security scanning throughput via decreasing the load on operators scanning objects passing through carry-on luggage CT scanners. To our knowledge, there have been no efforts to develop imaging-based footwear screening, although there have been several attempts to create a capability that looks for chemical, magnetic, and low-frequency electromagnetic signatures. The objective of this project is to develop a millimeter-wave imaging based footwear screening technology that can be licensed to a third party for development and field deployment.

Initial work included performing an imaging system parameter optimization, which included system bandwidth, antenna polarization, and physical scanning geometry. The optimal single element system was used to scan multiple shoe types containing various types of threat simulants. An initial imaging database was constructed, and manual detection of abnormal metallic and non-metallic materials in a wide variety of shoes was performed. Algorithms leveraging the 3D nature of the data were developed and have been shown to increase contrast between nominal and anomalous shoes in many cases. New volumetric visualization techniques have been tested and used to increase ability to visually detect abnormalities embedded within shoes. Some of these methods use preprocessing techniques that format 3D point cloud data to surface meshes, which is the standard for the application of feature extraction algorithms based on surface characterization and intrinsic geometric distances.

During FY 2013, we made significant technical progress. The design and modeling of antennas radiating into a dielectric half-space was performed. The antenna design was chosen used based on simulated electromagnetic parameters. An impedance matching layer located in between antennas and footwear soles was researched, constructed, and tested. A hybrid fast Fourier transform (FFT)-based generalized synthetic focusing technique (GSAFT) and pure GSAFT imaging algorithms were developed and used to reconstruct data obtained from the new sparse array design. A half-space focusing algorithm was written to focus images generated by system located free space illuminating objects in a dielectric half-space. A real-time footwear imaging system has been created via integration of the final impedance matching configuration with a new millimeter-wave sparse imaging array. A high frequency structure simulator was used to model different antenna configurations radiating into a dielectric half-space. Based on these results, a final antenna configuration was determined and is being used to image footwear through an impedance matching layer.

Multiple impedance matching layers have been researched, constructed, and tested. Image volumes of multiple types of footwear with each impedance matching configuration have been obtained. The images were evaluated qualitatively to determine the optimal impedance matching configuration. The final configuration has been shown to improve image quality significantly and increase the detectability of multiple threat objects.

The final integration of a sparse millimeter-wave imaging array with the impedance matching layer has been performed. Due to the physical requirements of imaging within footwear, a hybrid focusing technique using GSAFT and FFT has been developed and implemented in imaging software. This algorithm reduces/eliminates artifacts due to the quasi-monostatic assumption violation in close range imaging applications. An additional half-space focusing algorithm has been developed and is being integrated into the software. Both of these focusing techniques will increase imaging capabilities. The final integrated footwear imaging system will be tested on actual humans wearing a diverse set of footwear. These results will test the effectiveness of the new imaging algorithms and impedance matching layer.

We have shown that an impedance matching layer significantly improves image quality in the footwear scanning application. With the results obtained with this effort, marketing is underway with a company to license this unique technology. The next step will be to create another impedance matching layer that would be resilient to a realistic footwear scanning scenario.

PN13079/2560

Miniaturization of Multi-Modal Regenerative Feedback Sensor

A. Mark Jones

The objective of this project is to develop and miniaturize a class of multi-modal sensing platforms based upon regenerative feedback resonator circuits. It is envisioned that this technology could be deployed as unattended wireless sensors for national security applications.

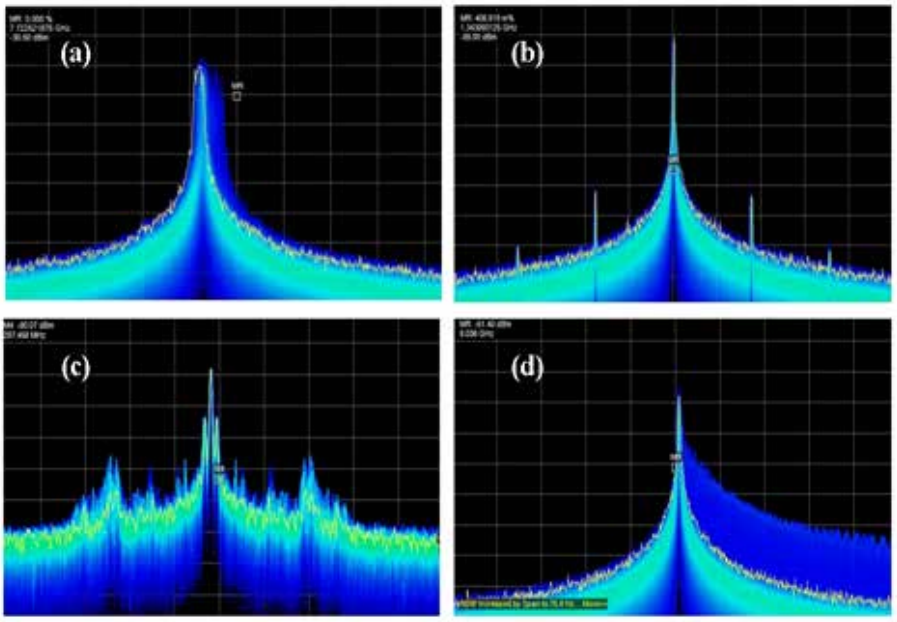
This project was initiated because it was discovered that material-loaded regenerative feedback resonator circuits under evaluation as radiation detectors exhibited unique responses to other phenomena during characterization. The circuits were originally intended to sense transient changes in detector material electromagnetic properties induced by ionizing radiation. However, prototype configurations demonstrated surprising sensitivity to a variety of other stimuli, including acoustic pressure, human proximity and motion, visible light, electromagnetic fields, and temperature. Recognizing the potential for multi-signal detection, we began to explore the use of these circuits in the broader context of national security missions.

The simplest implementation of the circuit being investigated consists of a microwave amplifier with a material-loaded cavity resonator in the feedback path and a directional coupler to measure the resonant behavior of the loop system. Other components may include a phase shifter for precise frequency control and a band-pass filter to reject undesired resonant modes. The sensing mechanism is produced by interaction of the resonator with the external environment. The circuit is compatible for use with any two-port resonator element, and we have utilized 3D and planar split-ring resonators as well as resonant patch antennas.

The regenerative feedback technique has been widely used to increase sensitivity and frequency selectivity of radio receivers. These circuits are attractive because of their simple, low-cost topology and ability to generate dynamic quality factors many orders of magnitude higher than those quality factors available from resonators used in a passive configuration. For example, we have measured dynamic quality factors as large as 15 million for material-loaded circuit prototypes. The sensing circuit can be compact and designed to transmit wirelessly real-time frequency response to a receiver located a safe distance away from the deployed sensor.

Our first step was to perform a baseline evaluation to assess the response of early prototype sensors to acoustic, electromagnetic, and proximal inputs. For each modality, we identified relevant sensor characterization methods to collect the appropriate standardized data. After reviewing the results, we selected human proximity and electromagnetic emissions as the modalities for detailed research and design efforts. It is expected that investments in these sensing configurations can be readily leveraged for other modalities.

We made significant progress in understanding and miniaturizing the multi-modal sensor in FY 2013. Using a combination of 3D finite element electromagnetic simulations and experimental investigations, we examined sensor phenomenology, including resonator modal behavior, dielectric material loading effects, and transduction mechanisms. Excellent agreement was obtained between simulated and measured characteristics of the various



Example instantaneous frequency response of multi-modal sensor prototype configurations captured using a real-time spectrum analyzer for various inputs: (a) carrier frequency shift due to human proximity, (b) sideband structure due to electromagnetic emissions, (c) frequency modulation due to human voice, and (d) broadband frequency shift due to camera flashbulb.

PN13031/2512

resonator configurations. Additionally, research revealed that the proximal sensor modality operates as a short-range radar transceiver, where the resonator functions as a reciprocal antenna that transmits and receives propagating waves. The phase shifts corresponding to reflected signals from nearby objects are converted into excursions from the nominal carrier frequency. In addition to proximity sensing, direct detection of incident electromagnetic signals is enabled by the self-mixing operation of the amplifier as it is driven into its nonlinear regime. The received signals appear as frequency-modulated sidebands on the carrier frequency. We performed a literature review of active integrated antennas and self-mixing oscillator antennas to understand resonator radiation characteristics and phase-noise stability for these related configurations.

The sensor form factor was reduced from bulky initial prototypes consisting of coaxial components to a small printed circuit board and resonant patch antenna operating near 2 GHz. This more compact design exhibited significantly improved performance for both proximity and electromagnetic inputs when tested using the characterization methods from the baseline evaluation. In addition, we confirmed that the sensor can detect human presence and movement through internal walls of typical office buildings. We performed a preliminary demonstration of the use of multiple identical sensors operating in wireless exfiltration mode with external directional antennas to detect and track movement through an area. This simultaneous operation was possible because the sensor was designed to enable precise tuning of the frequency response of the regenerative feedback circuit.

After confirming the successful operation of the 2 GHz printed circuit sensor, we designed the next iteration of the multi-modal sensor. To miniaturize the form factor and maintain performance, the operating frequency was increased to 5.8 GHz and the resonant patch antenna and feedback circuit were placed on opposing sides of a multilayer circuit board. The overall dimensions of the 5.8 GHz sensor are 50 mm × 50 mm × 2 mm. The resonator element was designed using electromagnetic simulation.

Our efforts on this project during FY 2013 resulted in the publication of a peer-reviewed journal article in *Review of Scientific Instruments* on the use of this technique for detecting transient changes in electromagnetic properties of semi-insulating and magnetic materials. We will pursue additional technical publications during the second year of the project. Additionally, the technology from this project relates to patent pending IPID 30048-E.

For FY 2014, we plan to fabricate and test the 5.8 GHz miniaturized multi-modal sensor. In addition to performance characterization, we will examine further the operation of multiple resonators in concert to monitor an area of interest, which will require determining the optimum approach to integrate a directional antenna for wireless transmission of the real-time signal. Other planned research areas include investigating options for low-power sensor operation, such as power harvesting and energy storage, devising methods to separate desired signatures from undesired and environmental background effects, and exploring the use of non-radiating resonators for applications in measurement science.

Mössbauer Spectral Imaging

Lucas E. Sweet

Actinide Mössbauer spectral imaging is a novel concept targeted at improving the limit of detection and chemical characterization of actinide-containing solids. This development has the potential to be a new tool in the field of radioactive material detection.

Mössbauer spectroscopy has been an underutilized chemical characterization technique for a couple of reasons. First, the radiation sources required to probe actinide nuclear resonances have been difficult to prepare and required significant radiological material handling infrastructure. Second, this has been a bulk solid analysis technique that required up to several days for a single spectrum to be collected on a sample. At these sample quantities and data collection times, other analytical techniques are more effective for chemical characterization.

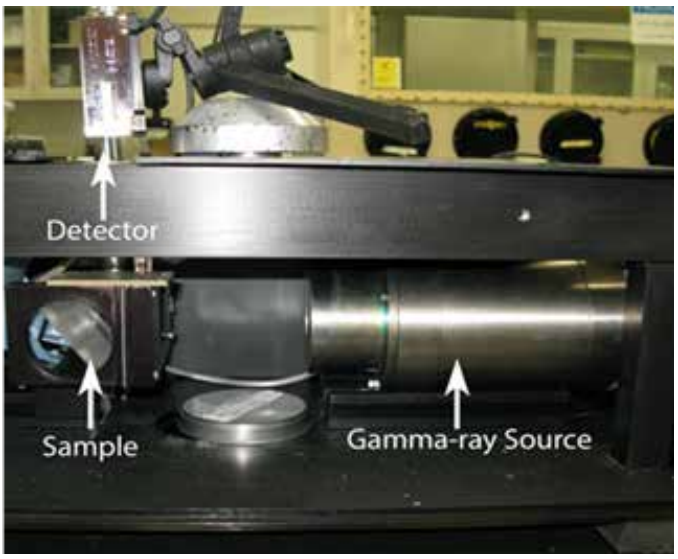
Recent advances in iron (^{57}Fe) Mössbauer spectroscopy have drastically improved the detection limits and data collection efficiency. These advances include the use of gamma detectors that have high spatial resolution, gamma-ray focusing optics, and technologies for Mössbauer measurements using high flux synchrotron light sources. Our focus is to apply these new advances in iron to actinide Mössbauer spectroscopy. These developments are intended to produce a novel ultra-sensitive characterization technique for actinide-containing solids that can have applications in nuclear safeguards and nuclear non-proliferation.

During FY 2013, we purchased the typical components used in Mössbauer spectrometers, made modifications, and explored alternative measurement geometries aimed for improvements in detection limits and decreases in data collection times. One of the key and most significant steps this year was to evaluate options for gamma detectors in this instrument. The typical gas proportional detectors used for ^{57}Fe Mössbauer have poor counting efficiencies for the energy of the gamma rays analyzed in actinide Mössbauer measurements. After carefully evaluating the efficacy, speed, and energy resolution of our detector options, we chose a CdTe detector, which was then integrated into the spectrometer originally designed for a specific gas proportional counter. We confirmed that this new detector configuration would allow for the high efficiency detection of the correct gamma-ray lines for performing all of the actinide isotope Mössbauer measurements that we are targeting.

After the CdTe detector was integrated, we explored alternative measurement geometries hoping to add additional efficiency and improve system detection limits. We used a computer model to test several different measurement geometries and compared real data collection efficiencies of some of these geometries using our spectrometer. Because our Mössbauer spectrometer was set up for transmission mode measurement geometry, we reconfigured the system to perform measurements in a backscattering geometry that improves detection limits for small particulate sample types, even though this is less efficient for bulk powder samples.

We made progress on several other fronts. A sample holder was made to analyze radioactive samples in a contained, safe manner, along with custom adaptors to incorporate a variety of gamma-sources necessary for different actinide Mössbauer measurements. A cryostat was custom designed for low temperature measurements essential for actinide Mössbauer. We also tested ways to increase the signal to noise ratio using lock-in-amplification algorithms.

For FY 2014, we plan to complete the Mössbauer spectrometer that is capable to measuring nuclear resonances of ^{237}Np and then add the capability for ^{238}U nuclear resonances. Able to measure multiple isotopes simultaneously or serially, we will also gain the ability to perform isotopic ratio quantification, which will be a new use for this technique. Looking further ahead, we plan to add the imaging component through the combination of detectors with high spatial resolution and gamma-ray focusing optics.



The current state of the actinide Mössbauer spectrometer being developed has a CdTe detector and backscattering geometry to improve the limits of detection and data collection efficiency.

PN13032/2513

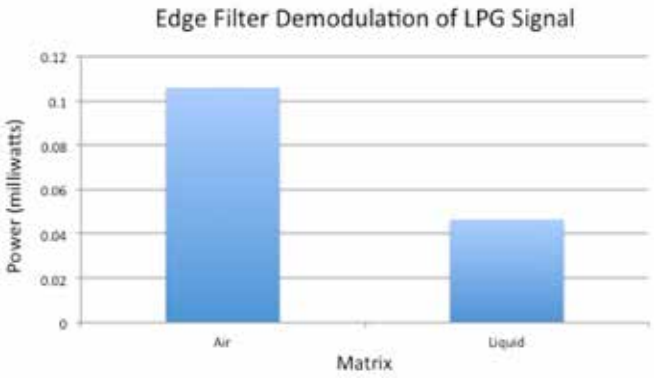
Optical Fiber Sensing System for Process and Health Monitoring (Stain and Strikes)

Mark E. Jones

This project investigated optical fiber sensors toward the goal of improving the composite manufacturing process specifically used to produce blades but also applicable to manufacturing nacelles and hubs.

To enable offshore wind efforts and gain acceptance for renewable offshore projects, the cost competitiveness of power generation needs to be demonstrated. While many factors influence cost, a large part of manufacturing expense is tied to scrap materials that do not meet specification, and single- component failure to meet specification is a significant issue. Composites are primary materials used in offshore structures because they meet size, operational environment, and lifetime/servicing constraints. Traditionally, these large composite parts are made through a molding process where epoxy or polyester is injected into molds that contain glass or carbon fibers. After wetting, the parts are cured. A primary cost and/or performance issue results where voids form in areas where resins fail to flow. By inserting optical fiber sensors, these fibers inherently become a part of the composite structure.

For this project, our research team performed small-scale experiments on how embedded optical fibers can be used without modification through the lifecycle of the components to determine strain for health monitoring and performance-driven maintenance. With no modifications and minor changes to the demodulation algorithms, high frequency point sensors can be created to monitor for bird strikes. We will use the results of this project to develop a comprehensive optical fiber approach to measure and improve the manufacturing process of turbine blades and to demonstrate sensor and fiber interrogation methods that span multiple frequency regimes. New demodulation concepts will allow spatial evaluation of strain loads and high frequency monitoring of bird strikes.



Measured intensity through a spectral edge filter with the sensor in air and water.

During FY 2013, we determined the optimal grating specifications using MATLAB modeling approaches. The sensitivity that an optical fiber has toward the surrounding materials is dependent on how much of the light is able to extend out of the fiber. This condition is determined by the type of optical fiber used, the wavelength of light, and the grating period. We designed gratings using commercial communications-grade fibers that leverage low cost due to a high volume production market. Based on design parameters, we identified COTS vendors able to produce the gratings, and we built an interrogation scheme to measure performance.

Sensor being interrogated using an optical spectral analyzer (OSA).

Another accomplishment was demonstrating two interrogation techniques. Our team illuminated gratings with white light sources and used a spectrum analyzer to measure spectral changes with fiber wetting. We proved that edge filters can be used for a high frequency, intensity-based signal interrogator. Specifically, the team used a scanning narrowband laser to filter the grating signal along its linear sideband. This configuration can be used to measure high frequency events (such as strikes) that occur in composite structures.

Because of delays in grating fabrication and receipt of epoxy resins, the in-structure tests were restricted. We did, however, perform limited composite layups for resin transfer molded parts. Fiber placement techniques were leveraged from other active projects, limiting the optimization of grating placement. Resins were identified that will interact most optimally with the fiber cure grating-based sensors.

PN13059/2540

Platform for High-Throughput Determination of Enzyme Kinetic Parameters for Hemicellulose Saccharification

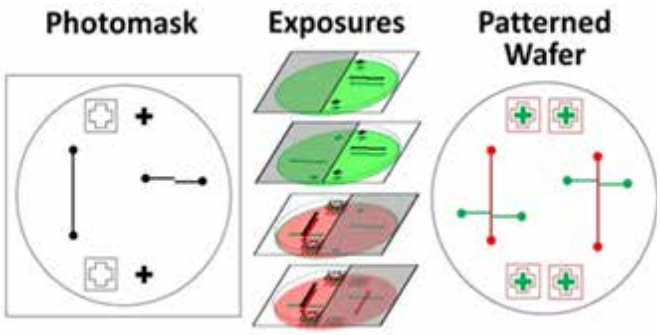
Ryan T. Kelly

We are developing an analytical platform based on droplet microfluidics coupled with ion mobility spectrometry/mass spectrometry (IMS/MS) to monitor the reactions involved in the enzymatic breakdown of biofuel precursors.

With increasing efforts to develop viable biofuel alternatives to fossil fuel energy sources has come the need to understand the fundamental enzymatic reactions involved in natural lignocellulose degrading systems as well as to screen rapidly recombinant libraries of engineered glycoside hydrolases for feedstock degradation efficiency. Existing methods to monitor enzymatic saccharification are low in throughput, often involving liquid chromatography separations or fluorescent labeling, and require large amounts of material. We are developing a platform to monitor degradation efficiency with much higher throughput and with minute amounts of reagents. The platform will enable lignocellulose degradation to be measured in a label-free, high throughput fashion, providing improved fundamental understanding and a means of optimizing processes relevant to biofuel production.

The platform is based on droplet microfluidics coupled with IMS/MS. Picoliter to nanoliter sized volumes comprising enzyme, substrate, and dilution buffer are controllably dispensed into an immiscible oil stream. Rapid mixing of enzyme and substrate droplets upon contact initiates the degradation process and the reaction takes place as droplets travel through the microchannels towards the detector. We aim to demonstrate the ability to monitor reaction of multiple (hemi)cellulases with a panel of cellodextrins with high throughput and to detect the products in a label-free fashion using IMS/MS. This will be the first time that droplet-based microfluidics have been combined with IMS/MS, which should prove powerful for high throughput screening and online reaction monitoring.

In FY 2013, we laid the groundwork for the data collection and analysis that will be carried out in the second and final year of this project. The nature of the project requires that a number of iterations in microfluidic device design be carried out. These iterations have been needed to produce properly sized droplets, achieve rapid mixing, ensure adequate incubation time, and optimize the extraction of droplets from the oil stream for IMS/MS analysis. Each successive generation required multiple photomasks to accommodate different microchannel heights and cross-sectional shapes. The production of three to four photomasks, which each take a day on average to produce, for every design change became a major impediment to optimizing the microfluidic platform. We solved this problem by devising a novel strategy enabling designs that would normally require up to five photomasks to be produced using a single photomask, while still maintaining the same alignment tolerances between layers achieved using separate photomasks. The results of this innovation have been accepted for publication in the international journal RSC Advances.



Schematic depiction of the production of aligned multilayer microfluidic devices from a single photomask. This development enables us to develop and evaluate devices much more rapidly.

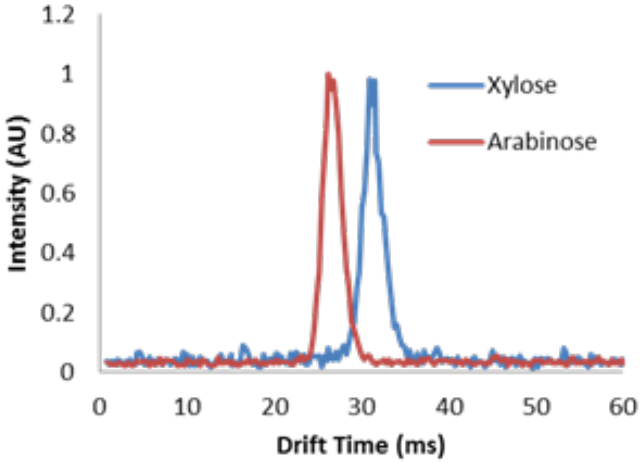
In terms of microfluidic design, we have made several advances. First, the microchannel surface needs to be hydrophobic to allow aqueous droplets to travel smoothly without adhering to the walls. Our initial protocol of simply baking the devices overnight to recover the PDMS hydrophobicity provided inconsistent results. We found that adding a small amount of surfactant (0.75% perfluorooctanol) greatly enhances device stability and minimizes droplet adhesion to the walls, while still enabling droplets to merge upon contact for mixing operations.

Another issue we encountered was conventional analog pressure regulators, which we use to control the pressure with which droplets are introduced into the oil stream, tended to drift over time and with changing input pressures. We purchased a digital pressure controller that quickly responds to such changes and allows for consistent droplet generation performance over extended periods of time (hours). Finally, we explored different channel geome-

PN13074/2555

tries to enhance mixing and found that increasing the tortuosity of the flow path while varying the cross-sectional area indeed increases mixing efficiency. With the above developments, we made great strides toward finalizing device design, which we expect to complete before the end of the calendar year.

We have concurrently focused on optimizing electrospray IMS/MS conditions for the detection of sugars. Sugars are more challenging to analyze by electrospray ionization (ESI) than peptides, for example, due to their hydrophilicity a lack of readily chargeable sites. We explored different buffer conditions (i.e., ammonium acetate with and without added sodium chloride) and concentrations for a variety of monosaccharides and oligosaccharides and found that the majority were readily detectable with strong signal to noise. We determined that ammonium acetate buffer without added sodium chloride generally provides slightly stronger signal than when sodium chloride is added and, as such, salt additives will be avoided moving forward. We also found a very small increase in signal for increasing concentrations of saccharides from in the range of 10 µM to 200 µM, indicating that even 10 µM is beyond the linear dynamic range of the IMS/MS platform. This finding has important implications for this work, as it sets a fundamental limit on the range of steady state kinetic parameters that can be determined using our proposed approach. The reactions under study in the present work have K_M values in the micromolar to millimolar range, indicating that they will be challenging to measure by IMS/MS. On the other hand, we found that IMS was able to differentiate clearly between sugars having identical masses (i.e., xylose and arabinose), which will be important in analyzing complex mixtures.



Normalized ion mobility spectra showing that sugars having identical masses are easily resolved by IMS that would be indistinguishable by MS alone.

Given the inherent challenges that we face in determining kinetic parameters in the micromolar to millimolar range, we will instead develop our platform for FY 2014 as a tool for comparing degradation efficiency among glycoside hydrolases and mixtures of glycoside hydrolases rather than determining fundamental rate constants. This will provide significant benefit, as large numbers of genetically engineered cellulase variants can be generated but methods to characterize them in terms of yield and stability in harsh environments are low in throughput. Additionally, we plan to complete microfluidic platform development and demonstrate the online reaction monitoring of cellodextrin degradation.

Prototype Fast Neutron Detector

Sean C. Stave

We are developing a high-rate capable, non-volatile fast neutron detector that may have a role in the future of active interrogation of special nuclear materials.

This project is developing a stand-alone tool for the characterization of neutron sources. The detector being developed maximizes the different time scales for neutron and gamma ray interactions in layers of plastic scintillator, a non-volatile material that can be used anywhere, and uses precision timing to pull the neutron signal out of gamma ray noise. Previous efforts used large, single pieces of plastic scintillator and focused on very fast readout of the light. Our modeling efforts showed that approach to be inadequate due to pile up of multiple events and the sheer size of the detector smearing out the light signal.

The modeling and testing performed in FY 2013 indicated that the multiple slab design would generate sufficient light and maintain the appropriate timing resolution to separate the neutrons from gamma rays. All components for a 20-slab module were assembled and tested. Two main challenges were addressed. The first was to decide upon a configuration for mounting the scintillator slabs. The mount needed to reduce the amount of material around the scintillators (fewer places for neutrons to downscatter without being detected), maintain a precise geometry, and allow for flexibility in placement of the slabs. The original design at the end of FY 2012 had placed the slabs in close proximity to one another, but a study at the beginning of FY 2013 showed the advantage in efficiency and gamma rejection ratio (GRR) that would be gained by separating the slabs more. The final mount design had fixed slot locations using the longer distance between slabs. Even greater separations can be studied with the design by removing slabs (e.g., every other slab can be removed to roughly double the slab spacing).

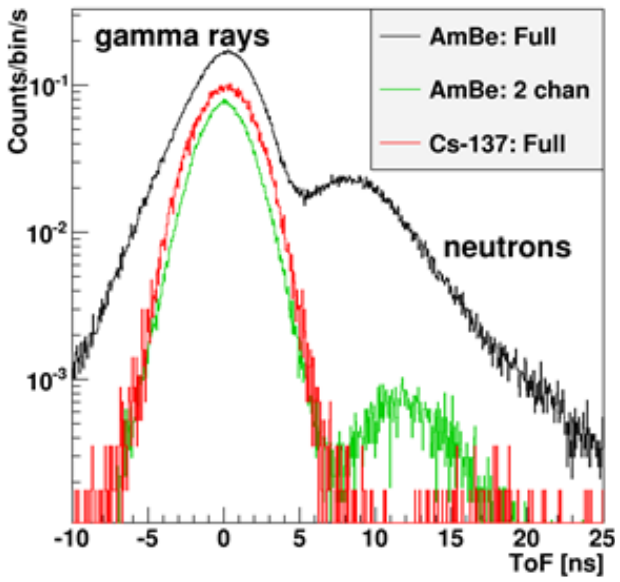
View of the PoC detector with custom circuit board on top.

The second major design challenge was the mount and associated readout electronics for the silicon photomultipliers. This challenge was addressed in parallel with the scintillator mount design and resulted in a compact, low noise package that interfaced well with the scintillator mounts. Each channel had an individual high voltage control for hardware gain matching and a special sync pulse was included to allow sub-nanosecond timing resolution between any two channels in the data acquisition system. Additionally, every pulse from each scintillator slab was digitized and stored. A key part of reducing the data rate and improving the live time of the system was implementing a hardware coincidence trigger, which could be set to acquire single events or exclude singles and record only events that are in coincidence within a particular time window.

The limiting factor for the number of detector elements in the proof-of-concept (PoC) detector was the cost per channel of digitizer. Nanosecond scale timing is critical for this application and required significant spending for hardware and design work. It was demonstrated that the additional channels of readout contributed more than an order of magnitude to the sensitivity of the detector. However, there was a tradeoff between efficiency and GRR. Liquid scintillators typically have a GRR of 10000 and a neutron detection efficiency near 20%. The PoC detector achieved a best GRR of 2250 with an efficiency of 0.08%. Higher efficiencies were achieved, but with the cost of reduced GRR.

The simulations indicate that the efficiency can be increased to 2% by collecting more light. However, the noise in each channel must be kept to a minimum to maintain a low accidental coincidence rate that dilutes the GRR. It is also possible that additional geometry information could be used as part of a neutron/gamma ray identification algorithm. In the first part of FY 2013, this algorithm work was advanced as well. The conclusion was that through timing and geometric cuts, there may be an additional factor of 10 that could be achieved in the GRR. Certainly, laboratory testing showed that the distributions of distances between coincident events for neutron and gamma ray induced events are different, so there should be additional selective power available.

PN12064/2465



Example time-of-flight spectrum illustrating the full PoC detector response to neutrons compared with the two-slab detector from the previous year, and the response of the full detector to a gamma-ray source.

Two additional features of detector performance were discovered through the testing process. The first is that the detector exhibits neutron direction sensitivity due to the preservation of the relative timing between hits. The second feature is that there is sensitivity to the neutron energy spectrum. A single liquid scintillator detector has no directional sensitivity and minimal neutron energy sensitivity. The segmented nature of the PoC detector allows for the single-scattered neutron energies to be reconstructed into an energy spectrum that can be used to identify the neutron source.

In the end, the PoC detector showed the validity of the method and the various choices of hardware. The detector performed within a factor of 2 of the simulated efficiency and was able to achieve a GRR that was about 25% of the goal. Various improvements have been identified with the collection of more light being the most promising. The detector can be scaled up to increase efficiency, but to be affordable will require fast, cheap digitizers that should be available in the near future. More light will also open the possibility of using the energy deposited in each slab to separate further the neutrons from the gamma rays.

Reactor Materials Degradation During Plastic Deformation and Creep: High-resolution Experiments and Integrated Models of Microstructural Evolution and Materials Response

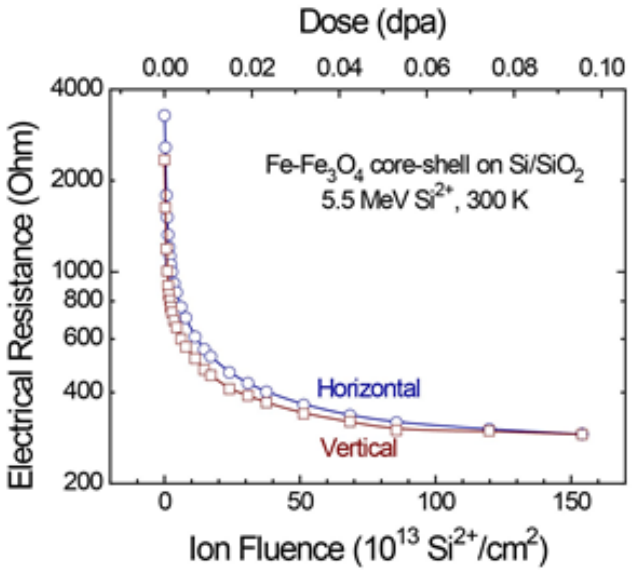
Weilin Jiang

This project explores the feasibility of in-situ nuclear radiation monitoring with emphasis on the performance and understanding of sensor materials.

The operation and new designs of nuclear reactors for power generation demand the highest safety standards. The development of an *in-situ* monitoring system will significantly enhance safety by providing "real-time" degradation data for structural materials, leading to a rational prediction for their residual operating life. The goal of this project is to understand the performance of sensor and reactor materials to support development of continuous online monitoring.

***In-situ* study**. Granular nanostructured films were studied *in situ* using an ion accelerator and a helium ion microscope (HIM). The results show that under 5.5 MeV Si^{2+} ion irradiation at 300 K, there is a super-exponential decay of the electrical resistance of Fe-Fe_3O_4 core-shell and Fe_3O_4 magnetite granular films. In the low-dose regime, the material is extremely sensitive to ion irradiation: even at a low dose He^+ ion irradiation at 300 K, there is a decrease in the resistance by ~15%. This behavior provides a potential for high-sensitivity detection of nuclear radiation. Resistance decreases more gradually in the higher-dose regime, allowing for a long-term nuclear monitoring. Compared to 300 K, irradiation of core-shell granular films at 473 K leads to a more dramatic decrease.

In-situ HIM study of microstructural evolution under 25 keV He^+ ion irradiation shows that upon irradiation, loosely interconnected nanoparticles tend to aggregate. With increasing dose, crystallites grow at the expense of the surrounding amorphous material and smaller crystalline particles through irradiation-induced interfacial expitaxy and coalescence, respectively. X-ray diffraction (XRD) for irradiated *in-situ* samples indicates a significant increase in crystallite size. Larger particles are observed to interconnect and gradually form a nanowire-like structure. Phase transition from shell Fe_3O_4 to FeO also occurs.



In-situ study of electrical resistance of an Fe-Fe_3O_4 core-shell granular film on Si/SiO_2 using the van der Pauw method as a function of ion fluence or dose in dpa during ion irradiation at room temperature. A super-exponential decay behavior of the electrical resistance is observed. In low-dose, the material is extremely sensitive to ion irradiation, providing a potential for high sensitivity nuclear radiation detection. The resistance decreases more gradually in a higher dose regime, allowing for a long-term nuclear monitoring.

Separate thermal annealing and *in-situ* electron irradiation experiments suggest that both electronic and thermal processes can hardly alter the initial microstructure of the granular films. Nuclear collision is concluded to be primarily responsible for activation of the nanostructural evolution process. The result indicates that the granular films are likely to be susceptible to neutron irradiation. Despite their inability in activating nanostructural evolution, both thermal and electronic processes can play a significant role once the structural evolution starts. A manuscript based on the *in-situ* data is currently under preparation for publication.

***Ex-situ* study**. A number of materials have been studied *ex situ*. It was found that the particle size increases in Fe-Fe_3O_4 core-shell granular films after ion irradiation. The irradiation also reduces the oxide shell to lower-valence Fe. The magnetization curves of double perovskite

PN13053/2534

La_2BMnO_6 (B = Ni and Co) nanoparticles show paramagnetic-ferromagnetic transitions at T_C ~275 and 220 K for La_2NiMnO_6 (LNMO) and La_2CoMnO_6 (LCMO) nanoparticles, respectively. The results from the films of $Ni_{1-x}Co_xO_y$, where x=0.50, 0.67, and 0.75, indicate that ion irradiation induces a loss of oxygen, a reduction of nickel, and an increase in low-temperature ferrimagnetism.

Fifty-seven materials were considered for candidate sensor materials and down-selected to five. One of the candidates, $BaTiO_3$ polycrystalline films (~400 nm thick), was grown on platinized silicon using pulsed laser deposition, and subsequent XRD characterization confirmed the crystal structure. The electron backscatter diffraction (EBSD) images and the domains determined from magnetic force microscopy of HT-9 steels are obtained and correlated.

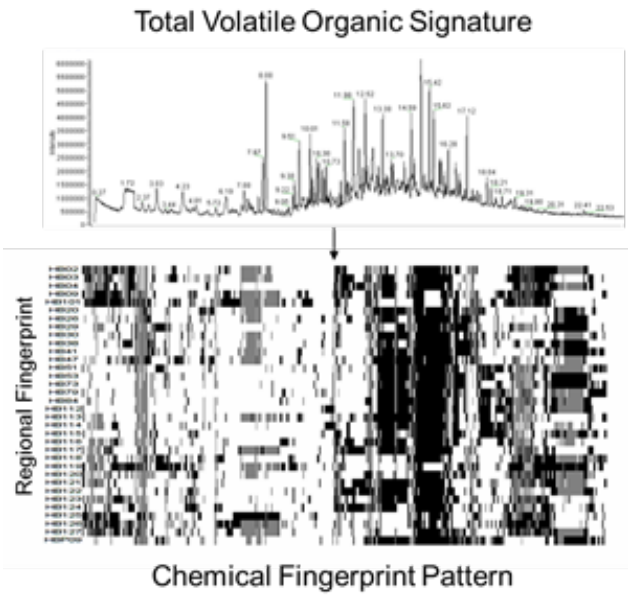
Modeling and simulation. FeCr alloy is taken as a model system to demonstrate the model capability in studying the effect of Cr precipitates on magnetic response. The results show that both the nonmagnetic particle and 360° Bloch domain wall are nucleation sites of an anti-direction during domain switching, and the latter is the easier nucleation site. The existence of nonmagnetic particles could suppress the domain wall formation and cause magnetic hardening. The nonmagnetic particle morphology has great impact on magnetic domain switching and coercive field. Spherical particles result in largest coercive field comparing to the same volume ellipsoidal or cylindrical particles. The frequency of applied magnetic fields also affects the magnetic hysteresis loops. The lower loading rate makes smaller coercive field. Because the model accounts for the microstructure, local fields around defects, long-range fields (e.g., stress), and electric and magnetic fields, the simulation is useful to design radiation sensor materials.

Secondary Signatures for Provenance Attribution

Jon H. Wahl

This project investigates "natural" (secondary) tagging and tracking signatures present in the atmosphere for correlating nuclear process signature detections to a specific source. Our research will improve the U.S. government's ability to associate the nuclear signature materials in a given sample with the environs of or the effluents from a target of interest.

Many nuclear measurements are highly dependent on meteorological data and modeling input, which may be limited, unavailable, or in many cases unreliable. Consequently, additional provenance information would be valuable to improve and bolster confidence. The overall objective of this project is to develop a robust methodology to determine signatures or features that inherently exist in the environmental background (i.e., features of opportunity), which can be volatile organic compounds (VOCs), particles (in the form of soot, dust, or pollen), and microbial cells (earth microbiome). Importantly, a certain fraction of all these types of backgrounds will be due to transport/dispersion to a region. Identifying and utilizing this fraction, which will likely be a feature set of the background, can offer an indication of the directional region or origin of an air sample.



Schematic representation of an extracted chemical fingerprint from a gas chromatogram to various regional chemical fingerprints.

One of the most readily available chemical background features are VOCs, the collection and analysis of which are at the heart of our chemical approach. Our technical approach is based on operational use of proven technologies and analytical methods, including large volume air sampling and concentration utilizing sorbents, sample extraction, and analysis by gas chromatography/mass spectrometry (GC/MS). Modern day GC/MS is the "gold standard" for volatile compound analysis. In general, a robust volatile compound map needs to be screened because neither the type, number, or amounts of components are known *a priori*, although certain chemical classes (e.g., hydrocarbons) may be anticipated. As such, GC/MS is used to examine specific chemical characteristics within a variety of environmentally collected samples to determine the statistically significant signature components. This chemical and environmental data, as well as any other additional material such as regional information of known industrial activity or endogenous backgrounds, will be mined for "features of opportunity."

The chemical organic methods for provenance signatures have been developed, and the signature statistical framework has been established. Specifically, a robust air collection and chemical analysis approach has been examined utilizing a thermal desorption unit GC/MS system. In addition, a similar approach utilizing an electron capture detector was examined briefly.

A statistically derived method for denoising complex GC/MS data files was developed to analyze a collection study that contained numerous GC/MS generated chromatograms. In general, this approach is based on the assumption that well-separated, higher intensity compounds are the most significant and reproducible. Additionally, this method evaluates each chromatographic peak in terms of its combined retention time, *m/z* value, and intensity. A probability or significance level for the peak's location and intensity is calculated. If the combined significance level is greater than a pre-determined threshold, then both the location and intensity are typical of background peaks, and the peak is eliminated from consideration. If it falls below the significance threshold, then either the location or intensity or both are unique, and it is retained as a signature peak for the compound.

For FY 2014, we will continue to assess those signatures measurable by GC/MS based upon sampling scenarios, and we will assess the likelihood of those signatures being present in the concentrations needed for detection. Depending on the outcome and findings of the above assessments, a series of field trial will be conducted.

PN13024/2505

Super Resolution Fluorescence Barcoding and High Throughput Microfluidics for Quantifying the Expression of Multiple Genes in Individual Intact Cells

Galya Orr

With this research, the ability to interrogate individual cells in complex communities accelerates the discovery and design of new molecular functions for sustainable biofuels and carbon storage.

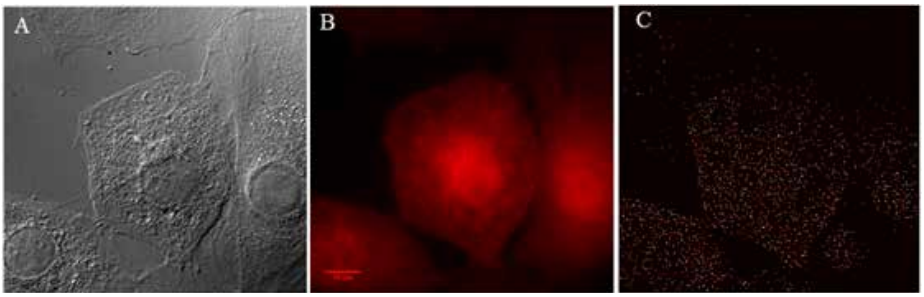
Fundamental limitations in averaged cell population measurements, often dominated by the most abundant molecules or processes together with recent technological advances, have led to new developments of single-cell analysis. These techniques can identify variability between cells rapidly and accurately unmask molecular details buried otherwise in population measurements. However, these methods are available at few expert laboratories and mostly in low throughput.

This project provides an automated high throughput platform for large-scale applications in biodefense or medical care. Fluorescence *in situ* hybridization (FISH) will be used, where fluorescent probes are designed to recognize and bind mRNA molecules for specific genes. While the number of fluorescent colors that can be differentiated within a single cell is limited to about seven, the combinatorial use of these colors to create fluorescent barcodes can significantly increase the number of probes that can be differentiated in a single cell to more than 60. Our newly developed technology will provide a powerful tool for rapid, accurate detection and interrogation of single pathogens or diseased cells for immediate diagnosis and treatment, and support research in environmental microbial communities, where complex relationships make it impossible to regrow isolated cells in culture.

We established combinatorial barcoding of FISH probes and their detection and quantification within intact individual cells using super resolution fluorescence microscopy. Building on our expertise in molecular biology, we established protocols and optimal probe configurations for generating barcoded FISH probes and inserted them into the cells. We have been using eukaryotic cells that are larger and easier to work with and will focus on bacterial cells later in FY 2014. Building on our expertise in single-molecule fluorescence imaging techniques in living cells, we assembled a multicolor stochastic optical reconstruction microscopy (STORM) system, which currently includes four lasers. By adding a fifth laser, it will be possible to detect nearly 30 different barcoded FISH probes in a single intact cell. By systematically increasing the number of barcoded FISH probes imaged in a single cell, we can work our way through the technical and computational challenges in order to achieve unambiguous quantitative evaluation of the expression of multiple genes in individual intact cells.

In addition to our above progress, we established the trapping of single cells in a microfluidic array for subsequent high-throughput tagging and imaging. The array is fabricated using multilayer soft lithography and consists of multiple inlets, each has a corresponding microvalve to control fluid delivery. The cells flow into the array and are confined in geometrical traps designed to accommodate just one cell each. Once the traps are filled, microvalves close around each trapped cell to hold the cells in place for imaging. In our current array, approximately 75% of the cell traps contained a single cell held in a precise location for subsequent imaging, while the remaining 25% contained more than 1 cell or were unfilled. Further refinement of this process will increase the fraction of traps containing a single cell to exceed 90%.

For FY 2014, we will establish the trapping of single cells in a microfluidic density array and produce delivery of all reagents needed for preparation and fluorescence labeling in an automated fashion. Additionally, we will scale up the numbers of detectable barcoded mRNA probes and trapped cells in the microfluidic array, combining the two systems.



(A) A differential interference contrast (DIC) image; (B) a wide field fluorescence image; and (C) a super resolution STORM image of cells tagged with FISH probes specific for GAPDH, a key enzyme in the conversion of glucose to energy.

PN13087/2568

Ultra Low Noise Detectors: Building an Advanced Capability

Michael E. Wright

This project combines the latest digital hardware with the latest physics analysis software into a field deployable platform capable of handling the highest data rates and large data sets – all in real time.

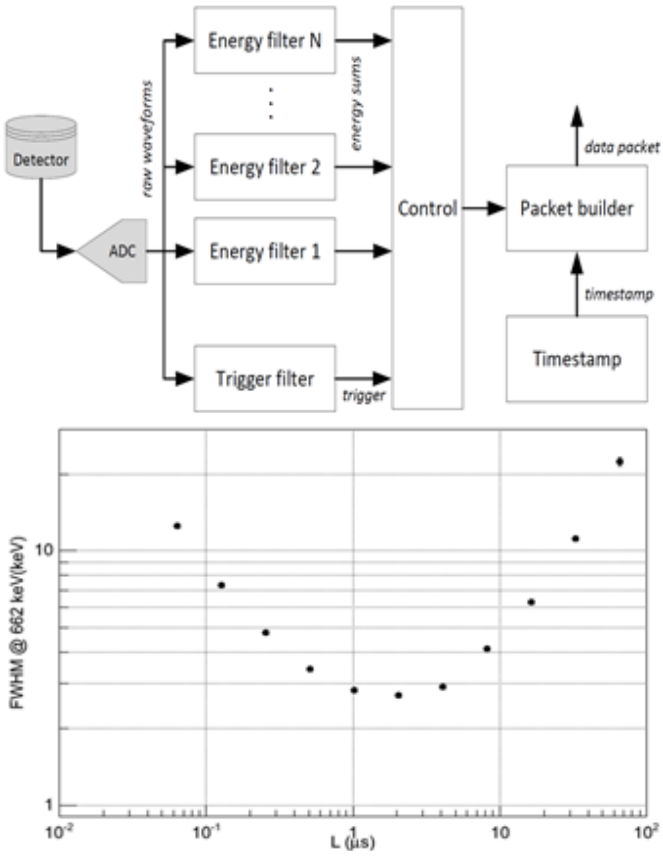
Air, land, and sea-going radiation detectors serve a range of important missions from monitoring radioactive shipments at ports to mapping dispersed radiation in the environment. However, high speed radiation detector electronics compatible with the high purity cryogenic germanium crystal radiation detectors have been severely limited to the speed of their legacy computer using a two decade-old personal device electrical communication standard in peripheral component interconnect. Instruments being developed require far higher data rates than this outdated hardware in its dotage can support, yet nothing faster that could support germanium precision deployed was available.

In this project, today's physics analysis software has been ported into hardware on a 3 billion transistor integrated circuit, transmitting data over the VPX computer data bus, achieving roughly a factor of a million times speed-up, depending on the algorithm. One of the goals of this project was to build and demonstrate a system that was capable of data rates required in finding radiation sources in real time, as opposed to days or even months later than the existing post-processing. Another goal was to build a system that could be easily field deployed, enabling searching or mapping of large geographical areas much faster than previously possible.

Initially, the project required finding an existing computer solution; there was no time or budget to develop new digital hardware. The solution would have to be creative and well considered, as it would be setting a new path forward for what is expected to be decades to come. Specifically, the OpenVPX computer platform was selected. Commercial and military suppliers created OpenVPX for military applications, including high speed high precision analog to digital conversion and real time data processing. Physics codes for post-processing data require a near total re-write when ported to a digital hardware description language for implementation in configurable digital hardware – each step had to evaluate in less than 4 ns. Eight parallel trapezoidal digital filters, temporal selection, event sensing, time stamping, scratch data memories, and storage of measured time stamp and particle energy to non-volatile memory was demonstrated.

The new platform was connected to a high purity germanium radiation/particle detector. Hundreds of thousands of events were analyzed, and excellent energy resolution of 2.4kEV at 662 keV was accomplished. Final steps included demonstrating better energy resolution using differential signaling from the radiation detector to the digitizer. Ultimately, the goals of developing and demonstrating a suitable path forward to far higher data rates and germanium-compatible energy resolution in a field-deployable system, with a per-channel cost target of only a couple thousand dollars, was achieved.

With this system, vastly more data can be collected and analyzed than before, promising to create far more complete pictures of radiation for a wide range of scientific programs. Our ultra-sensitive multi-channel radiation detector with read-out capability supports fundamental physics discovery, ultra-sensitive measurement, environmental science, and nuclear security applications.



Top: A block diagram of the multi-rate filter system implemented in configurable hardware. **Bottom**: A multi-rate system energy spectrum.

PN13055/2536

Biological Sciences

Analysis of Functional Potential from Metagenome Data

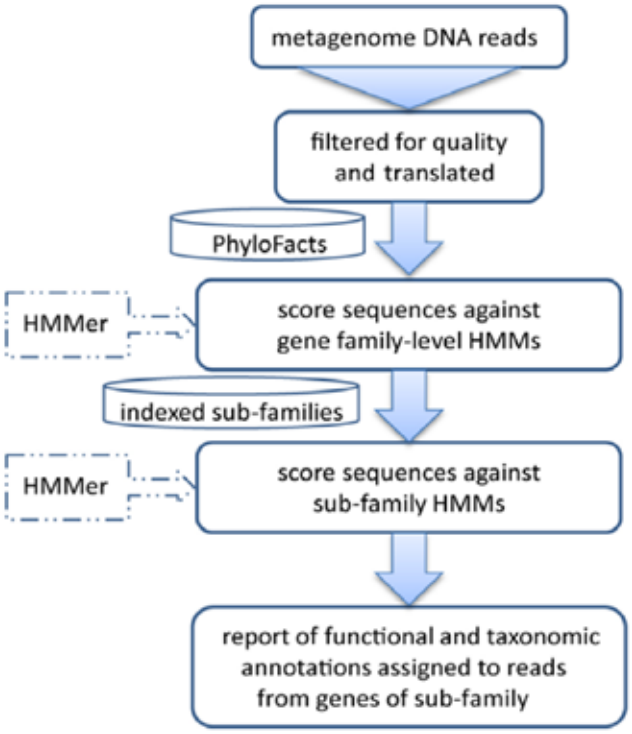
Lee Ann McCue

In recent years, the technology to generate DNA sequence data has out-paced its analytical tool development. Novel computational methods to analyze the large quantity of generated sequence data are needed, and will be essential for DNA-based studies designed to decipher the functional role of microbial communities in the environment.

Microbes exist and function in communities – studies of natural environmental or human-associated sites often reveal hundreds to thousands of microbial species co-existing. Microbe interactions within such a community are often complex, as is the impact the community as a whole has on its environment. Thus, deciphering microbial community activities in their natural environments is a prerequisite to understanding their functional role in environmental processes. Studies of natural microbial communities have increasingly maximized the advanced technologies now available to sequence DNA in a high throughput manner. It is now feasible to sequence extracted DNA from natural microbial communities consisting of many species. Such samples consist of a composite mixture of genomes of all the community members (i.e., a metagenome). While some progress has been made using metagenomic data to infer the presence and abundance of microbial taxonomic groups, inferring their functional contribution to the community has proven challenging.

This research developed a computational analysis approach for analyzing individual (unassembled) DNA sequence reads generated by current high-throughput sequencing technologies to provide information on the taxonomic composition as well as functional potential of the DNA source. This approach utilized several established computational tools, combined with a database developed by our collaborator, Dr. Kimmen Sjölander. The advantage of this approach is its ability to provide functional and broad-level taxonomic information for a metagenome in a more accurate and less biased manner than is currently possible using existing computational methods.

The PhyloFacts database provided by Dr. Sjölander consisted of 69,409 functional gene families, each represented by a multiple sequence alignment, a hidden Markov model (HMM), and a phylogenetic tree. For this project, we expanded and indexed the database by generating multiple sequence alignments and HMMs representing the data located below every branch of the phylogenetic tree for each gene family. We have further specifically indexed each family as representing a local protein domain (~1/3 of the families), or a full-length protein (~2/3 of the families).



Flowchart of the computational analysis approach for metagenome data. Existing software tools are shown in hashed boxes.

We have implemented a multi-step analytic pipeline that 1) reads a DNA sequence file in fastq format; 2) trims the sequences for quality; 3) optionally filters the sequences for matches to rRNA gene sequences; 4) translates the DNA to protein sequences; 5) searches the PhyloFacts HMMs (all 69,409 or the local or global subset) for a matching gene family; 6) searches the indexed sub-family HMMs to identify a specific subsection of the family that best matches the query; and 7) writes a report of the results for every query sequence match to a gene family. The report provides general metrics about the input, such as DNA sequence quality and length of the translated protein, as well as information on the query sequences' match to a PhyloFacts gene family. The report predicts a taxonomic origin for the sequence by identifying the least common ancestor of the genes within the branch of the phylogenetic tree that yielded the best HMM match. The report also provides the UniProt gene functional annotations from those genes.

The analysis pipeline has been implemented on the PNNL Institutional Computing System. Analysis of validation datasets for known microbial sequence data has shown high sensitivity (99%) and positive predictive value (> 85%).

PN12008/2409

Characterization of Signaling Networks in Single Cells

Bryan E. Linggi

Our experiments characterize signaling networks in single cells using microfluidic systems to decrease sample input and increase molecular measurement resolution. Characterization and enumeration of signaling network variability "states" will greatly advance our understanding of cellular responses and community behavior.

Cell communities exhibit complex responses to changes in their environment. Studying associated molecular changes (e.g., RNA and protein) is likely to improve predicting the behavior of these communities and engineer them for human benefit. However, traditional assays that measure changes at the population level obscure the underlying cell-to-cell variability that exist even in the homogenous population. By understanding differences in signaling at the cell (rather than population) level, we can begin to unravel the apparent heterogeneity in cell response (so called "rare-events") that mediate many biological processes. As the mechanistic basis for cell responses is limited by the inability to measure molecular changes at the single cell level, there is a critical need to measure molecular changes at this level to understand the inherent variability among cells and appropriately link them to cellular behavior.

In FYs 2011 and 2012, we optimized multiple approaches to capture and analyze single cells using microscopy and microfluidic devices. We created several microfluidic designs that were typsinized, loaded into devices, and captured using a variety of different "capture wall" geometries. Additionally, we designed devices that allowed for precise control of two-fluid interfaces, increasing the throughput of the system. Finally, we created four different clonal populations (single cell derived) cell lines for future use and allowed for dual color imaging for quantitating cell cycle and transcription factor activity in a single cell.

We continued our analysis efforts to capture individual cells and analyze their phenotype, protein, and RNA expression patterns. At the beginning of FY 2013, we were developing a method using polydimethylsiloxane (PDMS) membranes with evenly spaced holes of 20, 30, 40, or 50 µm in diameter placed on polystyrene plates and exposed to oxygen plasma for 30 sec. We found that we could retain cells on exposed areas, but regions at which the membrane did not have direct plate (wrinkles in membrane) contact were exposed to the oxygen plasma and were therefore hydrophilic. These areas bound many cells and created a heterogenous cell environment. After literature review and scientist consultations, we were not able to overcome this limitation and concluded that this method was not viable for our goals.

As an alternate approach that had potential to overcome the PDMS membrane limitations, we explored the application of UV patterning. Unfortunately, the UV chamber malfunctioned during this step and required off-site repair. While we made progress on the method, given the remaining project time available (~7 months) and the uncertainty of when the chamber could be repaired, we changed approaches to a goal that could be accomplished by the end of the funding period. We did learn that cell viability is extremely low when cells are isolated, and current available technologies (PDMS stamping, stencils, UV patterning) are not robust enough for our application as a tool to limit cell heterogeneity. While some areas of a microscopic field were sufficient quality, there were other cell regions that were either unhealthy or contained clusters. Extensive development will be required to advance these approaches to meet standards required for this application.

To conclude this project, we are defining the relationship between single cell protein localization and gene expression. We will focus on two proteins, ERK1/2 (ERK) and SMAD3, both which have well defined changes in localization, moving from cytoplasm to nucleus upon activation. The change in localization can be visualized at a single cell resolution using antibodies to total activated ERK and SMAD3. There is also a crosstalk between the amount of nuclear SMAD3 and the ERK activation state: active ERK phosphorylates the linker region of SMAD3 leading to its export from the nucleus. However, others have demonstrated a requirement for ERK activity for functional SMAD3, in a cell context dependent manner. We performed preliminary experiments and optimized the visualization of ERK and SMAD3 in human mammary epithelial cells (not shown).

The second component of this project is gene expression analysis under diverse environmental conditions. We performed preliminary experiments to identify genes that are activated or repressed in cells with activated ERK (caused by exposure to EGF) and activated SMAD3 (caused by exposure to activated TGF-β). We identified genes that are regulated by EGF/ERK (IL-8) and T (p57), and identified antagonism between ERK and SMAD at the level of gene expression. Inhibitors to either ERK (inhibitor of MEK, U0126) or SMAD3 demonstrate that ERK has a negative effect on SMAD3 activity, but this is overcome by exogenous EGF. We are in the process of exploring how these combinations of activators and inhibitors effect localization of ERK and SMAD3. If changes in localization do not explain the changes in nuclear localization, it would suggest that alterations in the activity of nuclear localized ERK or SMAD3 as a mechanism of crosstalk.

PN11011/2333

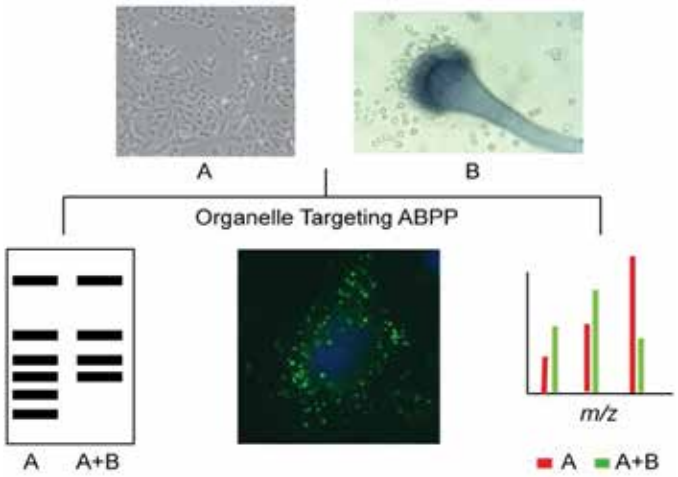
Combining Proteomic Technologies to Create a Platform for Spatiotemporal Enzyme Activity Profiling

Susan D. Wiedner

We are combining proteomic technologies of activity-based protein profiling (ABPP) and subcellular proteomics. Individually, these two technologies study different proteome "fractions" but together would provide spatiotemporal information of the functionally active proteome.

ABPP and organellar proteomics are techniques used in proteomics to decrease the complexity of the natural cellular proteome, which often contains thousands of proteins. ABPP isolates the proteome by tagging functionally active enzymes with chemical activity-based probes. In contrast, organellar proteomics reduces complexity by focusing on proteome subcellular compartments. Combining ABPP and organellar proteomics into the spatiotemporal enzyme activity profile (STEAP) will provide information about the functionality of active proteome on the organelle level. On a cellular level, this information would help elucidate fundamental biological processes such as protein translocation, signaling, and cellular response to environmental stress. The STEAP platform will be used to compare relevant proteomes such as diseased systems.

In FYs 2010–2012, we applied ABPP to the pathogenic fungus *Aspergillus fumigatus* grown in the presence and absence of human serum. We found that the activity-based probe (ABP) reactive proteome of *A. fumigatus* significantly responds to human protein. The three fungi compared share almost 90% of their genomes, including known virulence factors; however, ABP-reactive virulence proteins were measured only in the pathogenic organism. Thus, ABPP provided a catalog of ABP-reactive proteins in response to an environmental change compared across multiple conditions and organisms, thereby providing valuable information about the underlying biology and improving ABPP capability.



STEAP application of an organelle targeting ABP to stressed (A+B) and non-stressed (A) organisms, enabling ABPP of proteins within native subcellular environment. The multi-modal organelle targeting probe allows analysis of subcellular ABP reactive proteins by gel electrophoresis, fluorescence microscopy, and mass spectrometry.

The prior years' enhancements became the foundation for FY 2013 project work, where we shifted from studying *A. fumigatus* to host-pathogen interactions. Proteomic analysis of host-pathogen interactions entails identifying pathogen proteins in the milieu of host proteins. To reduce proteome complexity, we used STEAP to analyze a specific subcellular proteome. The lysosome and acidic organelles is one such subcellular location that plays a key role in microbial defense by macromolecule and protein degradation: *A. fumigatus* evades cellular defense by residing in acidic organelles where it can germinate and escape, leading to infection. We developed acidic organelle targeting ABPs to accumulate and tag proteins within acidic organelles.

Fluorescent microscopy of cells treated with probe, fixed, and conjugated to fluorescent tags showed that the presence of weakly basic amine allowed rapid accumulation in punctate vesicles within a mammalian cell. We synthesized two probes with different enzyme reactive electrophiles and both were able to accumulate in specific subcellular regions. A multi-modal probe design allowed analysis of tagged proteins by SDS-PAGE as well as mass spectrometry. With comparative analysis optimized for ABP-reactive proteomes, we applied our organelle accumulating probes to a perturbed and unperturbed mammalian cell culture. Mass spectrometry analysis showed the probe electrophiles covalently bound to different enzymes, with only modest overlap. ABPs provided a readout of ABP-protein reactivity dependent on culture condition. Our cell-permeable, targeting, multi-modal probes combined organelle proteomics and ABPP into a STEAP platform for probing proteins in subcellular native environments.

With further optimization, such as different electrophiles for a greater number of enzyme families and incorporation of different organelle targeting moieties, the STEAP platform can be applied to a variety of biologically relevant systems.

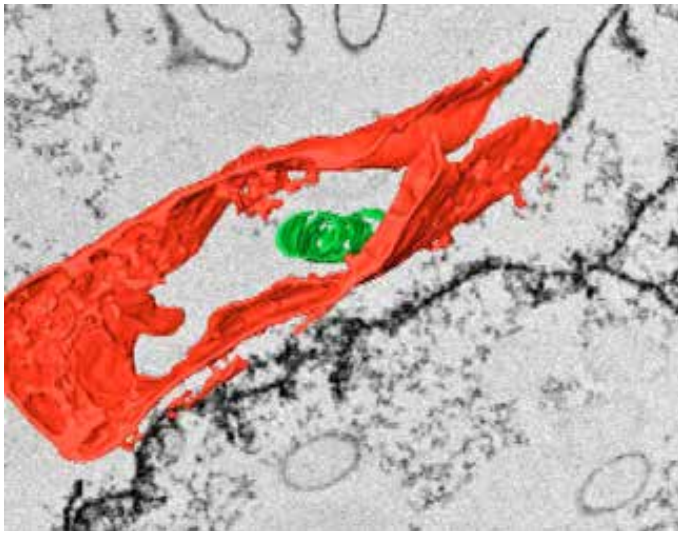
PN10011/2253

Correlative High Resolution Imaging and Spectroscopy to Characterize the Structure and Biogeochemical Function of Microbial Biofilms

Matthew J. Marshall

This research develops correlated capabilities for state-of-science imaging, compositional analysis, and functional characterization of microbial biofilms. It also seeks to understand biofilm influences on biogeochemical processes, such as the fate and transport of radionuclides in the subsurface of carbon sequestration.

In both natural and engineered environments, the majority of microorganisms live in structured biofilm communities. In addition to microbial cells, biofilms are composed of a poorly characterized organic matrix commonly referred to as extracellular polymeric substance (EPS) that may play roles in facilitating microbial interactions and biogeochemical reactions, including extracellular electron transfer (EET). Understanding how biofilm EPS interacts with inorganic substrates (i.e., metal ions and mineral surfaces) connects the molecular-scale biogeochemical processes to those at the microorganism-level and provides insight into how microorganisms influence larger, pore-scale processes. Using a multi-faceted, multi-scale approach of imaging and analytical techniques to elicit the complex microbial and biogeochemical interactions, biofilm communities will be surveyed in its nearest-to-native state. With several techniques, our project will provide detailed, high-resolution visualizations and compositional data of biofilms in their nearest-to-native state as they influence local biogeochemical reactions in their environment.



Amira 3D reconstruction of *Shewanella* cell and EPS matrix associated with UO_2. The cell can be seen in close proximity with UO_2-EPS.

In FY 2011, our perceptions of the size of nascent reduced uranium (UO_2) nanoparticles were limited by the high-resolution electron microscopy (EM). The addition of aberration-corrected Titan scanning/transmission electron microscopy-enabled atomic-scale discovery of biogenic UO_2 nanoparticles and the observation of nanoparticles smaller than 1 nm. For FY 2012, we generated highly reproducible biofilms for imaging studies and chemical analysis, as *Shewanella* cultures were grown using anaerobic minimal medium containing lactate and fumarate in a constant-depth (bio)film fermenter (CDFF). The amount of chemical information collected during EM analysis was limited and complicated by the addition of chemicals for EM fixation, dehydration, staining, and embedding.

Correlative soft X-ray-based imaging techniques presented an excellent opportunity to gain chemical and/or elemental information from ultrathin biofilm sections at relevant spatial resolutions. We developed a new cryosample preparation technique that omits traditional sample processing chemicals that produce data interference. EM analysis confirmed that excellent depth-resolved biofilm morphology and cell ultrastructures were well preserved. Samples were imaged by scanning transmission X-ray microscopy (STXM), X-ray fluorescence micro-imaging (μXRF), or nano-secondary ion mass spectroscopy (nanoSIMS) to obtain nm-scale chemical information that could be correlated to high resolution EM images. We studied carbon chemistry using STXM and found several differences between the cell surface and EPS encompassing biofilm cells. This high-resolution visualization of biofilm chemistry will help determine how biofilms influence local biogeochemical reactions.

Phase contrast imaging can provide an enhanced contrast of low-Z materials, and some methods also offered texture-based contrast modes. We used three physically distinct signatures—absorption, phase, and scatter—for imaging biofilms on complex matrixes. Preliminary results were obtained on the characteristics of several phase contrast methods: absorption/propagation-based, Fourier X-ray scattering radiography (single grating), and Talbot interferometry (double grating with phase stepping)

PN11015/2337

with tomography data acquisition. Our findings suggested that phase contrast enhances the visibility of unstained hydrated biofilms (relative to absorption contrast), while scatter provides little additional information on hydrated samples. Using osmium as a contrast agent enhances visibility for absorption-based imaging. For relatively thin biofilms (~50 µm or less thick), spatial resolution loss from the use of a 2-µm fringe pattern were weighed against the ability to image without a contrast agent.

The segmentation of reconstructed slices visualized pores (or channels) in the hydrated biofilm, which indicated regions where water, nutrients, or electrolytes can flow through a biofilm and may influence pore-scale biogeochemical reactions such as contaminant fate and transport. We significantly narrowed the technology gap between EM and X-ray capabilities. Our experiments facilitated the initial construction of both µm and nm scale biofilm models.

During FY 2013, we refined our X-ray microtomography- and microscopy-based biofilm models produced through FY 2012. We completed segmentation across a ~250 µm thick [z-direction] sample of reconstructed biofilm and constructed a detailed visualization of the channels that facilitate fluid perfusion through the hydrated biomass. Using segmented fluid perfusion channels and the biofilm external surfaces, we are now computing the distribution of biomass as a function of distance from the nearest point of nutrient advection (channel or external fluid interface). This distribution will demonstrate the relative importance of nutrient advection and nutrient diffusion for sustenance across the biomass. This dataset facilitated our construction of a micrometer scale 3D biofilm model and provided a community-scale visualizations for how cells within a microbial community can thrive in a complex structure such as a biofilm. Understanding how water, nutrients, or electrolytes flow across/through a biofilm provides important insight towards how microbial activities can influence larger-scale biogeochemical processes such as the cycling of multivalent metals, biocorrosion, or biofouling.

We improved our gratings-based phase contrast imaging by applying an improved reconstruction software to process projection images. This allowed us to image fully hydrated, unstained biofilm samples by Talbot interferometry and visualize the reconstructed images on site at the 2-BM beamline. Rapid visualization permitted for *ad hoc* sample modifications to more efficiently collect data. A third method, phase retrieval for propagation-based phase contrast, was tested this year. In this method, the detector is moved downstream of the sample, allowing edge enhancements due to refractive effects to be more clearly visualized. Following data acquisition, phase retrieval is performed. A material composition is assumed and used along with experimental parameters to infer the local densities required to produce the observed effects. This results in an image with slightly reduced spatial resolution but much greater sensitivity to low-Z materials and variations in bulk densities, and well-suited to segmentation for further processing.

To study biofilms further at the µm and nm scale, we used µXRF to image biofilms grown in a semi-porous bead matrix that mimics highly heterogeneous subsurface microenvironments. In these studies, exogenous metal electron acceptors were added to the growth medium to study metal localization across the bead matrix and within the biofilm. Initial µXRF images failed to show biofilm co-localized with exogenous metals; however, OsO4-staining successfully facilitated biofilm imaging within bead pores. Importantly, biofilm observation at this scale will reveal how microbial activity in the subsurface may influence pore-scale biogeochemical reactions such as contaminant fate and transport or the precipitation of carbon.

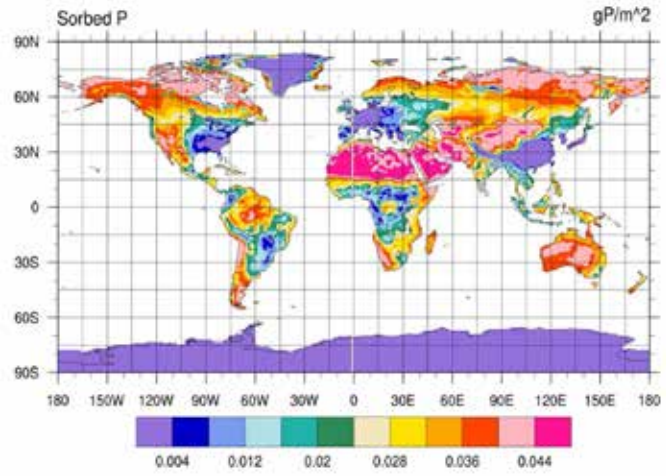
Developing A Next Generation Biogeochemical Module for Earth System Models

Yilin Fang

This research will bridge the gap between soil biogeochemical research and earth system modeling research so that the representation of terrestrial ecosystem processes in land models can be developed to improve the quality of climate model projections.

Biogeochemical processes regulate soil carbon dynamics and CO_2 flux to and from the atmosphere, influencing global climate changes. Integration of biogeochemical processes into the earth system (e.g., community land models [CLM]) currently faces three major challenges: 1) requires extensive effort to modify modeling structures and rewrite computer programs to incorporate biogeochemical processes with increasing complexity, and these efforts must be repeated to integrate new knowledge; 2) endures prohibitively expensive computational cost to spin-up CLM to steady state of coupled carbon-nitrogen processes due to the long residence time of soil carbon; and 3) existence of various mathematical representations of biogeochemical processes, some of which are too simplified to reflect fundamental mechanisms while others are too complex for pragmatic application. Thus far, systematic evaluation of different mathematical representations has not been performed in terms of their impact to climate prediction. The object of this research is to develop a generic biogeochemistry module to address the above challenges that will enable significant future modeling research with reduced cost and minimize coding error. The model can facilitate the entire community to test different mechanistic process representations and data to gain new insight on what is most important in the system in response to climate change.

During the first 6 months of the project, we developed a new biogeochemical module that has a computational framework to incorporate mechanistic biogeochemical reactions readily into Earth system models for simulating carbon-climate feedbacks. The new module uses a generic algorithm in coding and a reaction database to collect and compile biogeochemical reactions. The advantage of the new module is that new and updated processes can be incorporated into any Earth system models without the need to set up manually mathematical equations for describing biogeochemical reactions. The compiled reaction database consists of nutrient flow processes through the terrestrial ecosystems in plants, litter ,and soil. The database also has a structure to incorporate updated reactions. The module was first integrated in CLM, the land component of the Community Earth System Model (CESM), and benchmarked against simulations from the original biogeochemistry code in CLM. We incorporated a phosphorus cycle in CLM using the new module. The phosphorus-incorporated CLM is a new model that can be used to simulate phosphorus limitation on the productivity of terrestrial ecosystems. A journal article has been published in *Geoscientific Model Development Discussion* and is under open access peer review.



Sorbed P distribution with weathering rate of 0.05 g P m^{-2} $year^{-1}$.

For the first time, the developed biogeochemical module allows Earth system models to have a generic computational framework to simulate biogeochemical processes mechanistically. With the generic algorithm structure and reaction database, the new module has significant potential to be adopted widely by the climate modeling community for advancing Earth system model development, yielding broad impacts in community modeling and earth system research. The module will enable significant modeling research to be performed efficiently that is otherwise costly and error-prone. For example, the module framework developed in this project will facilitate effective comparison studies of biogeochemical cycles in an ecosystem using different conceptual models under the same land modeling framework, allowing structural uncertainties related to biogeochemistry models to be assessed with minimal coding efforts. In addition, our reaction database

PN13007/2488

approach provides a mutually understandable venue to communicate with biogeochemists for model process representation improvement and even inspire new research.

We are currently working on the second task of this project to develop a new approach and algorithms to accelerate the computation of biogeochemical processes including coupled carbon and nitrogen cycling in the Earth system models. This work is proceeding well, and a manuscript is in preparation.

In FY 2014, we will continue to expand the database for new and updated mathematical biogeochemical models. Specifically, we will work toward identifying a site with rich measurements of microbial biomass and community composition information through the databases of the North American Carbon Program Site Synthesis project. We will also start evaluating different mathematical biogeochemical models and their system-level response in CLM.

Development of Bromotyrosine Antibody Assay for Blood and Sputum Samples

Richard C. Zangar

This project seeks a quantitative empirical test for assessing inflammation in asthmatics.

Asthma is a common disease that currently afflicts about 9% of the United States (approximately 23 million Americans) and world populations with a lifetime risk of about 15% and substantially decreases worker productivity and quality of life. Asthma is particularly problematic in children ages 1 to 4, such that 0.24% of these are hospitalized as a result of this disease, a rate that is higher than any other cause except respiratory infections. Along with no medical cure, asthma symptoms may disappear on their own for unknown reasons. The inability to cure this condition means that treatment of this disease is designed to control symptoms. Currently, there is no practical test for assessing inflammation in asthma patients, which means that this information is inferred from patient questionnaires regarding asthma symptoms. However, patients have differing perceptions of similar symptoms, so questionnaires are not truly reliable.

We propose to address the need for assessing asthmatic patients by indirectly assessing the eosinophil peroxidase (EPO) activity. Eosinophils appears to be the sole source of EPO, and this activity seems the sole physiological source of bromine bleach (HOBr or hypobromous acid). HOBr reacts with proteins to produce bromotyrosine residues, which appear to be specific markers for inflammation in asthma patients. Notably, a variety of studies have demonstrated that total levels of protein bromotyrosine are altered in asthma and are promising biomarkers for this disease. Even so, these studies did not evaluate bromination on individual proteins, such as we can now accomplish, and the previous analytical protocols are not well suited for a routine clinical screening.

We are developing a novel antibody that detects a specific protein modification and allowed us to assess airway inflammation in asthmatics. The goal of this project is to assess further the study of the halotyrosine antibody to determine its ability to characterize asthma and thereby guide clinicians in the identification and treatment of this disease. We therefore examined sputum (airway fluid) samples from two sets of subjects. One sample set will allow us to characterize biomarker levels in asthmatics over time and the other which will further assess differences between asthmatics and healthy controls. We also analyzed blood samples from asthmatics and controls.

Obtaining the human samples for this research required a substantial effort. Transfer of both types of samples involved approvals and agreements, which took a considerable amount of time. After receiving both sets of blood and sputum samples, we performed and completed analysis. Specifically, we analyzed halogenation of 13 proteins in each sample. Currently, the data from the blood samples are being analyzed by the statistician. Preliminary examination results suggested that there were no major increases in protein halogenation in these samples. The sputum samples were blinded, however, and we have yet to obtain the unblinded results from those samples.

For FY 2014, we will continue working to confirm the relationship between halogenated proteins and asthma severity, as assessed by exercise challenge. In addition, we will develop a brominated antigen for quantitative analysis for the most promising brominated proteins identified during the previous year. For future studies, we hope to extend our analysis to children, who are particularly susceptible to the adverse effects of asthma but are particularly difficult to diagnose using current methods.

PN13082/2563

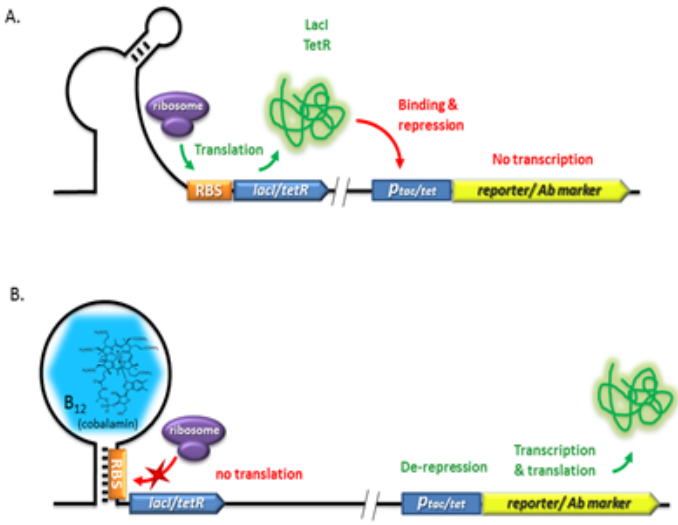
Directed Strain Evolution through Riboswitch-Controlled Regulatory Circuit

Alex Beliaev

We are exploring state-of-the-art metabolic engineering approaches for scalable and cost-efficient biofuel and bioproduct production through development of new synthetic biology tools that will drastically accelerate the metabolic engineering process using principles of directed evolution.

Current metabolic engineering strategies involve modification of specific targets (reactions, pathways) that lead to significant improvements in product yield and productivity. As this approach requires a clear understanding of the organism's genetic and metabolic makeup, its success is often confounded by unanticipated regulatory features, feedbacks, or interactions. In contrast to direct engineering, strain evolution does not require a priori knowledge of the system and takes advantage of genetic diversity and evolutionary forces. With few exceptions, however, selective pressure rarely enhances product yields, as natural selection forces target functions essential for survival.

We addressed the above challenge in this project by using riboswitch-controlled synthetic circuits that transform cellular input into a signal that triggers expression of growth-essential function. By placing genes essential for survival under the control of these regulatory circuits and applying selective pressure, the system will yield strains that have evolved to overproduce the riboswitch-specific metabolite. The proof-of-principle work proposed here will exploit the applicability of riboswitch-controlled circuits for directed strain evolution and product yield optimization. The expected outcomes of our work include engineering of synthetic regulatory circuits; developing approaches and toolkits for directed strain evolution; and conceptually developing metabolic engineering and synthetic biology strategies to apply to the design of novel platforms for cost-effective biofuel production processes.



Development of riboswitch-controlled cobalamine-dependent synthetic regulatory circuit. (A) In the absence of cobalamine, the LacI/TetR protein shuts down gene expression by binding to tac/tetO promoter. (B) When cobalamine is present, the sequestration of ribosome binding site (RBS) blocks the translation of LacI or TetR, subsequently derepresssing the tac- or tetO-driven gene expression.

Starting in mid-year FY 2012 and continuing in FY 2013, we made significant strides toward developing synthetic regulatory modules that could modulate gene expression in a ligand-dependent manner. Specifically, we selected the naturally occurring riboswitches of *Shewanella oneidensis* MR-1 involved in cobalamin (vitamin B_{12}) and lysine biosynthesis. The riboswitch control elements were inserted upstream of DNA sequences encoding transcriptional repressors, LacI or TetR. These constructs were subsequently incorporated into a two-component circuit to control the expression of reporter (GFP, YFP) or antibiotic marker (kanamycin or gentamycin) cassettes in a ligand-dependent manner. The latter was achieved through translational attenuation of LacI and TetR by cobalamine or lysine, either of which positively affect the expression of the output module (reporter or antibiotic markers). Additionally, we modified the phenotype of the host organism (i.e., *S. oneidensis MR-1*) to start testing the dynamics of the constructed circuit function and evaluating its performance based on the response curve quantitatively linking ligand (cobalamin, lysine) concentration and protein levels.

Both B12 and LYS controlled circuits were tested and modified to display a broad dynamic range, low background response levels and specific sensitivity in both *Escherichia coli* and *S. oneidensis MR-1*. The availability of riboswitch-controlled genetic device to confer antibiotic resistance only in the presence of specific ligand provides us with new ways for direct strain engineering. It also provides foundation for future efforts to utilize and strengthen PNNL expertise in computational biology, molecular modeling, and combinatorial chemistry. The range of applications involving synthetic riboswitch-based regulatory circuits include biofuels production, biological monitoring, and national security.

PN12025/2426

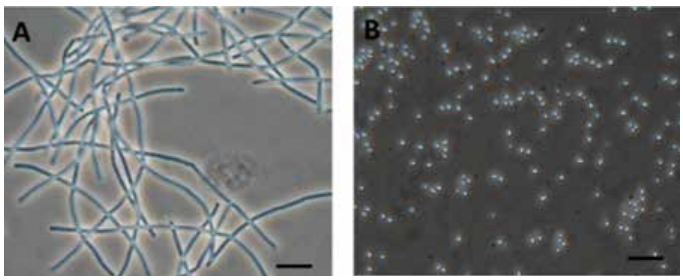
Elucidating the Natural Spore Phenotype

Marvin G. Warner

The project uses PNNL expertise in antibody evaluation to develop optimal methods for spore capture and retrieval from complex environmental samples (i.e., soil and blood soil mixtures) prior to proteomic analysis. In doing this, we will enable direct mass spectrometric determination of the protein profile differences between natural vs. man-made organisms.

A persistent question in microbial forensics is whether there are methods to determine if a pathogen detected within or isolated from the environment is naturally occurring. The phenotypic (i.e., expressed trait) state of an organism is impacted by the environmental cues that it experiences. In this way, assessing the phenotypic characteristics of an environmental pathogen may provide information about the environment in which it grew. The proteomic methods used have the potential to detect these characteristics; however, performing an analysis of environmental samples is challenging due to the microbial background and biochemical complexity.

Our project is evaluating optimal methods for *Bacillus anthracis* retrieval from a complex matrix such as soil with the goal of recovering sufficient spore cell mass for analysis and comparing to laboratory cultivated spores of the same strain. A laboratory mesocosm is used to simulate environmental factors that influence sporulation following death of a natural host. Spore mass retrieval occurs with state-of-the-art affinity reagents alone or in conjunction with cell sorting to reach a minimum estimate for proteomic analysis. A comparison of soil-produced spores with blood sporulation or the laboratory can provide useful insight into how these data are used to differentiate naturally and laboratory produced *B. anthracis* spores.



Bright field microscopy of (A) *Bacillus anthracis* sterne vegetative cells and (B) *B. anthracis sterne* spores. Vegetative cells were grown in tryptic soy broth, with spores were prepared in new sporulation medium. Scale bar is 10 µm.

Methods to grow and prepare *B. anthracis* sterne spores were identified, tested, and optimized during FY 2013. Spores were prepared on either a solid agar surface such as tryptic soy agar (TSA) or brain heart infusion (BHI) agar plates. A liquid sporulation media (new sporulation medium [NSM]) produced the purist spore preparations (> 95% spores) free of vegetative cell mass. For spore purification, spores and cells are pelleted by centrifugation and resuspended in sterile water. This solution was stored for 7 days at 4°C to lyse any remaining vegetative cells. Spores were then washed 3 to 6 times in sterile water prior to enumeration and characterization.

Evaluation of the affinity and specificity of the commercially available *B. anthracis* spore antibody was done by immunomagnetic separation. Antibody purchased from multiple commercial sources were coupled following manufacturer recommendations to Dynabeads®MyOne™ tosylactivated beads. Capture of 100 or 1000 colony-forming units spores or vegetative cells was performed in peptone buffered water. We determined that several available commercial antibodies bind to spores and vegetative cells with low specificity, with these antibodies not binding a high number of cells (low affinity). To improve capture efficiency, we are investigating additional antibodies and optimizing coupling and loading onto magnetic beads. The immunomagnetic approach should lead to capturing sufficient numbers of spores from complex environmental matrices for proteomic analysis without the need for a culturing step that could lead to increased microbial background.

We developed a spore lysis protocol with sodium dodecyl sulphate-compatible with mass spectrometry (MS) analysis and that allows recovery of the captured spores from the beads. We also tested a flow cytometric assay that allows us to screen spores rapidly from complex samples. Initial results are promising and indicate that flow cytometry will be a valuable tool for screening spores from samples either directly on the capture bead or after separation and elution. We anticipate that this will provide a rapid method to prescreen sample prior to workup for MS analysis.

In FY 2014, we will complete our evaluation of the capture antibodies, continue development of our spore recovery and lysis protocols, fully develop outflow cytometric assays, and continue to collect MS-based proteomic data for our laboratory mesocosm grown spore samples.

PN13022/2503

Engineering of Oleaginous Yeast for Production of Advanced Hydrocarbon Biofuels

Shuang Deng

We are using an innovative synthetic biology approach to develop a yeast system as an ultimate biofuel for primarily diesel and jet fuels. The results will provide extraordinary capabilities for advanced hydrocarbon biofuels development.

Limited fossil fuel supply and environment concern bring about urgent pressure to produce alternative fuel from renewable resources. Compared to ethanol, hydrocarbon is more promising because it has similar properties to gasoline or jet fuel. Oleaginous yeast has attracted attention recently because it accumulates lipid via de novo fatty acid biosynthesis using carbohydrates and could be a possible biofuel feedstock alternative. However, the lipids extracted from the biomass have to be treated by catalytic cracking or thermochemical liquification in order to convert to fuels such as gasoline and diesel.

It was recently discovered that *cyanobacteria* acyl-ACP reductases in combination with aldehyde decarboxylases is necessary and sufficient for converting fatty acid to alkanes or alkenes. *Lipomyces starkeyi* naturally produces high amounts of fatty acids; hence, it is a logical platform for engineering high rate alkane/alkene production. Our project is working to developing an innovative way to produce advanced biofuels (alkanes, alkenes) in vivo using a synthetic biology approach. Cyanobacteria acyl-ACP redutase and aldehyde decarboxylase will be engineered into *Lipomyces* and direct naturally synthesize fatty acid to make alkanes or alkenes. This work will provide extraordinary capabilities for advanced hydrocarbon biofuel development, which is expected to reduce petroleum usage significantly.

In FY 2013, we tested different antibiotic sensitivity for genetic screening in *L. starkeyi* and found that the hygromycin is the best selection maker. After we made numerous different transformation method attempts with *L. starkeyi*, including LiCl, protoplasting, and electroporation, we successfully developed the *Agrobacterium tumefaciens*-mediated transformation system that allows us to introduce genes into *L. starkeyi*. The transformation system is the foundation of this research.

Also in FY2013, different promoters from *L. starkeyi* were tested for their expression level by engineering a hygromycin selection marker under control of those promoters, respectively. Three promoters were chosen, including the *pyrG* gene, the transcription elongation factor (*tef*) gene, and Glyceraldehyde 3-phosphate dehydrogenase (*gapdh*) gene. The *pyrG* promoter drove expression for the hygromycin selection marker.

Two genes from *S. elongatus PCC7942*, acyl reductase (orf1594) and aldehyde decarbonylase (orf1593), were the first codon-optimized for high expression in *L. starkeyi*. Next, acyl reductase was built under the control of the *tef* promoter, and aldehyde decarbonylase was driven by the *gapdh* promoter. All of these gene fragments were built together using yeast gap repair. Finally, the plasmid construct was isolated and cloned into an agrobacterium binary vector that carries cyanobacteria acyl-ACP redutase and aldehyde decarboxylase into L. starkeyi. By the end of FY 2013, the alkane biosynthetic pathway will have been reconstructed in *L. starkeyi*. This engineered *L. starkeyi* culture is being analyzed for the hydrocarbon production. Additionally, the engineered strain will provide an excellent starting point for our proposed production of oleochemicals.

In FY 2014, we propose to optimize the production of hydrocarbons in our genetically modified *L. starkeyi* system and insert other genes necessary to make acids and alcohols. A plant acyl-ACP thioesterase (TE) gene is needed to liberate the fatty acids from the fatty acyl carrier. Different TEs with different substrate specificities determine the fatty acid chain length, so TEs can be used to engineer chain length composition. Fatty alcohols are usually produced by hydrogenation of plant oil or by synthesis from petrochemical precursors. However, genetically expressing the fatty acyl-COA reductase in *L. starkeyi* should lead to the production of fatty alcohols by the yeast.

PN13052/2533

Exploring and Engineering Phototrophic-Heterotrophic Partnerships

Hans C. Bernstein

This project aims to identify controllable, cooperative ecological phenomena employed by phototrophically driven microbial communities for conceptualizing and engineering multispecies biocatalytic platforms. The target application for this work is energy capture and transfer from renewable resources, light, and carbon dioxide.

Microbial consortium engineering has become an established scientific discipline populated by interdisciplinary biologists, engineers, and ecologists. The methodology is based on assembling microbial communities through enabling, encouraging, or enforcing interactions between distinct cell populations and their respective environments. Applications have the widely accepted potential to contribute technology toward key social benefits, such as biofuel production, carbon sequestration, and environmental remediation. The soundness of the consortium concept for biotechnology applications is supported by direct observations in nature. Optimized by eons of evolution, naturally occurring ecosystems are almost ubiquitously organized as interacting mixed communities. Of these, photoautotrophic microbial consortia are of keen interest to both chemical and biological engineers as these are promising catalytic systems capable of utilizing renewable resources light and carbon dioxide. Naturally occurring phototrophic organisms often couple with heterotrophic counterparts and engage in mutualistic interactions such as syntrophic exchanges of mass and energy originally derived from the environment.

The objective for this research is to dissect metabolically coupled interactions within naturally occurring biofilms from unique environments (i.e., high temperature and hypersaline) to discover the ecological cooperation strategies that can be used to build engineered microbial consortia. The focus will be on identifying natural biological phenomena between photoautotrophic and heterotrophic microorganisms that can be controlled and harnessed inside of the applied biological catalyst systems. This research differs from previously reported microbial community studies because it will advance technical understanding in the context of both fundamental and applied science. This study will focus on microbial interactions inside biofilms, which are surface-associated accretions of microorganisms encased within polymeric matrix materials. Biofilm-associated chemical reactions are non-equilibrium processes that both create and are dictated by gradients in mass, energy, and momentum. To date, the mechanisms by which microbial community interactions are coupled within resource gradients have yet to be exploited fully for microbial biotechnology.

Despite a late FY 2013 start, the project team established and secured the appropriate microbial laboratory space and chose a research field site: Hot Lake hyper-saline benthic microbial mat system. Field experimentation has commenced and will continue into FY 2014. The first field experiment (FE-1) is being designed to provide the essential data on the relevant ecophysiological parameters (i.e., spatially resolved chemical and energy requirements) governing the native phototrophically driven microbial community. The FE-1 experimental procedures are currently being conceptualized to provide a basis for interpreting the flux of energy (light, photochemical, and thermal) and mass (inorganic and reduced carbon species) into and within the native microbial mat/biofilm system.

The FE-1 experiment will employ combined microsensor and laser ablation isotope ratio mass spectrometry measurements for sub-millimeter spatial resolution of light attenuation, photochemical production, and carbon assimilation. The quantitative interpretation of these relevant energy and mass conservation processes in the native ecosystem will provide a benchmark to begin designing *ex situ* photo-biofilm reactor (PBFR) experiments in the laboratory, which is the first specific aim of our research. The expected outcome from the FE-1 experiments will provide, to the best of our knowledge, the first ever attempt at coupling the spatially resolved energy and carbon balances associated with a photoautotrophically driven hyper saline microbial mat community.

The first laboratory experiment will begin in early FY 2014. The goal will be to design, construct, and characterize a novel biofilm community culturing method employing a PBFR. This work will use the FE-1 data to provide an essential tool for highly controlled *ex situ* culturing and experimentation that are required for the duration of our project work.

PN13101/2582

Fungal Siderophore Production for Recognition of Uranium Compounds

David S. Wunschel

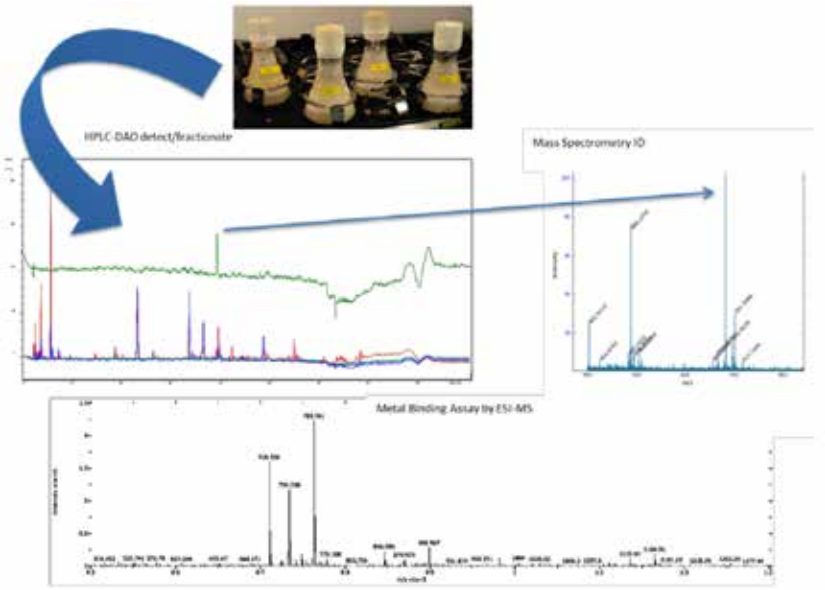
This project uses fungal siderophores as adaptable metal recognition molecules that can be engineered to function as ligands for uranyl species. These ligands will enhance the selectivity of detection platforms and focus on the products of human enrichment activities.

A number of mechanisms for binding metals have been suggested for remediation, but these approaches lack the specificity for definitive metals. Likewise, the organic synthesis of ligands is a potential solution for binding uranyl compounds; however, the organic synthesis of complex ligands requires a time consuming, expensive, and laborious process. Naturally occurring mechanisms of metal immobilization provide an attractive alternative because they can perform multi-step syntheses that are far faster and easier to scale up than synthetic chemical routes. Siderophores are one type of compound that microorganisms use to capture essential amounts of Fe and Cu for cellular metabolism and are capable of binding U and Pu compounds with affinities similar to the likely target metals. The siderophores are particularly attractive because there are a wide variety of chemical structures that potentially have diverse specificities within sequenced strains of fungi.

We sought to identify candidate fungal siderophores using separations of mass spectrometric and spectroscopic methods. Initially, we developed methods to detect and characterize siderophores from fungal cultures. Next, we investigated the tools for characterizing uranium compound-binding specificity. Having the siderophore production system in place as well as the tools to characterize the siderophores is vital in order to determine their binding specificity. The technical outcomes of this project are methods for siderophore expression, purification, and characterization. Measuring methods for metal binding specificity by mass spectrometry (MS) will be established for examining siderophore-metal binding. A major outcome is a collection of fungal siderophores characterized for uranium binding with the goal of identifying those with potential to recognize uranyl produced from anthropogenic activities.

Culture conditions for siderophore compound production were investigated through multiple fungi species, including strains of *Aspergillus niger*, *A. carbonarious*, *A. aculeatus*, and *Neosartorya fischeri*. A low Fe and standard medium was used to produce fungal cultures, followed by an organic extraction of the culture supernatant and then an extract concentration. Following extraction, a high performance liquid chromatography (HPLC) separation and UV/Vis absorbance measurement was used to screen samples for siderophore-like compounds using their spectroscopic properties as related to published surveys (noting absorbance in the 420–450 nm range) and commercial standards of ferrichrome and desferroxiamine A. In some cases, the HPLC-UV/Vis data contained a few clear peaks that were suspected siderophores. Comparison of low and high cultures illustrated that compound production appeared to be Fe regulated, consistent with previous reports.

Electrospray ionization (ESI)-MS was employed to examine siderophore metal binding ability. Different solvents were investigated for the ESI-MS analysis to determine a combination suitable for solubilizing the siderophore and any metal solutions used. The low solubility and stability of uranyl species in water led to an investigation of ethanol, methanol, acetonitrile, and dimethyl sulfoxide (DMSO) with water at various pH values. The goal was to determine the best solution for the solubility and ionization efficiency of both the free metal and siderophore; an acidified



General work flow of the project for fungal culture siderophore production, separation, and screening, followed by characterization of siderophore metal binding.

PN13025/2506

methanol solution appeared to favor metal and siderophore stability. Focused studies of ESI-MS uranyl solutions found that efficient ionization of uranyl nitrates was observed in 50% of the methanol solutions, sufficient to view uranyl complexes with themselves and nitrate species. These findings demonstrated that a methanol solvent system would be compatible for both siderophores and uranyl species.

Analysis of the uranyl species then focused on instrumental conditions (e.g., source collision energy) and solution pH to understand the impact of these factors on the observed species in the mass spectrometer. The non-acidified methanol solvent system was suitable for general analyses, but the pH can be varied to study binding affinity. This lays the foundation where each siderophore will be analyzed in this solution using a pH titration series to study gas association constants to understand siderophore metal binding selectivity. Secondary phases of analysis may utilize laser-induced fluorescence analysis to determine solution-phase binding constants.

Plans for FY 2014 are to apply this discovery approach to a broad set of fungal strains for siderophore production and evaluation. Those strains producing compounds determined to be the most selective for uranyl species will be candidates for larger scale production or genetic modification. The biosynthetic apparatus for siderophore production in fungi has been described for a broad number of species and strains. A long-term goal is to use a selection of strains and their siderophores as the groundwork for investigating ways to enhance the binding specificity of their siderophores through engineering the genes that encode them.

Genome-Enabled Systems Approach to Predict Immobilization of Technetium in the Subsurface

M. Hope Lee

Our research is an important first step for demonstrating the effectiveness of applying high throughput molecular techniques and metabolic models for monitoring and predicting technetium (^{99}Tc) transformation. Data and modeling will provide important insights into coupling microbial activity to reactive transport and hydrological models currently used for remediation.

Technetium (^{99}Tc) has been released into the subsurface over the past three decades through fallout from weapons tests and discharges from active nuclear processing plants and other facilities. ^{99}Tc has a relatively long half-life, is soluble and mobile in groundwater, and can be absorbed by plants and animals. If left untreated, ^{99}Tc poses a risk to the environment and human health for thousands of years. A variety of technical, scientific, and financial challenges complicate efforts for ^{99}Tc remediation due to the complex nature of the subsurface and associated biogeochemical cycles regulated primarily by microbial activity.

Although much has been learned about the physiology and metabolic potential of single microbial species (pure cultures) that immobilize ^{99}Tc, major gaps exist in our understanding of the function of these and other microorganisms in natural and contaminated ecosystems. As such, tools that integrate the chemical and biological reaction network influencing ^{99}Tc mobility in the subsurface need to be developed and tested. A first step in this development will be testing the utility of using cultivation-independent molecular tools, such as high throughput sequencing techniques coupled to genome enabled community-level metabolic models as a method for predicting and controlling the biotic component of ^{99}Tc immobilization. While other metabolic models have been tested and validated (some for contaminated subsurface environment)s, our stoichiometric metabolic modeling approach simplifies the microbial community into key or sentinels metabolisms and is not based on cultured surrogates or well characterized single microbial members. This type of modeling has been used successfully to analyze natural and engineered microbial systems on both a pure culture and community level and defines a network's metabolic potential based on a complete listing of the simplest, non-divisible pathways.

During FY 2013, the project achieved the following accomplishments. First, cultures of dissimilatory metal reducing bacteria (DMRB) isolated from Oak Ridge Field Research Center (FRC) sediments (*Geobacter sulfurreducens* strain PCA, *G. daltonii* strain FRC-32, and *Anaeromyxobacter dehalogenans* sp. strain 2CP-C), from Hanford subsurface sediments (*Cellulomonas* sp. strain ES6), and from freshwater lake sediments (*S. oneidensis* MR-1), were obtained and cultured. These specific bacteria were chosen because they have been shown to reduce metals and radionuclides (such as ^{99}Tc) chemically. In addition, genome sequence information that will be used to inform the metabolic models is available for all species but the *Cellulomonas* sp. strain ES6.

Initial experiments were performed to examine metal reduction using ferric iron (Fe(III)) as a surrogate. The *Cellulomonas* sp. strain ES6 was grown under aerobic conditions using trypticase soy broth. Experiments were performed under batch conditions in a chemostat under non-growth conditions, and lactate and iron were monitored. In addition, single-use chemostats and extensive chemostat control system modifications were performed in preparation of work with ^{99}Tc.

Finally, the construction of a genome-enabled metabolic model for individual species was initiated using genome information available for the above cultures in DOE's Joint Genome Institute. In addition, growth information from the literature and experiments performed has been used to estimate biomass production rates. Because the *Cellulomonas* sp. strain ES6 has not been sequenced, an initial model will be constructed next year using a related *Cellulomonas* species.

During FY 2014, experiments involving ^{99}Tc reduction under growth and non-growth conditions will be completed for each species. Once these experiments are completed, species will be combined, and similar ^{99}Tc reduction experiments will be performed. Metabolic models for individual species will also be completed, with this information subsequently used to generate a community level metabolic model to understand carbon and electron flow within the community and how this affects ^{99}Tc reduction. This model will be based on elementary mode analysis, and different methods of combining each species metabolism will be performed.

Finally, the metabolic model will be experimentally tested using stable isotope probing. Carbon sources that have been labeled with ^{13}C will be fed to the combined culture, with monitoring of genomic and metabolomics data commenced. Using the experimental data, the metabolic models will be adjusted accordingly to predict environmental systems responses more effectively.

PN13048/2529

Imaging and Monitoring the Initial Stages of Biofilm Formation

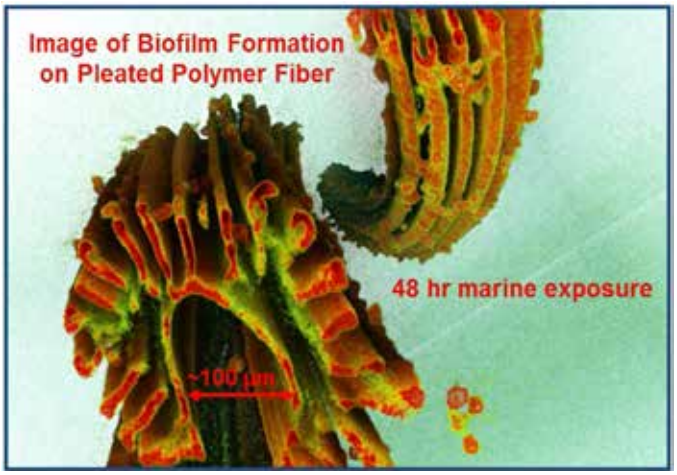
Raymond S. Addleman

The analytical capabilities developed in this project will enable unprecedented visualization, characterization, and understanding of the initial molecular and cellular processes involved in biofilm formation.

Systematic studies of the initial stages of biofilm formation – the conditioning film and primary colonizers – are lacking in literature largely due to the complex nature of the problem as well as the lack of available tools for studying the process. We are leveraging advanced chemical imaging techniques recently developed at PNNL to improve initial biofilm formation understanding and subsequent surface biofouling. This interdisciplinary effort will provide better tools and mechanistic understanding of the initial biofilm formation processes on existing and novel material surfaces. Having new tools to study the processes that lead to biofilm formation will support the development of new, innovative, non-toxic means to combat fouling relevant to DOE missions in hydropower, efficient desalination, ship transport, bioremediation, and improved separations and catalysis as well as biomedical and other applications.

Key issues to be explored in our research include determining which new analytical methods are best suited to visualize and characterize the initial stages of biofilm formation. We will investigate how surface structure at the nanoscale and microscale affects the composition and structure of the initial biofilm as well as how surface chemistry and physical properties (e.g., hydrophobicity, hydrophilicity, surface roughness, surface patterning) affect the composition and structure of the initial biofilm and rate of formation. Another topic of consideration is whether the initial biofilm is deposited at random, as a uniform film, or is concentrated around identifiable nucleation points. Additionally, we will examine how the composition and structure of the resulting biochemical-organic layer affect the initial deposition of colonizing cellular organisms and consider how improved imaging data can provide a mechanistic understanding of how to construct multiscale materials to minimize and prevent biofouling.

A critical challenge for the study of biofilms has been the lack of methods and instrumentation that can perform detailed measurements and monitoring of wet surfaces. However, the recently developed methods for *in situ* environmental imaging that can work with wet samples do not require electron-dense coatings or high vacuum conditions



X-ray microtomography image of marine biofilm forming on a pleated polymer fiber. After 48 hr of marine exposure, the initial cells colonizing the fiber can be observed. X-ray microtomography provides 3D imaging with a resolution of ~1.4 µm.

and provide an opportunity to characterize near-nanometer scale resolution of surfaces, biomolecules, and cells as they interact and assemble on surfaces. Our mid-year start in FY 2013 focused on evaluating these methods for biofilm applications. The imaging and analytical techniques explored included

- MALDI-TOF, shown to be effective for biochemical characterization during biofilm formation
- X-ray microtomography, shown to be effective for structural characterization of biofilms
- ultrasonic imaging, shown to be effective for in-situ monitoring of the growth of biofilms
- optical microscopy, with efforts underway to provide dual wavelength monitor for correlation of cellular luminesce with surface nanostructure.

For FY 2014, we will continue to explore new methods to monitor the composition of the biofilms as they form. In addition to finishing the development on the promising methods from FY 2013, we intend explore other methods such as AFM-Raman, nano-DESI, TOF-SIMS, E-STEM, and X-ray and electron spectroscopy. We will image and monitor the impact of surface composition (structure and chemistry) on the initial cellular settling and colonization processes and determine sequential dependencies of biomolecular deposition onto surfaces. With demonstrated methods, this project will explore biofouling on selected materials in two focused, high impact areas: marine biofouling and biomedical-related biofilm formation. The biofouling processes as a function of time and surface composition will be studied quantitatively.

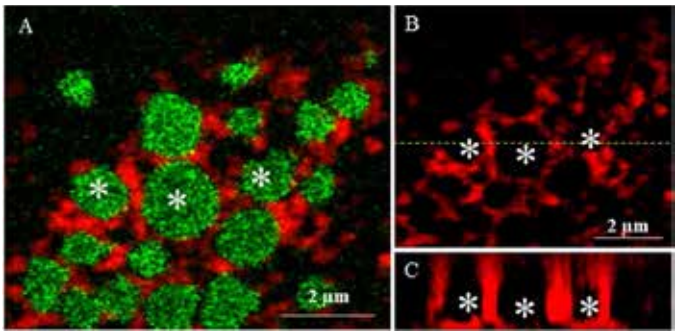
PN13067/2548

Integrated Nano-Scale Imaging for Investigating Applications and Implications of Nanomaterials

Galya Orr

We developed integrated super-resolution fluorescence and X-ray microscopy techniques to investigate mechanisms underlying cellular interactions and fate of individual nanoparticles in the intact hydrated cell with nanometer (nm) resolution. This approach enables molecular-level insights into the spatial distribution of specific proteins, molecular complexes, individual organelles, or nanoparticles in the cellular environment, leading to a new understanding of their behavior and function in cell response and survival.

This project provides new imaging approaches to investigate individual nanoparticles and their interactions with living cells, providing data needed to accelerate safe applications of nanotechnologies in industry and medicine. The unique properties of materials at the nano-scale have been explored for multiple applications, but the unknown impact of nanomaterials on human health and the environment has limited their use. The potential of engineered nanomaterials to impose adverse effects on living systems has been observed, but a great deal of confusion remains about the properties that make a particle toxic or biocompatible. The potential toxicity or biocompatibility of nanoparticles is governed by cellular interactions and particle fate, which dictate cellular response and ultimately determine the impact on human health. Investigations of these relationships have been challenged by the limitations of current research technologies and by the strong tendency of nanomaterials to agglomerate under experimental conditions, providing information that is irrelevant to nano-scale materials and the exposure in vivo.



A. Correlative STXM and super resolution structured illumination fluorescence microscopy (SIM) of a small area in the cytoplasm of an intact hydrated alveolar epithelial cell showing aggregates of cerium oxide nanoparticles (green) inside lysosomes (red); B. The presence of the particles within lysosomes is determined using a specific fluorescent marker; C. Side view along the dashed line in B, confirming tight fit of the red lysosomal marker around three of the particles (stars) in 3D.

Our main objective was to develop new approaches that allow us to study individual nanoparticles within the intact cellular environment one nanoparticle at a time to delineate particle properties and cellular processes relevant to nanomaterial and in vivo exposure. Maximizing our expertise in single-molecule fluorescence techniques and working with the Advanced Light Source (ALS) at Lawrence Berkeley National Laboratory, we developed approaches to obtain chemical and structural images of individual nanoparticles within organelles in the intact hydrated cell with nm resolution. These approaches were applied to investigate the cellular interactions and fate of individual nanoparticles to identify particle properties that are toxic or biocompatible and delineate the underlying mechanisms.

In collaboration with ALS scientists, we investigated the chemical state of individual nanoparticles in the intracellular environment of intact hydrated alveolar epithelial cells grown in culture. We established experimental parameters for investigating this system with 20 nm resolution using scanning transmission X-ray microscopy (STXM), correlative STXM, and super resolution fluorescence imaging to identify the location of the nanoparticles within distinct organelles with nm resolution. Our recent studies focused on cerium oxide (CeO_2) nanoparticles, which have been reported to induce both protective and toxic effects on the cells. These conflicting results might reflect selective shifts in the reactivity state of the particles in different cellular environments. To investigate this possibility, we have applied correlative STXM and super resolution fluorescence imaging and determined the oxidation state of CeO_2 nanoparticles inside the cytoplasm. We found a clear shift in the oxidation state of the nanoparticles inside lysosomes, which in turn, pointed us to distinct molecular interactions likely to underlie the cellular response. In summary, by establishing correlative STXM and super resolution fluorescence imaging in intact hydrated cells, we achieved new understanding of molecular interactions that occur between untagged pristine CeO_2 nanoparticles and the intracellular environment with unprecedented resolution.

PN11043/2365

Mapping and Characterization of Organic Matter in Soil Aggregates using Laser-Ablation Sampling

M. Lizabeth Alexander

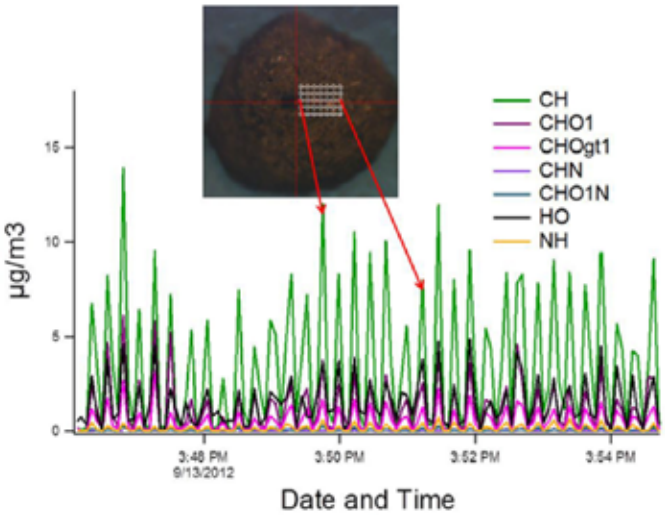
We are applying existing technologies in new ways to examine the relationship between the location of microbes and soil organic carbon (C) as well as the relative age of soil organic C in specific areas.

Although critical to a complete global picture of C cycling and the relation to climate change, the current understanding of C deposition, microbial utilization, and sequestration of organic C in soil is based almost entirely on analysis of bulk samples with no clear understanding of the underlying distributions and physical-chemical controls that occur at the micron-scale in intact soil aggregates. C cycling in soil is primarily a function of microorganisms interacting with their environment. A true mechanistic understanding of C cycling in soil must therefore be at the sub-millimeter scale. Because of historical technological difficulties in collecting the necessary date at this scale, most modeling approaches use empirical data collected at batch and field scales. However, direct transfer of these empirical relationships into models may not adequately represent mechanistic processes occurring at the microscale, and may not lead to correct model predictions when the soil system is perturbed. Many modelers believe that an improved fundamental understanding of mechanistic drivers at the microscopic scale is highly appropriate and would be easier to scale up in predictive models than are purely empirical relationships determined at larger spatial scales. A major hindrance to a micron-scale mechanistic understanding of microbial organization and C cycling processes in soils is the current inability to link physical structure to distinct microbial and chemical properties. This inability stems from multiple problems associated with the size-scale of the observations desired, opacity of the soil, chemical heterogeneity of soil organic C, and how utilization of soil C sources or pools varies spatially and temporally.

This project applies existing technologies of laser ablation (LA)-aerosol mass spectrometry (AMS) and LA-capillary absorption spectrometer (CAS) to study the speciation and spatial variation of C in soil. These sophisticated methods have potential to overcome observation scale and soil opacity problems, map elements spatially, and produce molecular information at a lateral resolution between 1–10 μ and depth resolution of as little as 100 nm. This capability opens a variety of avenues for examining the relationship between the locations of microbes or soil organic C and the relative age (young/labile versus old/recalcitrant) of soil organic C in specific areas.

The bulk of FY 2013 work focused on understanding and resolving issues arising from the details of how ablation laser samples a soil aggregate and transports particulates to the spectrometer, particularly the lateral spatial broadening in the transport tube that leads to temporal broadening in the AMS data output and ultimately determines spatial resolution of the LA-AMS technique. Another critical issue is how the tightly focused ablation laser interacts with and is affected by the uneven surface presented by soil aggregates and how this affects both the absolute signal level and relative levels of specific components for example the C/N ratio.

Initial measurements were made using the same technique employed successfully for imaging a fungal hypha. The technique employed was to run the laser at a constant repetition rate of 10–20 Hz and scan at a constant but slow rate with continuous data collection. When this approach was attempted with aggregates, the data showed little variation in absolute signal level or component ratios, even though the aggregates appeared heterogeneous under the microscope. A primary difference between the measure-



LA-AMS signal for soil aggregate taken by sampling with 1-sec bursts of 20 laser shots using 4-sec acquisition windows at adjacent locations using a laser spot size of 20 m.

PN12047/2448

ments of single fungal hyphae and soil aggregates is that a hypha is isolated on or near the surface of an agar medium. It is also possible that the hyphae have more lateral variability than the aggregates on the scale probed.

A series of studies was conducted to determine the temporal spread of particulates arising from the process of sucking them from the ablation site and from lateral spreading in the transport tubing to the AMS. These were performed by firing a short, 1-sec burst of 10–20 laser pulses and running the AMS in a very fast data acquisition mode so the rise and fall of particulates could be observed. Initial results were obtained using the 3/8-inch tubing standard with the AMS and running at standard flow rates and pressures used in the normal operation of the LA system in conjunction with ICP-MS. Substantial broadening was found on the order of 4–6 sec for the hypha/agar sample matrix and 10–20 sec for the soil aggregate. Data from the soil also frequently had a tail that lasted much longer.

A number of iterations were tried using various combinations of ablated particle sample pickup geometry, flow rates, and tubing diameter and lengths. A primary issue is that the commercial sample cell is designed to operate with and ICP-MS having an inlet flow rate of about 1 L/min, while the AMS aerosol inlet runs at 0.085 L/min. Running at the standard flow rate would reduce the signal level by over an order of magnitude but using the flow rate compatible with the AMS increases the length of time required to sweep out the volume of the ablation plume, producing temporal broadening. The linear velocity of the carrier gas in the transfer tubing is also decreased at lower flow rates allowing more lateral spreading which translates into even more temporal broadening. The observation that these effects are more pronounced for the soils than the hypha/agar matrix suggests that the particulates from laser ablation of soil have a larger aerodynamic diameter which is a function of size, density and shape. This is reasonable given the gelatinous nature of the agar which could well result in the production of "fluffier," more fractal particles which would not be as subject to broadening due to a lower aerodynamic diameter.

Modifications to the laser ablation sample chamber including consisted of changing the design in the sampling pickup, which uses a cyclone effect to entrain all ablated material and to allow the pickup to function properly with smaller, 1/16-inch O.D. transfer tubing to maintain the same linear velocity in the transfer tube and minimize lateral spreading and therefore temporal broadening in the signal. These modifications reduced the broadening so that over 90% of the signal from a single 1-sec burst of 20 laser pulses arrive in a 2-sec window with 100% of the signal arriving in a 4-sec window. The approach was then to ablate discrete 20 μ spots with a 1 sec burst and wait 2 sec between. The area on this particular soil aggregate was chosen because it was relatively flat when viewed under the CCD microscope of the laser ablation system and was "milled" with the ablation laser prior data acquisition to minimize any effect of surface variations.

Data presented from the AMS represent the organic components of the soil and does not include minerals such as oxides so that the variations in total signal are a true representation of the total organic fraction plus small amounts of less refractory inorganic compounds such as ammonium, nitrate, and sulfate. The aggregate showed a variation from 40 to over 400 with the actual upper bound over 200. These ratios indicate that there was little microbial mass present, and the high ratios indicate that very little humic material was present as well. These results show a need to perform characterization of soils with well-known compositions from a variety of organic material. It is unfortunate that the necessity of focusing on technical details of the LA-AMS process precluded this type of detailed analysis. The technical focus also did not allow any demonstration of high-throughput imaging but was a critical step in moving the LA-AMS technique toward those goals.

The results from this project have been encouraging enough to stimulate significant interest and support from the scientific community and as well as an internal offer to continue this development into FY 2014. In addition, a collaboration with Aerodyne Research has been established with a proposal to characterize the soil composition for these studies. An initial paper is in preparation and will be submitted by the end of FY 2013.

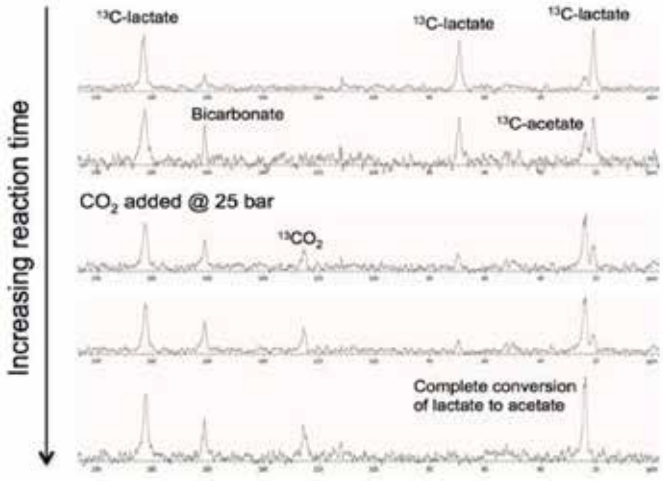
PN12047/2448

Microbial Processes Accompanying Deep Geologic CO_2 Sequestration

Matthew J. Marshall

This project aims to increase understanding of the impact of geologic CO_2 sequestration on microbial biomass and viability. Deep subsurface locations for CO_2 storage contain microbial populations that may impact gas generation, caprock integrity, and other parameters that need to be better understood when planning and modeling the behavior of these sites.

The capture and storage of carbon dioxide in deep geologic formations represents one of the most promising options for mitigating impacts of greenhouse gases on global warming, owing to the potentially large capacity of the formations and their broad regional availability. CO_2 is injected into formation fluids, nominally saline aquifers, as a supercritical fluid ($scCO_2$) at depths >2500 m that maintain its supercritical state. Studies on "deep biosphere" have demonstrated that microbial communities are present in all environments that have been suggested as suitable locations for geologic carbon dioxide sequestration. In some industries, $scCO_2$ is used as a sterilizing agent; however, field trials have detected microbial activity after CO_2 injection, suggesting that indigenous bacterial strains are able to tolerate these extreme conditions. The persistence of bacteria in these environments is important because these microorganisms may affect long-term storage of CO_2 in a number of ways. Bacteria may act as reaction sites for increased rates of mineral precipitation, generate mineral precipitates as a result of their metabolism, and may also generate methane gas within the storage location.



A temporal series of NMR scans showing the utilization of lactate by *Desulfovibrio vulgaris* and conversion to acetate and CO_2 (as bicarbonate) under 25 bar CO_2 at 37°C.

We focused our current studies on the physiology and metabolism of a model subsurface sulfate-reducing bacterium (SRB), *Desulfovibrio vulgaris*, commonly found in deep subsurface environments that are promising locations for CO_2 storage. This research has leveraged unique experimental capabilities that PNNL has developed. High-pressure nuclear magnetic resonance (NMR) analyses of growing D. vulgaris cultures have been performed to determine the in situ metabolic state of these cells under varying CO_2 partial pressures. Biomass is incubated with ^{13}C-lactate solution that acts as a source of both carbon and electrons for growth. As the biomass utilizes this compound, increasing concentrations of daughter products (e.g., acetate, bicarbonate) can be monitored to determine rates of metabolism. To date, we reproducibly demonstrated microbial growth in this system under both ambient (atmospheric pressure, 37°C), and pressurized (25 bar, 37°C) conditions. As this research progresses, the response of biomass to increasing pressures will be determined. Additional NMR experiments are also planned using a methanogenic (methane-forming) bacterium that can generate methane (CH_4) from H_2 and CO_2. These are of great interest in storage environments where CO_2 will be in excess and H_2 may be present from organic material fermentation. By using ^{13}C-labeled CO_2, we anticipate the ability to track methanogenic bacteria activity under a range of pressures found in deep subsurface environments.

In tandem with high-pressure NMR investigations, the physiological state of *D. vulgaris* under elevated CO_2 pressures has been investigated at the mRNA level using a shotgun RNASEQ approach. A comprehensive suite of samples is currently being sequenced for metatranscriptomics; biomass was exposed to CO_2 pressures of 25, 50, and 80 bar before being flash frozen to preserve mRNA signatures that may indicate potential stress responses to CO_2 pressures. While preliminary data have indicated that some aspects of CO_2 stress response is shared across other stresses (e.g., pH, metal exposure), other genes suggest that there may be some novel aspects to the cellular response. In particular, the up-regulation of a series of hypothetical genes warrants further attention. The sequencing of additional metatranscriptomic samples will allow gene network responses to be elucidated and may reveal roles for specific poorly annotated genes.

PN11052/2374

During FY 2013, we continued experiments utilizing SRB strains in tandem with high-pressure NMR and commenced metatranscriptomic dataset analyses. Consistent results are being achieved using the 300 MHz Tesla high-pressure NMR system. Briefly, we can incubate a representative sulfate-reducing bacterium in a small volume within an NMR rotor. Cells are grown on 3-^{13}C lactate, which is metabolized to acetate. Using NMR, we are able to detect all three carbon peaks arising from lactate, and can track conversion of lactate to acetate. Baseline rate measurements for this conversion have been gathered at 1 bar CO_2 pressure, and recent experiments have collected similar data over a range of CO_2 pressures at 37°C. Preliminary results suggest that even at low pressures, CO_2 exposure can have significant deleterious effects on lactate metabolism; experiments at 25 bar CO_2 resulted in no conversion of lactate to acetate. Slower rates of conversion were observed at 15 bar CO_2. Currently, further measurements are being taken to generate consistent rates for these conversions, which can then be compared across a range of pressures. Coupled to high-pressure NMR experiments, complementary studies are being performed using *D. vulgaris* in batch reactors to study CO_2 exposure effects on physiology and metabolism. Biomass generated in these reactors has been used in RNA-SEQ studies in addition to TEM investigations. Imaging has indicated changes in cell membranes following CO_2 exposure, generating further hypotheses being investigated in the laboratory.

Cultures of *D. vulgaris* were incubated at a range of CO_2 pressures (atmospheric, 25, 50, and 80 bar) and subsequently harvested for RNA-SEQ analyses. Coupled measurements of cell viability indicated that although cells weren't actively metabolizing at these pressures, they remained viable over many hours. RNA-SEQ data have identified a number of mechanisms that may enable *D. vulgaris* to remain viable under these conditions, including the regulation and expression of certain amino acids. Transcripts for leucyl and Isoleucyl-tRNA synthetases (DVU1196; DVU1927) were both up-regulated at 25 bar, while expression of a tRNA modification GTPase (TrmE; DVU1079) exhibited the same pattern. Further evidence for a significant amino-acid based response to these stresses was inferred from genes associated with leucine biosynthesis. In addition to the leucyl-tRNA synthetase, transcripts from a 2-isopropylmalate synthase (DVU2981), and both the large and small subunits of a 3-isopropylmalate dehydratase (DVU2982 and DVU2983, respectively) were similarly up-regulated. This pattern was additionally maintained in the 50 bar and 80 bar CO_2 exposures. However, it is worth noting that the up-regulation of leucine biosynthesis may be a general stress response in these species; increases in mRNA transcripts associated with leucine biosynthesis were reported under both osmotic stress and alkaline stress conditions.

Micro-Fluidic Models for Studying Microbial Communities—Integration of Micro-Fluidic Model Experimentation, Multimodal Imaging, and Modeling

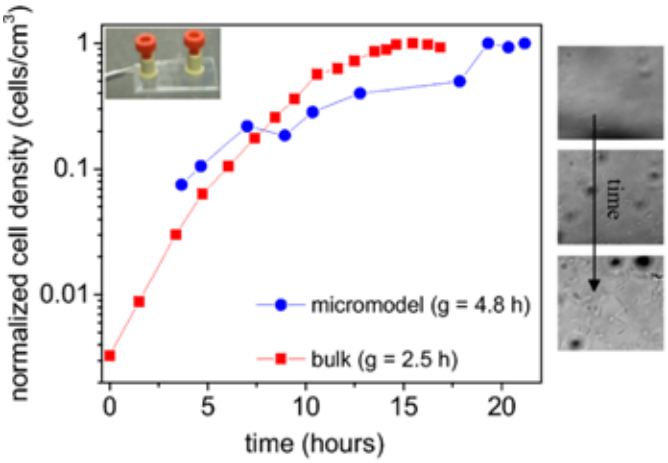
Michael J. Wilkins

Understanding soil carbon flux is key for developing modeling tools to predict carbon cycling in the environment, which has implications for climate change and land use policy.

Cellulose is the primary structural component of plants and the most common organic compound on earth. The global carbon cycle is an interplay between carbon reservoirs in the atmosphere, oceans, subsurface, plants, and soil, with microbial communities mediating many transformations between reservoirs. The flux of microbial-generated CO_2 from soil to atmosphere remains the least constrained component of the terrestrial carbon cycle. Approximately 10% of atmospheric CO_2 cycles through the soil annually, yet the temporal and spatial dynamics of soil respiration remain difficult to model or predict. Within subsurface environments, cellulose degradation occurs in pore spaces where a range of advective and diffusive biogeochemical processes produces local microenvironments, the conditions of which can affect the distribution of microbial communities in the subsurface and subsequent rates of cellulose degradation. Understanding pore-scale microbial processes affecting rates of subsurface cellulose degradation is fundamental to predicting and modeling the global carbon cycle.

This project studies controls on rates of microbial cellulose degradation across spatial scales and chemical gradients. Data obtained from these experiments can be used to refine modeling approaches for terrestrial C cycling, with technologies arising from a number of projects integrated into these data sets. Our general efforts are aimed at developing microfluidic platforms containing bioavailable carbon substrates used for pore-scale investigations into microbial-mediated cellulose transformations and forming methods for tracking enzyme activity on novel carbon substrates using multi-model nondestructive imaging techniques.

In FY 2013, our work focused on developing cellulose thin films that could be deposited within micron-scale pores. In addition, specific cellulose-binding dyes were used to label TMS-cellulose and allow visualization of these thin films using fluorescence microscopy, the intensity of which was proportional to the amount of cellulose deposited. As cellulose was degraded by the action of microorganisms or pure cellulase enzymes, a corresponding decrease in fluorescent



Cell growth dynamics for Clostridium thermocellum. While this strain grows within the microfluidic pore networks, rates (g) are slower relative to cells grown in a well-mixed batch media. This suggests that chemical gradients and other variables that develop in the pseudo 2D micromodel network influence growth rates, as might happen in porous natural environmental media.

intensity was detected. The degradation of this material was demonstrated using the anaerobic thermophilic cellulose-degrading bacterium microorganism *Clostridium thermocellum*. This strain was able to grow within pore structures of the microfluidic model and degrade TMS-cellulose thin films. Additionally, these experiments demonstrated the ability of microfluidic platforms to operate at elevated temperatures (*C. thermocellum* grows at an optimum temperature of ~65°C) and under oxygen-free conditions.

Subsequent experiments utilized the microfluidic platforms and cellulose thin films for experiments investigating rates of cellulose-degradation by purified cellulase enzymes. In collaboration with Washington State University, novel cellulase enzymes from the cow rumen system were cloned into *E. coli* vectors and over-expressed. Two novel cellulases were harvested, purified, and added to microwells that had been installed on top of a cellulose thin film. This setup allowed the concurrent monitoring of a cellulase-free control well and two cellulase wells. Results indicated that cellulase-degradation kinetics were variable between the two cloned enzymes and demonstrated the use of cellulose thin films coupled with multi-model non-destructive imaging as an accurate, high-throughput technique for assessing cellulase enzyme activity. These results have supported ongoing research using similar microfluidic platforms and oxygen-sensing optode materials to measure regions of microbial activity at the pore scale.

PN10055/2297

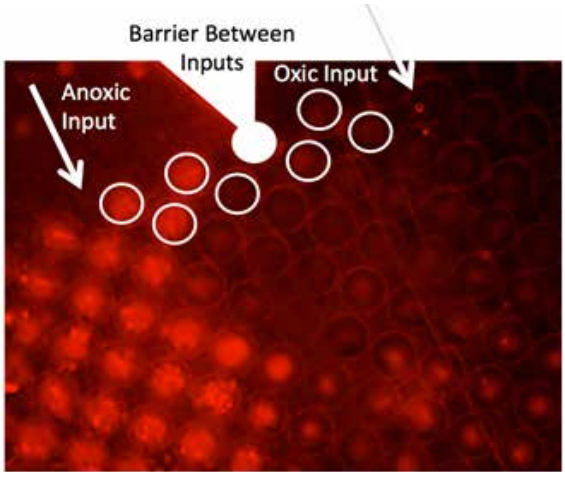
Multimodal Imaging of Cellulytic Microbial Activity in Oxygen-constrained Microenvironments

Jay W. Grate

This project integrates sensing and imaging techniques with microfluidic devices as spatially structured microscale models for carbon cycling processes in soil and sediments.

Soil is a heterogeneous, spatially structured environment subject to temporal variations in saturation, temperature, resource availability, and connectivity between pore spaces. Important biogeochemical processes occur at microsites in soil aggregates and pore spaces, where certain microsites at certain times may have a disproportionate effect. These are known as "hot spots" and "hot moments." Hot spots tend to occur at interfaces such as an aggregate surface, the boundary between 'oxic and anoxic environments, or along hydrological flow paths providing a high supply of reactants. Accordingly, it is essential to understand processes at the pore and soil aggregate scale (i.e., microscale) where fundamental processes occur and hot spots or moments may arise.

Conventional tools, techniques, and materials for studying microbial carbon cycling in terrestrial soils do not always lend themselves to carbon cycling studies in spatially structured environments at the microscale, as these opaque media that do not permit bright field or fluorescence imaging. Conventional CO_2 gas efflux studies from wetted soil provide data averaged over larger scale soil samples and do not capture microscale heterogeneity. Batch microbial cultures can lack spatial structure and are not representative terrestrial environments. Cellulose hydrolysis in batch mixtures is also not representative of cellulolytic processes acting on solid cellulose at specific locations. In addition to spatially structured microenvironments suitable for microbial growth, new microfluidic models must provide solid cellulose carbon resources that can be imaged and provide methods to measure local concentrations because oxygen is used in aerobic respiration of cellulose to CO_2.



Fluorescent intensity image for a pore network device with two inputs. Pores and pore networks are defined by pillars made of an oxygen-sensing material. White circles indicate where pillars are located, and a structural barrier between input regions is solid white. Fluorescence intensity is quenched by oxygen; bright fluorescence on the other side indicates an absence of oxygen. Such a micromodel can indicate where oxygen is being consumed by respiring microorganisms.

To these ends, we developed three successful synthetic chemistry approaches to covalently dyeing cellulose nanocrystals with Alexa Fluor dyes. In each case, observations confirm that the dyes are covalently bound to cellulose rather than simply physisorbed. These materials can be deposited in pore network microfluidic structures, and we demonstrated fluorescent imaging of their spatial locations and cellulolytic consumption. We also developed new fluorescent oxygen sensing materials that provide new flexibility in designing and fabricating microfluidic structures that enable oxygen chemical imaging within the flow paths and pore spaces, and can reveal where microorganisms are depleting the oxygen. Finally, we designed microfluidic devices where microorganisms can live without being displaced by flow and the oxygen depletion can be observed using incorporated fluorescent optode materials. Oxygen depletion by a cellulolytic microbial culture in this microfluidic structure was successfully demonstrated.

A new microscope was installed to facilitate all aspects of this research. The microscope includes imaging in wide-field fluorescence, bright field, spinning disk confocal fluorescence, and a method for performing fluorescence lifetime imaging measurements at all pixels in the image field simultaneously. The fluorescence lifetime imaging method, which can image long fluorescent lifetimes typical of oxygen sensing fluorophores, is used in conjunction with optode oxygen sensing materials.

As a result of developing these unique tools and approaches that significantly differ from prior conventional tools for studying carbon cycling, PNNL is well positioned to design and conduct experiments to investigate hot spots and moments of solid cellulose respiration at the microscale.

PN13064/2545

Multiscale Simulation of Microbial Carbon Transformation in Soils: Connecting Intra- and Inter-Aggregate Scales

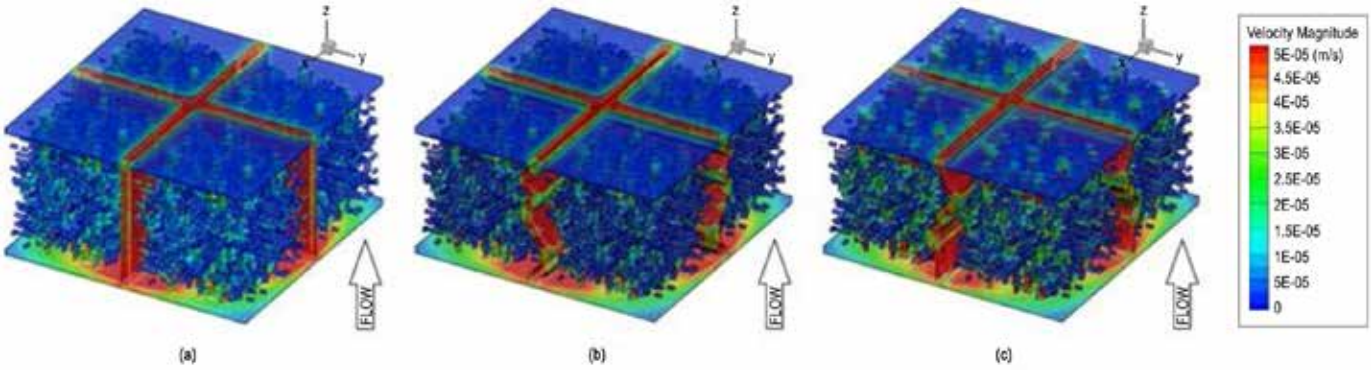
Timothy D. Scheibe

This project is developing multiscale modeling approaches to link the dynamics of microorganism communities acting within environments that vary at small scales (within individual soil aggregates) to the controls exerted by inter-aggregate flow and transport.

The earth's soils represent a major pool of carbon in the global carbon cycle estimated to be larger than the amount of carbon in the atmosphere and double the amount stored in vegetation. It has previously been demonstrated that individual soil aggregate particles form unique microenvironments within which variable microbial communities form and control rates of carbon degradation activity. The dynamics of these microenvironments are linked through fluid flow and solute transport in the inter-aggregate pore spaces at larger scales. Therefore, development of multiscale models that connect microbial community dynamics acting within microscale environments to controls exerted by flow and transport at larger scales is necessary to improve our understanding of factors that control rates of microorganism mediated carbon turnover in the soil carbon pool. Specifically, this project builds on earlier aggregate-scale model of carbon-degrading microbial communities work that simulates individual organisms and their interactions. Our work adds coupling within aggregate- to pore-scale models of fluid flow and nutrient transport to create a multiscale simulation capability that addresses the previously defined need.

A late FY 2012 project start yielded only development of whole, aggregate level models based on individual-based study results initiated, with work on modifying smoothed particle hydrodynamics software for pore scale flow calculations (to include microbial reaction kinetics) briefly started. For FY 2013, work progressed in three areas: 1) study of flow partitioning between inter- and intra-aggregate pore spaces; 2) development of upscaled whole-aggregate models based on individual-based aggregate-scale simulations; and 3) modifying a pore-scale multiphase flow simulator to accommodate two-phase (air/water) flow, solute transport, and upscaled microbial reaction kinetics. The individual-based aggregate-scale model assumes full saturation and no flow within intra-aggregate pore spaces. This assumption is reasonable because of small pore size (<10 μm) that leads to high capillary forces and strong water retention. A series of simulations was performed using a saturated pore-scale simulator to test and validate the assumption, considering variability in inter-aggregate spacing, shape of inter-aggregate pores, and internal aggregate porosity. It was demonstrated that only a small fraction of the total fluid transport through the model systems occurred through intra-aggregate pore spaces in all cases considered. A manuscript describing this work was submitted to a peer-review journal.

The individual-based aggregate-scale simulator was used to perform a large suite of simulations within which model parameters were varied systematically to cover the expected range of aggregate conditions in a natural soil system. From each individual-based simulation, averaged fluxes were computed and used to derive distributions of



Visualization of simulated water drainage in a model porous medium. Water phase is represented by green, granular solids are blue, and air phase is red. Note pockets of water trapped by capillary forces as the air front moves through. Simulation results from the SPH code.

PN12052/2453

upscaled reaction model parameters. These parameter distributions will be used in pore-scale simulations (which cannot explicitly resolve individual microorganisms due to computational limitations). In this way, the aggregate- and pore-scale models are explicitly coupled. A manuscript describing these simulations and the upscaled model parameter distributions is in development.

In natural soils, the dynamics of partially saturated water flow play a critical role in the redistribution of reactants such as oxygen and reaction products (e.g., soluble forms of organic carbon) among multiple soil aggregates. A pre-existing pore-scale simulator based on the smoothed particle hydrodynamics (SPH) numerical method was adapted to simulate multiphase (air-water) flow and incorporate upscaled aggregate-scale microbial reaction models. A series of simulation test runs was performed to establish SPH model parameters needed to simulate the movement of air-water interfaces accurately, which is controlled by surface tension between the two fluids and contact angles between the fluids and the solid grains. A reaction module was added to the SPH code to incorporate upscaled reaction systems defined based on the aggregate-scale simulations as described previously. The reaction solution was implemented using a multi-thread computational algorithm that provided significant speedup of the computationally intensive solutions. Test simulations have been performed using 3D model systems with several soil aggregates using millions of computational particles. Once simulations achieve satisfactory results for this test system, a small number of very large simulation runs will be performed using larger 3D systems. These runs will explore the impacts of wetting and drying on effective (pore-scale) reaction rates based on individual aggregate-scale reactions models.

The primary outcome of this project is a modeling system that links carbon degradation dynamics across three critical length scales: 1) the scale of individual micro-organisms; 2) the scale of single soil aggregates, and 3) the porous medium scale, composed of many soil aggregates. This work is a critical first step toward obtaining fundamental understanding of the impacts of microscale processes ("the world microbes see") on large-scale soil carbon dynamics and will support a primary DOE program objective to develop a robust predictive capability of carbon cycling in the natural environment.

Optofluidics and Microfluidics for Exploring Biofuel Production at the Single Cell and Molecule Levels

Andreas E. Vasdekis

This project investigates biofuel production at the single cell level. By treating cells as a single biochemical factory, biofuel production, efficiency and timing will be precisely measured, hence unmasking better-performing traits.

Our society's prosperity and growth is inevitably linked to reliable energy sources. Due to climate changes, increasing energy demands, and finite fossil fuel reserves, sustainable prosperity and growth necessitate the use of renewable energy sources such as solar-, wind-, or biofuel-based. The latter are of key importance due to their compatibility with our current transportation infrastructure. Despite substantial advancement during the past decades, there is still a plethora of unknowns related to such alternative energy sources, especially within the context of efficiency and species selection or design.

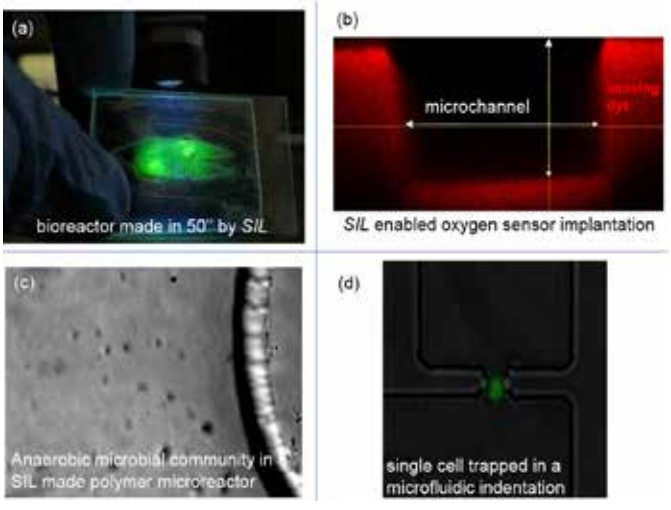
This project addresses such challenges by directly imaging biofuel synthesis at a controlled cell and molecule density down to the single entity level, at which the cell acts as an isolated chemical factory, while its synthetic performance can be precisely characterized, unmasking invisible spatio-temporal phenomena and better-performing individuals. The project's methodology is based on cell handling microsystems, chemical imaging of biosynthesis, and 'omic analyses (i.e., transcripotmics and proteomics) of such pre-analyzed microorganisms.

In FY 2012, the project reached two milestones, namely our extensive work with single cell microfluidics and chemical imaging methods. Key microorganisms were selected to this end, such as the cellulose fermenters as well as efficient fungi in lipid synthesis. These selections were made in collaboration with scientists at MIT and EPFL.

With regard to microfluidics, single cell isolation was developed and published using both bacteria and fungi. Such assays contain both micron- and nano-scale features capable of trapping single or few cells. In parallel, a novel microfluidic prototyping technique was developed, termed solvent immersion lithography (SIL), which exploits meso-scale polymer properties and enables microsystem prototyping in under 1 min in any polymer, thus overcoming limitations of traditional lithographic processes. SIL was critical in studying the thermophilic anaerobe *Clostridia thermocellum*, an efficient cellulose fermenter that is impossible to grow and study in conventional elastomeric microbioreactors. Additionally, the method enables the simultaneous functionalization of the imprinted features with chemical moieties, an aspect studied in the context of optofluidic oxygen sensing.

For imaging biofuel synthesis, two approaches were implemented. The first is single cell confocal imaging, for which a new imaging modality, vesicle photonics (VP), was developed and reviewed with colleagues. Employing VP principles, we are presently developing protocols for efficient fluorescent labeling of key metabolic signatures, such as intracellular lipids and photosynthetic organelles. Additionally, we are developing high sensitivity plant saccharification assays, using photonic approaches as well as microfluidic integration of cellulose and enzymes.

In FY 2014, we will apply the aforementioned methods directly to image biofuel synthesis at the single cell level. The latter will be performed using lipid synthesizing fungi and microbial and enzymatic cellulose saccharification. The integration of 'omic techniques such as transcriptomics and proteomics will be explored and evaluated in preparation of the final step:'omic analysis of pre-screened microorganisms by microfluidic and optofluidic platforms.



(a) An SIL-enabled microbioreactor in polystyrene; (b) An oxygen sensing microchannel enabled by SIL. The microchannel is the dark center, and the red-stained is the oxygen sensing areas; (c) An anerobic microbial community in a polystyrene microbioreactor; (d) Microfluidic trapping of a single photosynthetic cyanobacterium.

PN12005/2406

Predicting the Response of Complex Biological Systems

Karin D. Rodland

This project integrates data from multiple imaging and mass spectrometry (MS)-based technologies to predict complex biological system response to external stressors. In addition to furthering PNNL's capabilities in systems biology, we will explore applications relevant to human health.

Recent technical improvements in genomics, high-throughput proteomics and transcriptomics, and other analytical technologies provide biologists with powerful tools for describing the behavior of simple cellular systems. What has been missing is predictive capabilities that allow manipulation of biological systems for desired outputs in bioenergy, bioremediation, and mitigation of harm in response to stressors such as ionizing radiation. Achieving this predictive capacity will require both continued technological improvements in experimental systems and in analytical tools, including imaging and developing new computational approaches and mathematical models that allow molecular-level data to be scaled to cells, cell communities, tissues, and organisms. The goal is allowing us to predict, manipulate, and potentially design multi-cellular systems that contribute to DOE goals in bioenergy, contaminant fate and transport, carbon sequestration, and global climate change.

This project applies unique PNNL capabilities in 'omics measurements, imaging, and computational modeling and simulation to promote the transition of biological science from merely describing biological phenomena to predicting (and eventually controlling) the response of complex biological systems to perturbation, whether in environmental stressors, climate changes, or through deliberate manipulation to enhance performance (e.g., production of biofuels). Chief among the capabilities required are improved imaging systems that incorporate PNNL strengths in MS, nuclear magnetic resonance (NMR) spectroscopy, and emerging strengths in chemical imaging.

One of the specific subtasks of this project combines MS and NMR to provide a 2D image of biochemical changes in the brain. Integrating and interpreting data that arise from such combined imaging modalities will require application of novel computational tools for registering and adjusting multiple inputs. While existing research often focuses on simple model systems consisting of one organism or cell type, it is increasingly clear that predicting behavior of real biological systems requires model systems that incorporate realistic levels of cellular complexity while allowing accurate measurement of individual components. Application of systems biology tools to such complex model systems is another goal of this project. Finally, if the objective is prediction of response, then there is an urgent need to apply new computational tools that can simulate the behavior of complex biological systems accurately over time and space, and then extrapolate to predict the response to perturbation.

During FY 2012, predictive capabilities successfully developed included integrating 2D and 3D imaging technologies to understand spatial factors in biological function; identifying modifications to proteins that monitor the response to oxidative stress to predict changes in function or disease; developing unique nanomaterials for targeting medical isotopes; developing unique computational tools for integrating and interpreting complex datasets; and applying activity-based protein profiling to understand developmental changes in enzyme function. The specific tasks undertaken for FY 2013 and their outcomes are described below.

Multi-modal imaging of complex biological systems. We applied unique PNNL capabilities in imaging MS (nanoDESI and MALDI) to identify the specific brain regions affected by exposure to neurotoxicants. Improvements were made in the DTEM microscope under construction in EMSL to improve its applicability to biological systems.

Characterize the protein modifications indicative of inflammatory responses. A combination of MS and antibody-based technologies was used to identify protein modifications induced by environmental stressors.

Develop and test nanomaterials for delivery of targeted therapies. The ability of functionalized nanoparticles to deliver biological and radiological targeting agents was tested in vitro and in vivo.

Determine changes in detoxifying enzymes in early development to assess risk of fetal and neonatal exposure. We developed activity-based probes for the enzymes that detoxify polycyclic aromatic hydrocarbons and measured the activity of these enzymes in early and late neonatal stages and in mature mice, demonstrating a decreased detoxifying capability in the very young.

Development of advanced computational tools for the integration and interpretation of "big data" generated by 'omic technologies. A variety of computational tools were developed to increase the utility of large datasets generated by PNNL's 'omic technologies. These included improved tools for statistical analysis of missing data, knowledge environments for the seamless integration of different data types, and methods for simulating the complex factors regulating protein translation and turnover.

PN12061/2462

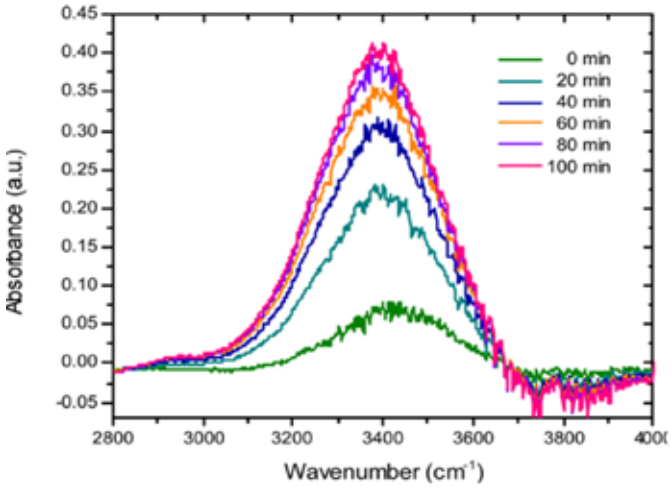
Predictive Chemical Bond Information in Living Cells

Thomas J. Weber

We developed a technology to interrogate live cells based on Fourier transform infrared (FTIR) spectroscopic measurements that can define changes in chemical bond patterns unbiased and non-destructively.

Many examples exist of emerging live cell applications for infrared (IR) technologies that correlate changes in chemical bond patterns with biological response. These patterns may or may not reflect specific chemical classes investigated, and methods to connect changes in chemical bond patterns with specific chemical classes are needed to address this knowledge gap. An approach being investigated in this project involved tagging chemical classes with deuterium as a contrast agent. Our test case focuses on aquaporins, a class of membrane proteins that function to traffic water in and out of cells. A real-time measurement technology for water mobility/trafficking would significantly benefit the aquaporin field. Water homeostasis is a critical feature of all biological systems, and a technology that can follow water movement will have important applications in eukaryotic and prokaryotic systems of importance to DOE mission areas.

Polar molecules strongly absorb IR radiation and water is the most polar molecule. Therefore, water absorption can be investigated as a definitive metric for how strongly a deuterium group functions as a contrast agent to improve interpretation of the FTIR spectroscopic measurement. To regulate water trafficking, we investigated aquaporin protein function. Aquaporins are integral membrane pore proteins that selectively conduct water molecules in and out of the cell while preventing the passage of ions and other solutes.



LLC-PK1 cells maintained in the IncubATR were treated with vasopressin and the water signal detected at 3400 nm was monitored over time. We observe a time- dependent increase in the peak at 3400, suggesting detection of vasopressin-stimulated water channel activity.

Our initial focus was on a regulated renal proximal tubule epithelial cell model expressing aquaporin: AQ2. This chemical functions as a vasopressin-regulated water channel that cycles between intracellular vesicles and the plasma membrane upon vasopressin stimulation. We obtained a renal proximal tubule epithelial cell line (LLC-PK1) that has been stable transfected with AQ2. LLC-PK1 cells grow to confluence and form tight junctions, which is an ideal model system expected to prevent interference from water contacting the crystal surface. In ATR measurements, the evanescent wave passes through the cellular layer but not the media and will therefore be highly sensitive to changes in intracellular water mobility/trafficking.

AQP2-GFP(NT) LLC-PK1 cells were treated with vasopressin, and redistribution of chimera from intracellular vesicles to the plasma membrane will be confirmed by epifluorescence microscopy. The plasma membrane localized AQP2-GFP retains water channel function and induces a marked increase in water flow following vasopressin treatment. We are monitoring the fingerprint region with emphasis on ~3400 cm^{-1} before and after vasopressin treatment, which is hypothesized to encompass the water signal. If this wavenumber reflects water mobility directly, then a marked increase in absorption at 3400 cm^{-1} will be observed. We are also monitoring all wavenumbers in the fingerprint region to identify additional peaks that may reflect increased water mobility/trafficking. In parallel, we are determining how deuterated water (D_2O) functions as a specific contrast agent for chemical bond measurements as previously demonstrated for coherent anti-stokes Raman scattering. Incorporation of aquaporin into the plasma membrane will reveal a clear temporal difference between aquaporin-mediated water trafficking and simple diffusion. This data will be used to identify the water peaks directly. We will then validate experimental findings in an identical experimental design using normal cell culture medium with and without vasopressin treatment to resolve water trafficking metrics in the absence of contrast agent.

PN13083/2564

Multiple project highlights achieved during FY 2013 included the following:

- A standard curve was generated comparing H_2O with D_2O using FTIR technology to determine detection limits. Water exhibited an absorbance at ~3400 nm and D_2O exhibited a spectral shift with peak at ~2500 nm. In quantitative terms, there was an inverse relationship
- between the concentration of H_2O and D_2O linear and proportional to the % D_2O in the standard. Based on this curve, it appears that as little as 0.1% D_2O can be detected by FTIR, which collectively suggests that there is both adequate sensitivity and spectral resolution for the proposed studies.
- The time required for data acquisition and analysis has been labor intensive and consumed a significant proportion of resources during developmental stages of this technology. A macro has been developed to automate data collection and data sets are now imported into MatLab for analysis. This has reduced labor needed for data acquisition and analysis time from a days/weeks time scale to a few hrs (~3).
- In live cell studies using the IncubATR technology, we observe an increase in the water signal at 3400 nm with time in LLC-PK1 AQ2 cells stimulated with vasopressin. This observation indicates that changes in water dynamics can be detected in live cell studies using the IncubATR technology.
- In studies using LLC-PK1 cells seeded on transwell inserts, we observe that D_2O passage across the monolayer is highly temperature dependent and significantly reduced at 4°C. Technical challenges with the model that are currently under investigation include constitutive plasma membrane localization of AQ2, which appears to result in a high water transport background.
- Ongoing studies at the time of this report are evaluating D_2O uptake mediated by AQ2 in live cell studies.

Our most important finding was the detection of changes in water dynamics in live cell studies using FTIR technology. If this result is properly validated, we will have defined a new approach for monitoring water transport by aquaporins, which will have important impact on mammalian aquaporin biology programs. Water trafficking is important to both eukaryotic and prokaryotic systems and technological advances that can facilitate water measurements such as under development here are expected to have broad biological applications in support of DOE mission areas.

Proteomics Measurements of Functional Redundancy and Stability Testing of Cellulose Degrading Anaerobic Microbial Communities Within Engineered Bioreactors

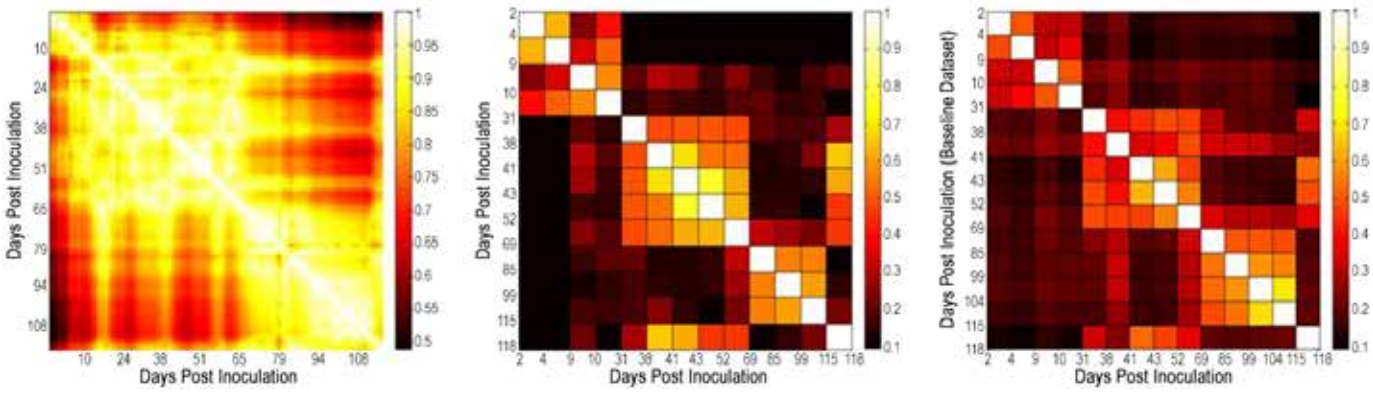
Stephen J. Callister

We are developing proteomics and bioinformatics capabilities for evaluating and monitoring microbial community populations relevant to biofuels production via insoluble cellulose degradation to measure the stability of cellulose degradation correlated to functional redundancy and environmental change.

Roughly 70% of plant biomass is composed of five- and six-carbon sugars, making it a primary resource for developing second-generation biofuels. In nature, an important service provided by microbial communities is the breakdown of cellulosic biomass. Understanding how microbial communities perform this process improves our knowledge about how biofuels are produced. Unlike a single microorganism, microbial communities work synergistically, often because of the diversity and redundancy multiple populations provide in a community setting. The enzymatic makeup and redundancy of the enzymes within a microbial community provide stability and synergism for the breakdown of recalcitrant cellulosic material. The focus of this research is the development of capabilities to measure the redundancy and dynamics associated with cellulosic degrading microbial populations extracted from a natural ecosystem but introduced into a bioreactor environmental setting.

During FY 2010 and FY 2011, 13 chemical probes were designed. Initial testing against commercially available enzymes showed that the probes are specific to enzymes capable of deconstructing cellulose and processing carbohydrates. Samples were collected over the period of bioreactor operation, capturing microbial population acclimation and establishment within the environment. In FY 2012, two cellulose-degrading bacteria, *Clostridium thermocellum* and *Fibrobacter succinogenes*, were chosen to test probe selectivity because of their different biological approaches applied to cellulose deconstruction. Complex protein mixtures secreted to growth media were extracted and labeled using each probe for *C. thermocellum* and a subset of probes for *F. succinogenes*. Samples of labeled and unlabeled proteins were then analyzed using mass spectrometry to supply protein identifications.

Results from the application of probes to *C. thermocellum* were recently published in FY 2013. Depending on the probe, a range of proteins selective for carbohydrates were observed with seven of these previously unknown to be associated with cellulose deconstruction and/or processing. The enrichment of these seven proteins using these probes



Three biological perspectives from the cellulose degrading microbial community housed within an engineered bioreactor monitored from start-up through perturbation and re-establishment of normal operation. The left is metabolite patterns from cellulose deconstruction and fermentation, the center is microbial population dynamics, and the right is microbial community protein profiles. Three distinct blocks are observed across all perspectives, denoting microbial community acclimation (smallest block upper left corner), establishment of dominant populations, and the impact of perturbation (lower block on right). Mechanistic links are being investigated between these three blocks to discover principles upon which better management of microbial communities within laboratory and industrial settings can be applied.

PN10069/2311

from the complex protein mixture could aid in hypothesis development concerning their role. Results from the application of the probes to *F. succinogenes* are ongoing.

The implementation of algorithms for analysis of liquid chromatography mass spectrometry probe data and global proteomics data derived from microbial communities in the engineered bioreactors were completed, and a scientific paper describing the algorithms was published in FY 2013. These algorithms are part of MultiAlign, a software tool to process liquid chromatography mass spectrometry data. Enhancements to MultiAlign include an alignment capability of mass and elution time features and a trace-back function that allows targeting of important mass and elution time features for tandem mass spectra data generation. Tandem mass spectra are critical for peptide sequence assignment and protein identification. These bioinformatics enhancements are significant as they allow proteomics analysis of probe-labeled proteins and complex protein mixtures from microbial communities without the initial need for genomic sequence information.

In addition to testing chemical probes against known cellulose-degrading organisms present in a liquid culture environment, tests were performed on the ability of the probes to label carbohydrate active enzymes within the soil environment. Crude extracts of the important cellulose degrading fungi *Trichoderma reesei* were added to a soil and successfully labeled. The action of native cellulases in the soil is a critical component in the cycling and sequestration of carbon. These findings were presented at an international conference focused on terrestrial ecosystems. Also during FY 2013, data generation from the cellulose degrading bioreactor was completed. While analysis of this data is ongoing, the implementation of developed capabilities have revealed interesting patterns of microbial community operation and function from several biological perspectives, including population dynamics; protein dynamics during initial bioreactor start-up, community acclimation, and perturbation; and metabolite dynamics. Redundant patterns between these three biological perspectives are being investigated, with findings currently being included in at least two peer-review manuscripts during FY 2014.

Pulmonary Injury from Acute Events Related to Nuclear Energy Production

Richard E. Jacob

This project will develop state-of-the-art imaging to enhance our understanding about the range of radiation exposure effects on pulmonary mechanics to facilitate development of timely, effective treatments in clinical and emergency response.

Catastrophic events such as that in Fukushima, Japan can result in significant radiation release and exposure to plant workers and large populations. Most studies of pulmonary radiation exposure effects have focused very long-term cancer occurrence from low-dose medical exams or acute tissue damage from high-dose radiation oncology. Few studies have examined radiation exposure impact on whole-organ mechanical function, which is critical to gas exchange. Recent work at PNNL suggests that lung disease that is not detected by conventional imaging or histology can induce mechanical changes that may be detected using a new dynamic imaging approach.

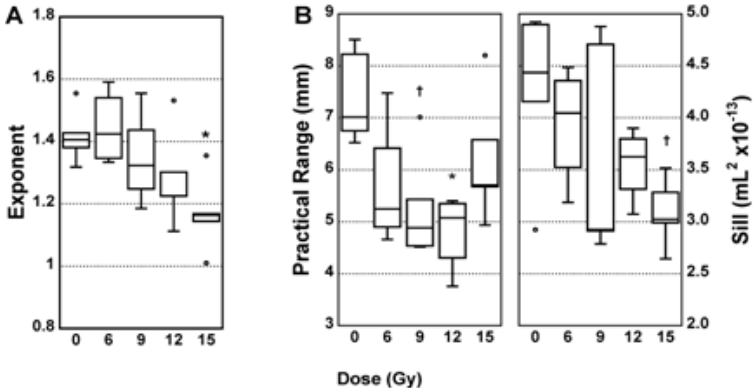
The objective of this project is to develop an improved understanding of the fundamental pulmonary mechanical changes due to radiation exposure. By studying animal disease models, we can begin to understand the effects of radiation on the lungs and develop a diagnostic approach translatable to humans critical to understanding the basic physiological response and disease course. We anticipate that this project will result in clinically translatable technical advances in dynamic pulmonary imaging and image analysis. This project will also develop a reliable animal model of lung radiation damage for future work of this nature.

During FY 2012, we standardized the radiation exposure protocol for rats, including development of contoured, rat-specific animal holders and a lead collimator to limit exposure to other organs. We also developed and refined methods of the CT image analysis, including non-rigid registration of image data for ventilation map calculation, which was published in the international peer-reviewed journal *PLOS ONE*. By the end of this first year of work, we found that pulmonary function tests and physiological measurements of body weight, lung weight, and the like generally did not correlate with radiation dose. However, close inspection of the ventilation maps indicated differences in ventilation homogeneity between exposure groups.

Few (if any) standard methods of assessing image homogeneity currently exist. Therefore, during FY 2013, we developed a new method to compare both the raw CT images and ventilation maps between exposure groups. This method combines a technique called "octree decomposition" – from computer science – and "variogram analysis" – from geostatistics. The application of this novel approach revealed differences between exposure groups that were not apparent with other measurements. In particular, a dose-response relationship was discerned through the very subtle differences in image homogeneity, as determined by the octree/variogram method, the results of which are being published in the journal *Academic Radiology*. We also successfully applied this technique to another animal model that was developed under a separate project.

Meanwhile, we are continuing our investigation in the rat lung. At the beginning of this year, we exposed 48 additional rats to 0, 6, or 15 Gy of radiation. The rats have been randomly divided into two groups: one was imaged at 6 months post-exposure and the second imaged at 8 months post-exposure. This study will be compared against the results of the first year study and importantly provide a second time-point (at 8 months) to help elucidate chronic disease progression. We have also added histological analysis of the lung tissue to this years' experiments to help illuminate the underlying changes to the lung tissue structure that were observed using the octree/variogram approach.

For FY 2014, we anticipate completing analysis of the data collected during this current year. We also plan to undertake a short-term study to investigate the acute effects of radiation exposure in our animal model by imaging the rats at between 2 and 4 weeks post-exposure.



A: Results from the variogram analysis of the raw CT images; B: The ventilation maps, showing a relationship between radiation dose and response as detected by variograms.

PN12065/2466

Quantitative Modeling of the Metabolism of Technologically Important Fungal Systems

William R. Cannon

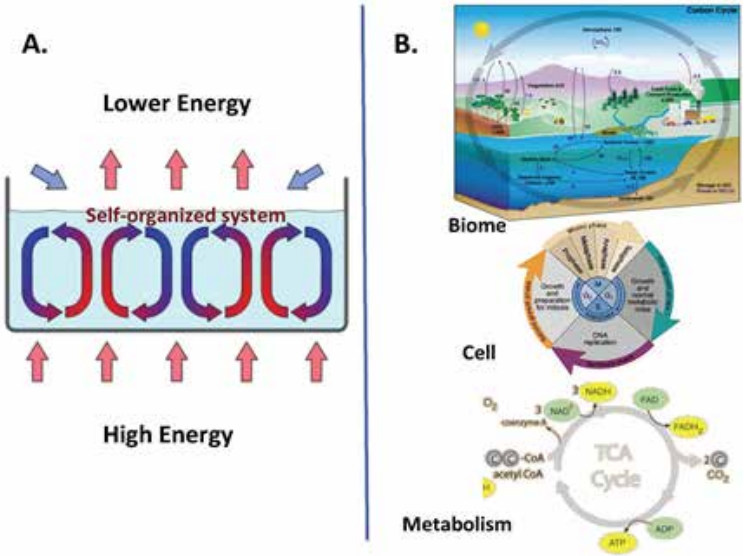
This project is evaluating a new simulation technology to address the challenges in engineering microbes for the production of industrially important chemicals and biofuels.

Currently, computational predictions are based on two approaches for modeling metabolism: flux-based approaches and kinetic simulations. Flux-based approaches are favored because they do not use rate parameters, which are required by kinetic simulation methods but are difficult to obtain. The basic idea in flux-based approaches, including cybernetic modeling, is to simply model flow through pathways. However, this convenience limits the predictive power of the methods in that the levels of chemical species (metabolites) are nearly impossible to predict from flux values, and the range of calculated fluxes consistent with the steady state assumption can span many orders of magnitude. Most importantly, flux-based methods do not provide any information on energy requirements of pathways, sets of pathways, or organisms in a community. Ideally, metabolism modeling would use kinetic simulations, but these simulations require knowledge of the thousands of rate parameters involved in the reactions. The measurement is labor intensive, hence rate constants for most enzymatic reactions are unavailable.

Specifically, we are constructing a state-based simulation model by which to compare other modeling methods. Important comparisons include the ability to predict metabolite levels as well as longer-term goals of integrating metabolite and proteome data for inference/prediction of enzyme activity, regulation, and expression. The practical goal is to construct and evaluate a model of an industrially important fungal system such as *Aspergillus niger*, *A. oryzae*, or *Trichoderma reesei*. Ultimately, we will use the model to characterize energetic requirements of the industrially important process of xylose utilization and bioconversion. Our approach involves modeling metabolic states rather than reactions. This new simulation technology provides predictions of metabolite levels and dynamic behavior, including relative reaction flux.

While calculating equilibrium constants is much easier than calculating rate constants, it is by no means trivial. We made significant progress in this regard by incorporating an automated procedure for calculating equilibrium constants into our code. This greatly streamlines the modeling process, as the user simply provides a file containing the reactions to be modeled with the reactants and products described using KEGG identifiers or standardized chemical names. At this early stage of development, we believe that the current approach is adequate. This work will be published in early FY 2014.

Our approach to modeling the central metabolism of fungal systems is to build the entire network by adding and testing each pathway independently. We completed testing of the pentose phosphate, glycolysis, and gluconeogenesis pathways. We have taken this approach because it became clear early on that simply having all the rate or equilibrium constants necessary for simulating a system is not enough. Biological systems do not achieve steady states as described in textbooks – by selective balancing of rate constants and concentrations of chemical species. Rather, a stable reaction system is achieved through regulation. Consequently, as we construct the central metabolism of fungi, we first evaluate the stability of individual pathways



Dissipative structures: A. A convection cell in which fluid becomes self-organized to dissipate energy efficiently from the high-energy source to a lower energy reservoir; B. Dissipative structures are common in nature, but living systems such as microbiomes are seldom modeled as such. Viewing living systems from this perspective will enable us to predict the optimal pathways for engineering cells to maximize the production of a target chemical.

PN13068/2549

and then combine these pathways into a larger system. Modeling stability in isolated, linear pathways has been straightforward – the flow of metabolites through the pathway is a mostly matter of non-equilibrium thermodynamics.

Cycles are ubiquitous in biology, and the ubiquity reflects their fundamental importance. The ability to model auto-catalytic systems using equations of state has made it apparent that biological systems are dissipative structures that form in nature when the environment is far out of equilibrium. These structures form because they relieve the stress in the environment that is a result of energy being input into the environment, either from sunlight or geo-chemical inputs such as thermal vents, with no place to go. Yet biological systems are seldom modeled with energy dissipation in mind as a primary driver of function. The biofuels literature, for example, has exceedingly focused on regulation as the primary hurdle to overcome when attempting to synthesize high-energy carbon compounds from lower energy components. Energy dissipation is the driving source for not only production of highly oxidized cell "waste" products but also for interactions within microbial communities, plant-microbe interactions, and interactions across levels of the food chain.

We now suspect that when a xylose utilization pathway is introduced into a fungal system, the xylose isomerase reactions may form a dissipative structure, essentially short-circuiting the pathway for producing biofuel precursors. This hypothesis will be among the first one tested when construction and testing of the central metabolism pathways are completed. Pending funding, we anticipate that this work will be completed in FY 2014.

Signatures of Environmental Perturbation - Microbial Community and Organic Matter Resilience

Vanessa L. Bailey

The overarching goal of this project is to use existing capabilities to discover one or more signatures that indicate impaired resilience of the soil carbon biogeochemical system following a system-wide perturbation such as simulated climate change.

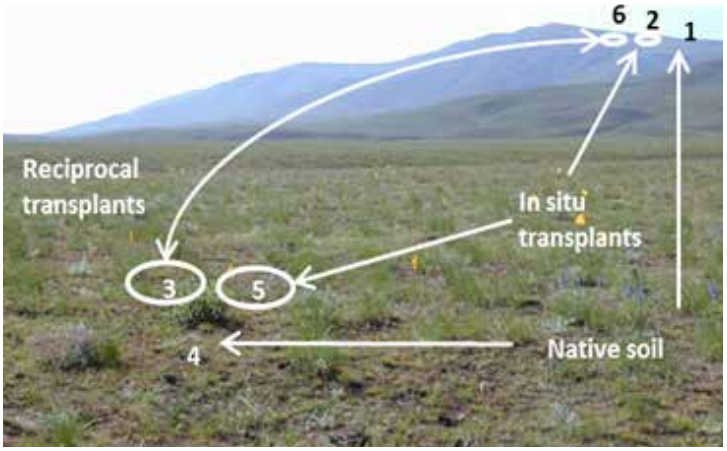
The microbial metabolism of soil carbon leads to the release of the greenhouse gases into the atmosphere, which directly impact global climate by exerting a strong positive feedback effect on temperature (i.e., global warming). Currently, no tools assess the vulnerability of soil carbon reservoirs to changing climate conditions, but it is likely that microbial metabolism of soil carbon will increase with temperature, resulting in even greater gas emissions. The structure and function of the native microbial community is intimately linked to soil carbon by both the deposition of new soil carbon and respiration of existing soil carbon as part of the terrestrial ecosystem carbon cycle. By integrating the key chemical and molecular signatures of the soil system, microbial community resilience and soil carbon reservoir vulnerability may be predicted with molecular resolution. Thus, the potential for more greenhouse gases to be released from native soil carbon can be better predicted as well, under these changing environmental conditions.

The focus on ecosystem stress and climate change is currently relevant, as researchers and policy-makers strive to understand the ecosystem consequences of climate change. Successful development of chemical/molecular profiles that link soil microbiology with soil carbon to ascertain soil vulnerability and resilience would have great impact on assessments of soil ecosystems in response to global climate change. Additionally, these integrated signatures could be used to support the design of sustainable agricultural and food/energy crop security practices. These efforts align directly with research performed by DOE, USDA, and the National Science Foundation. Ecosystem stress associated with pollution events may be of interest to multiple other federal agencies.

This project aims to move soil biogeochemistry and ecology toward a more informed understanding of the relationships between carbon chemistry, microbial community structure and function, and ecosystem perturbations. The demonstration that our analytical tools can be integrated to identify such signatures in a complex system such as soil has the potential to transform classical ecological research.

A late FY 2013 start, this project's activities focused on integrating current capabilities into the experimental plan. Our research leverages an existing long-term experiment at the Arid Lands Ecology Reserve in Richland, WA in which soil cores were reciprocally transplanted between the cooler, moisture upper slope position and the hotter, dryer lower slope position in 1994. This transplant was designed to mimic the stress of climate change with natural temperature and moisture gradients provided by elevation. A previous study of these soils has suggested that one of these core soils has lost robust microbial community function as a result of this transplant.

Researchers are evaluating the field experiment and the soil respiration data from the previous study to identify the subset of samples needed to strategically inform new, more highly resolved experiments and analyses. We prepared soil DNA from the previous study, and the full metagenome sequences will be available later. We also commenced characterizing the intrinsic soil carbon chemistry in archived soil samples from that study using high resolution mass spectrometry. A focused integration and analysis of these data in FY 2014 will further inform future sampling of the field site and direct the research toward follow-on, more refined chemical and biological analyses.



Schematic of the Rattlesnake Mountain reciprocal transplant experiment overlaid on a site photograph. The six soil types that will be examined are: 1) native upper soil; 2) upper soil transplanted in situ; 3) upper soil transplanted to the lower site; 4) native lower soil; 5) lower soil transplanted in situ; and 6) lower soil transplanted to the upper site.

PN13102/2583

Structure and Dynamics of Biological Systems

James E. Evans

Through this project, the ability to observe "live" biological systems across multiple platforms with various spatial, temporal, and chemical resolutions will provide unique insight for these model systems and validate the versatility of this combined approach.

As biological systems are highly complex, understanding how a single cell protein is related to the function of a whole organism requires spanning more than 10 orders of magnitude across spatial scales. This is a large feat on its own, but a complete picture is possible only if methods are added to probe the chemical and temporal evolution of the system to observe how it changes over time. Because no single instrument is currently capable of addressing structure, dynamics, and chemistry across these scales simultaneously, new methods integrating data from multiple instruments are required to achieve a clear understanding of biological processes relevant for the production of energy, cleanup of environment and improvement of human health.

Since the 1950s, biologists have strived to link biological system structure of to their physiological function. Yet countless details still remain obscure for most research foci because biological systems are highly complex and their processes occur over a wide range of spatial and temporal scales from femtoseconds to hours and angstroms to meters. In this project, we are developing, adapting, and employing state-of-the-art approaches toward enabling a better understanding of biofilm organization, enzymatic energy transduction, and epithelial cell interactions with nanomaterial. The overall objective is to integrate six emerging technologies to permit multimodal and multiscale spatial, temporal, and chemical analysis of biofilm organization, enzymatic energy transduction and nanotoxicology. We will combine dynamic transmission electron microscopy with femtosecond X-ray diffraction (XRD) at the linac coherent light source (LCLS), secondary ion mass spectrometry (SIMS), and atom probe tomography to interrogate the structure and dynamics of biological systems.

FY 2013 was the first year of this project, and considerable achievements have already been made. We had six published or in press manuscripts, with two additional manuscripts submitted for review. We also had eight invited presentations to domestic and international conferences along with six contributed talks and posters. While the publications help with the dissemination of our work to the larger scientific field, the conference talks help raise the visibility of the project and foster new scientific collaborations.

During FY 2013, significant progress has been made in each of our specific project aims. Highlights include acquiring the first transmitted XRD patterns from 2D protein crystals using LCLS, using nano-SIMS and in situ time-of-flight SIMS to study biological samples with chemical sensitivity, and developing novel methods for atom probe tomography to detect the individual protein macromolecules embedded in a resin matrix. Both the developments with biological atom probe tomography and femtosecond XRD of 2D crystals have never before been demonstrated. In addition, we developed a novel in situ liquid cell compatible with femtosecond XRD experiments to image the 2D protein crystals in a fully hydrated state in order to permit pump-probe dynamic observations.

In FY 2014, we will continue developing a multimodal and multiscale platform for interrogating biological systems. In particular, we plan to build off of the development phase that we performed during FY 2013 and implement a more integrated approach by looking at the same sample using multiple techniques. First, we will design and fabricate new *in situ* liquid cells for X-ray, electron, ion and optical imaging that will allow a single sample to be transferred between each technique. In parallel, we will adapt nano-SIMS for robust serial section analysis of biological systems and correlative microscopy. Previously, nano-SIMS has remained limited to observations of representative slices from a cell rather than each slice in consecutive fashion. The new approach should provide a comprehensive sliced 3D chemical map.

We will also extend our capabilities in atom probe tomography by implementing a cryo-transfer system that will allow us to run samples in both a room temperature embedded state or a cryogenic state. The benefit of running at low temperature is that the freezing process can be done in such a way as to prevent crystallization of water and optimally preserve the sample in its native state. Finally, we plan to evaluate the synergy of femtosecond X-rays and micro- to nanosecond electrons for acquiring movies of biological processes in real-time including conformational changes upon ligand binding and assembly mechanisms. This will maximize the new methods we developed with dynamic transmission electron microscopy and at LCLS using femtosecond XRD. These focus areas for FY 2014 will establish the interconnected framework to begin probing biological systems with high fidelity across multiple platforms.

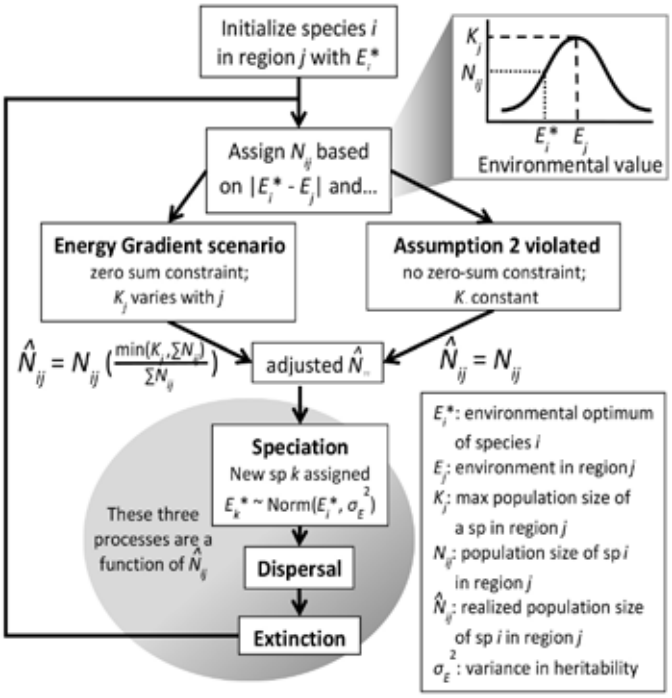
PN13009/2490

Understanding the Processes that Govern Subsurface Microbial Communities

James C. Stegen

This project ultimately aims to develop the theoretical tools capable of predicting microbial community composition and function through space and in response to environmental change. The resulting tools will be crucial for improving environmental and human health through remediation of contaminated sites across the United States and beyond.

Microbial communities play a central role in the functioning of natural ecosystems by heavily influencing biogeochemical cycles. It comes as no surprise that microbes can influence human-derived environmental contaminants and efforts to remediate contaminated locations. It has also been repeatedly demonstrated that ecosystem function (e.g., elemental cycling) is strongly influenced by the composition of ecological communities. As such, if we are to understand how microbe-influenced contaminants move through ecosystems and how effectively to remove or stabilize contaminants, it is vital that we initially understand what governs the composition of microbial communities.



Flow chart illustrating two eco-evolutionary simulation scenarios. Population size varies with the match of a species to its environment based on a trait characterizing environmental optimum. For the baseline energy gradient scenario in which a zero-sum constraint exists, population size is modified by the presence of other species in the region and their relative fitness such that the maximum regional abundance does not exceed Kj. To test zero sum constraint importance (Assumption 2), the simulation is conducted without a zero sum constraint in which species abundances are independent of other species' presence.

Historically, ecological community research is descriptive (e.g., how many species are in a given area and how the number of species changes across locations). Current research focuses on the processes that govern observed patterns. Although useful, most contemporary studies only infer the dominant class of process. This approach fails to describe the constituent processes quantitatively necessary for developing predictive models. This project fills that void by coupling observational, experimental, and stochastic simulation approaches, using the Hanford 300 Area subsurface as a model system.

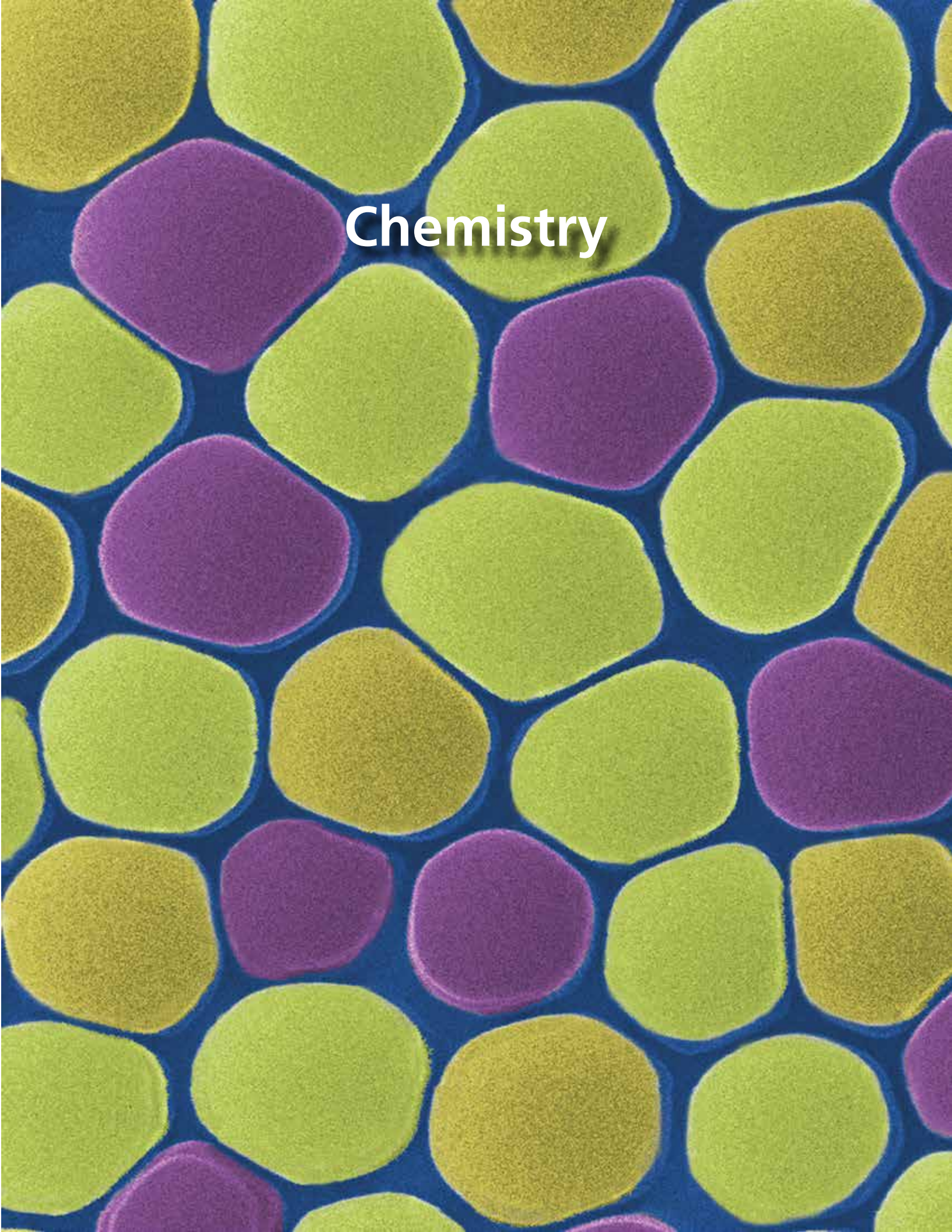
Work that began in FY 2012 set the stage for progress in FY 2013 in the following ways:

New statistical tools were developed that quantify the influences of ecological processes and identify unmeasured variables that influence microbial communities were extended to quantify an additional process class. In addition, application of the extended framework to the Hanford 300 Area subsurface led to a peer-reviewed article in a top publication, *ISME Journal: Multidisciplinary Journal of Microbial Ecology*, and to a manuscript currently in review at the journal *Ecology*.

New collaborations focus on understanding factors that govern microbial communities were leveraged to generate two additional publications in *ISME Journal*. From another collaboration established in FY 2012, a new simulation model has been developed that generates diagnostic fingerprints of ecological and evolutionary processes that govern spatial biodiversity gradients. This new model is general enough to be used to gain process-level insights from different ecosystems (e.g., subsurface, terrestrial, aquatic, or marine environments) and different taxonomic groups (e.g., bacteria, plants, or animals). A manuscript describing the new model is currently in review as an invited paper at *Ecology Letters*, and field data needed to run the simulation model for microbial communities will collected at the end of FY 2013.

Design and execution a field campaign in the Hanford 300 Area from last year laid the foundation for a proposal in review at the Joint Genome Institute (JGI). Last year's field study also helped to yield a substantially larger and currently ongoing field campaign across the river water-groundwater interaction zone of the Hanford 300 Area. Samples from this larger campaign will be studied using a variety of molecular-based analyses by leveraging resources at both JGI and EMSL. Work in FY 2014 will be primarily focused on analysis, interpretation, ecological modeling, and publication of the resulting data.

PN11078/2400


Chemistry

Bio-Inspired Actinides Recognition for Separation Science

Ping Yang

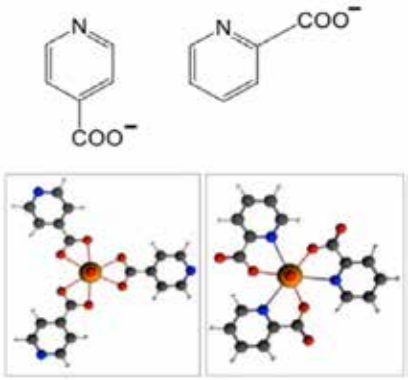
We are developing novel, more efficient strategies for selective separation to allow for better reprocessing of spent nuclear fuels that reduces the need for long-term waste storage.

Current techniques for the separation of transuranic actinides from lanthanides rely on multi-state liquid-liquid extraction. This approach is challenging due to the very similar chemical properties of actinides and lanthanides: they are all typically present under trivalent oxidation states in acidic solution. The fundamental understanding of the chemical and physical properties of actinide complexes is critical to design chelating molecular systems rationally and with high selectivity. Nature demonstrates that it is possible to engineer systems with remarkable molecular recognition capabilities among metal ions (e.g., potassium channels show a selectivity factor for K+ over Na+ as high as 104:1). The possibility of mimicking such natural design to improve the selectivity to actinides by collectively incorporating multiple binding sites with O, N, and S remains to be explored.

We propose to build a fundamental understanding of the physico-chemical interactions of binding motifs for actinides and use that understanding to design chelating reagents with high selectivity to actinides. The research will be based on two main thrusts: the thermo-structural investigation of actinides-bio-ligands complexes and the tuning of redox potentials of actinides compounds through interactions with ligands. Specifically, we are studying interactions between actinides and binding sites in select biomolecules using first principle electronic structure methods that include relativistic effects and electron correlation to validate calculations via gas phase collision-induced dissociation experiments and solutions.

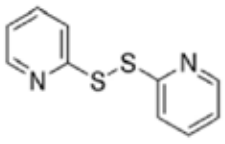
Reagents suitable for use in the separation process are typically based on bidentate chelating hard O-donor ligands. Recently, ligands composed of soft N- and S-donor atoms became the focus of research in tackling the challenging AnIII/LnIII separation. It is hypothesized that the interactions between soft donor ligands and 5f orbitals in actinides are more covalent than with 4f orbitals of the lanthanides. The exact origin of this covalency remains poorly understood and is the subject of ongoing debate. Fundamental theoretical studies are needed to further the understanding of bonding interactions.

N-donor ligands. To study soft N-donor ligands in actinide bonding interactions, we focused on a bidentate chelating ligand with hard O- and soft N-donor heterocyclic groups: alpha and gamma-picoline carboxylates. These ligands bind to transoxo complexes of An(VI) in different fashions as elucidated by our first-principle calculations using relativistic density functional theory (DFT) methods. We performed the analysis and computed binding/dissociative energies, including scalar and spinorbit coupling. The thermodynamically favorable structures of alpha picoline carboxylate actinyl complexes are O- and N-coordinated, in preference to bidentate binding with two oxygen atoms. The An-N distances are shorter compared with An-O bonds across the actinide series. In contrast, the stable configuration for gamma-picoline carboxylate is bidentate with the carboxylate group in the equatorial plane. Our DFT calculations show that the heterocyclic amine is a strong donor to the actinide center and more favorable compared to the hard O-donor. The calculated results were further verified by gas phase collision induced dissociation (CID) experiments from our Lawrence Berkeley National Laboratory (LBNL) collaborator. A manuscript describing these results is in preparation.

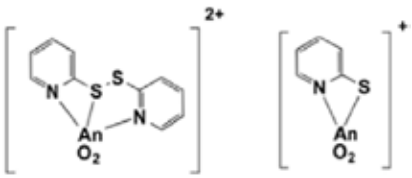


N-donor ligands showing the alpha- and gamma-picoline carboxylate binding to AnO_2^{2+} in different fashions. N,O-coordinated complexes are thermodynamically more favorable than the O,O-bidentate configuration.

S-S Bond Activation. Sulfur-sulfur (S-S) bond is postulated to reduce actinyl ions during the cellular biological remediation process. Therefore, we choose 2,2'-dithiodipyridine (DTDP), which contains one S-S bond that can reduce actinides under mild chemical conditions to study the role of soft S-donor ligands



2,2'-Dithiodipyridine



Both N and S coordinate to AnO_2^{2+} forming tridentate complexes followed by S-S bond activation and dissociation.

PN13070/2551

and the redox between actinyl ions and S-S. Relativistic DFT calculations show strong interactions between actinyl ions (AnO_2^{2+}, An=U, Np, Pu) and DTDP. We found that both N atoms and one S atom in DTDP coordinate to the actinide center, forming stable tridentate complexes and followed by S-S bond activation. Gas-phase CID experiments confirmed the dominant S-S bond cleavage in good agreement with theoretical calculations. As with the N-donor ligands work, a manuscript describing this bond activation work is in preparation.

Our calculations have shown that the heterocyclic N-donor is favorable with coordination ligands in preference to the hard O-donor group. We also verified that actinyl ions can coordinate to S atoms to activate the S-S bond. These results give us confidence to investigate our theory that redox potentials of actinide compounds can be tuned through interactions with chelating ligands containing multiple binding sites, where N, O, and S can coordinate to metal ions cooperatively. We will further study the ligand competition between soft S-donor and hard O-donor ligands. In addition, we will study the interactions of actinides with a redox-sensitive compound as a function of soft N- and S-donor atoms. Building on the above results, we will further our study in FY 2014 by validating the theoretical results using gas-phase CID and solution chemistry through a collaboration with LBNL and Washington State University.

Characterization of Catalyst Materials in the Electron and Atom-Probe Microscopes

Ilke Arslan

With our current global business economy, fuel is essential for car, plane, and boat transportation. The catalytic materials and processes studied in this project provide renewable and conventional energy resource conversion for less expensive, more environmentally friendly fuel.

This project aims to advance and combine nanoscale techniques to further the examination of zeolite catalysts. The characterization-synthesis loop is an important aspect of designing and controllably synthesizing a material of industrial significance. Likewise, characterization on the atomic and nano scales allows for fundamental understanding of the catalytic processes that aids in designing a better catalyst. We employ atomic resolution, 3D imaging, and in-situ fluid imaging in the scanning transmission electron microscope (STEM) combined with atom probe tomography (APT). Chemical information can be obtained in the STEM using electron energy loss spectroscopy.

Much of the experimental analysis in this project will make use of the Z-contrast imaging technique in the STEM. To form a Z-contrast image, the scattered intensity (electrons) is collected over a large angular range on a high angle annular dark field detector, averaging coherent effects between atomic columns in the specimen. This allows each atom to be considered to scatter independently with a cross section approaching a Z^2 dependence on atomic number (hence the name Z-contrast imaging). This detection geometry yields an incoherent image where changes in focus and thickness do not cause contrast reversals in the image (unlike conventional TEM) so that atomic sites can be identified unambiguously during the experiment.

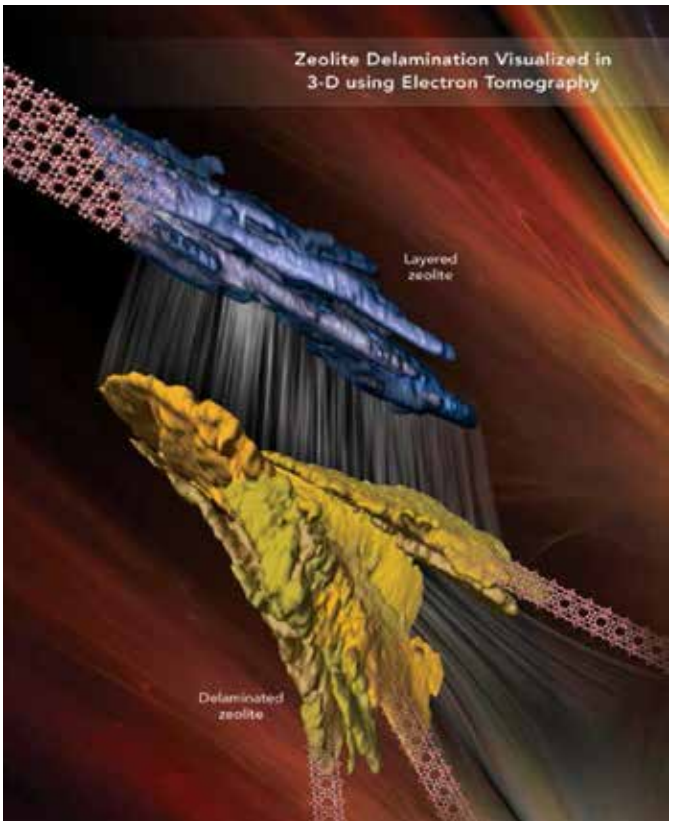
In electron tomography (ET), a series of images are taken at different tilt angles and are combined using reconstruction algorithms to form a 3D representation of the nanomaterial. The application of ET to inorganic materials using STEM has developed rapidly. A fairly large range of volumes of material can be analyzed in 3D in configurations as small as $10 \times 10 \times 10$ nm^3 to as large as $300 \times 300 \times 300$ nm^3, with the 3D resolution varying from just under 1 nm to nearly 5 nm depending on the volume size, number of images acquired, tilt range achieved, and alignment and reconstruction methods. APT is a controlled field evaporation technique that produces sub-nanoscale spatial and concurrent chemical resolutions. However, a number of artifacts can limit the overall resolution. Knowing the 3D structure of the material volume a priori by APT is therefore crucial. The most accurate way to do this is by matching in 3D with a non-destructive technique such as ET.

The artifacts in APT and ET are in different directions; therefore, the artifacts of each method can be understood and compensated. These methods are combined and applied to two sets of zeolite materials: MCM-22 (and related layered zeolites that can be delaminated) and ZSM-5. MCM-22 is used as a precursor material for delamination, with the goal of pulling apart the zeolite sheets to provide a higher surface area while keeping the active sites on the surface intact. The higher surface area allows access for larger bulky molecules to create reactions, opening up the materials for new processes such as cracking, alkylation, and 3D delamination. ZSM-5 is being studied as the acid catalyst for converting lignin-derived phenolic oil to produce gasoline range hydrocarbons and methanol. As this reaction occurs in water, it is important to understand the fundamental morphology and chemistry of the zeolites in water. We are applying a fairly new technology of imaging through fluids in the electron microscope to understand the structure of the zeolites in its aqueous environment. High resolution imaging and 3D ET will be combined with APT to understand the location and density of acid sites with nanometer spatial resolution.

The ultimate goal of delaminating zeolites is to investigate the formation of bimetallic clusters on supports from complexes of metals that are chemically similar and have a strong tendency to form clusters. The support for the metal clusters will be zeolites because its support allows precise control of the structures of supported metal clusters as a consequence of the regularity of their surfaces and cages and the opportunity to control the number of well-defined bonding sites (the metal precursors tend to react at the Al sites, as evidenced by EXAFS spectra). One of the challenges of zeolites is determining the metal cluster structures throughout the samples. An appealing alternative is zeolites in the form of extremely thin layers as support for our clusters. We expect to obtain high quality 3D and atomic-resolution samples and cluster images (even of the zeolite frame), as all the clusters will be located near surfaces. To realize this goal, the zeolite support must be fully characterized first to understand the morphology and determine whether synthesis is successful.

PN12015/2416

To accomplish our objective, we performed a thorough analysis of the 3D morphology of the zeolite at three stages of chemical and heat treatment to understand the delamination mechanism. The starting material was MCM-22, a precursor material. STEM tomography was performed on this material, and many thin sheets of zeolite were found bound together in large flat clusters. The distance between sheets is the pore size, which limits larger bulky molecules from entering. The next step is a chemical treatment, which in this case is performed at pH 9 using an aqueous solution at 353 K. This is a much lower pH than zeolite delamination in the literature, which yields no detectable amorphous areas that inhibit catalytic performance as those found in materials such as ITQ-2. The 3D imaging performed at this stage shows that the zeolitic layers are still flat and in clusters, but the distance between layers is slightly larger, suggesting swelling and breaking bonds.



The resulting work of our efforts to show the visualization of delamination was chosen as the program cover at the 2013 North American Catalysis Society Meeting in Louisville, KY.

At the final stage of synthesis, the material is calcined. The heat treatment does something phenomenal in that the zeolite sheets come apart in curvy layers and the surface area increases. The distance between clusters of the curvy sheets is larger (up to several hundred nm), thereby providing ample space for bulky molecule reactions. To date, there has been no published literature on a direct and 3D visualization of the delamination process. We performed the experiments and some reconstructions in FY 2012 and have used FY 2013 to focus on the new reconstruction algorithms. Using the discrete algebraic reconstruction technique (DART), compressive sensing, and other conventional methods, we optimized the reconstruction procedure for the delaminated zeolites.

In FY 2014, we will extend our analysis to a different set of delaminated zeolites in which the surface area is shown to increase (by other methods) greatly upon delamination, and we will perform the electron tomography quantitatively to show a direct correlation between surface area measurements on the nanoscale with surface area calculations from microns of zeolite. Whereas the last 2 years had MCM-22 as the starting delamination material, we will use SSZ-70 in FY 2014.

With zeolites for biomass conversion in FY 2012, we provided the first images of the ZSM-5 zeolite imaged in-situ in water. We found that when suspended in water, particle cluster size is smaller than when the material is dried. Upon ZSM-5 analysis on a dry grid with different temperature treatments, we found that the sample boiled to 250°C is more stable under the electron beam. The instability of zeolites under the electron beam is generally attributed either to the presence of trapped water in the pores or unstable cations in the structure, but this hypothesis has not been proven. For FY 2013, we continued imaging these zeolites in water and are using heating elements with the liquid holder manufacturer to perform in-situ leaching experiments. We have been imaging 50–100 nm ZSM-5 crystals in NaOH instead of a water solution, which should initiate leaching when heated. We are hoping to image the process of desilication in real time and show how the zeolites progress from solid to hollow particles, which has never before been shown. Heating in solution is new technology for the flow holders.

Lastly, we have been working with UOP and Utrecht University on studying features in ZSM-5 crystals with varying steam treatments to understand how the distribution of the acid sites changes with steaming. This is a challenging task that requires chemical and 3D morphological information. We are therefore correlating STEM tomography (morphological information) with atom probe tomography (chemical information) to use each method to minimize the artifacts in the other and to build an accurate representation of acid site distribution in these materials. We already demonstrated the morphology and chemistry of the materials with each technique separately and are currently focusing on the analysis of the very same sample with both methods. This correlation has never been done any material other than metals, and would provide much valuable information for the catalysis community.

PN12015/2416

Chemical Imaging Analysis of Environmental Particles

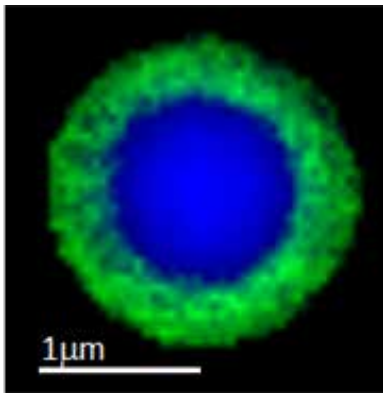
Alexander Laskin

To address concerns about the safe production and use of energy, it is necessary to develop and apply novel experimental approaches for an improved understanding of the origin, fate, chemical and physical properties of environmental particles.

The production and use of energy often leads to the generation of atmospheric particles that have a dramatic impact on a broad range of environmental issues. Environmental particles are typically multi-component; therefore, a detailed understanding of their complex chemistry is necessary and will reduce uncertainties related to their environmental effects. The application of complementary analytical methods is needed to provide comprehensive information ranging from microscopy level details of individual particles to advanced molecular characterization of complex molecules comprising particulate matter. Particular attention is being given to the application of a range of analytical techniques to experimental studies of heterogeneous gas-particle reactions pertinent to atmospheric and occupational environments, technologies related to energy production and emissions control, and particle health effects.

This project focused on the development of complementary experimental approaches for chemical imaging analysis of aerosol particles in realistic atmospheric environment. The research and development efforts were carried out along two major directions: the development and application of micro-reactor cells for in-situ studies of gas-particle reactions using complementary methods of spectro-microscopy; and the development and application of particle depth-profiling methods by single-particle mass spectrometry.

In FY 2011, we evaluated depth-profiling by single particle mass spectrometry, where data indicated that soot particles generated at different temperatures have remarkably different physical and chemical properties. For FY 2012, we worked with Advanced Light Source (ALS) beamline scientists on conducting laboratory studies focused on the chemistry of individual particles and their effects on hygroscopic properties, heterogeneous reactivity, and ice nucleation properties. Continuing our collaboration with ALS in FY 2013, we utilized the micro-reaction cell developed at the beginning of the project for in situ X-ray microscopy studies of hygroscopic behavior of model aerosol particles



Singular value deconvolution map of a particle of mixed AS/LSOA composition imaged using X-ray spectro-microscopy at atmospherically relevant conditions of 70% RH and T= 25C. Colored areas indicate liquid-liquid phase separation of SOA (green) and AS (blue) components.

with binary composition of atmospheric relevance: AS (ammonium sulfate)/LSOA (secondary organic aerosol from ozonolysis of limonene), $NaNO_3$/LSOA, and NaCl/LSOA. In these experimental studies, we monitored the separation of liquid phases in individual particles and quantified the extent of these effects as a function of relative humidity. These observations challenge present atmospheric models that assume homogeneous internal composition of wet (deliquesced) particles and cloud micro droplets. The effects of liquid-liquid phase separation change the intrinsic properties of particles' surface and their internal composition, and in turn may affects optical properties and particle-cloud interactions of atmospheric aerosol.

Also during FY 2013, we developed and applied a new depth-profiling approach to characterize the morphology of individual particles, which is based on infrared laser-assisted layer-by-layer evaporation followed by soft vacuum ultraviolet or ultraviolet-laser ionization. We applied this method to characterize the morphology of SOA particles formed in the presence of poly aromatic hydrocarbons (PAHs) and PAH-coated SOA particles. We found that the PAH coating remains on the surface and hence evaporates rapidly, whereas PAHs acquired during SOA formation become incorporated and trapped inside of viscous, nearly non-volatile spherical SOA particles. These results provide the first explanation of long-range transport of PAHs and other toxins far away from the sources into pristine environments.

In addition, we generated and studied the evaporation of SOA particles coated with liquid hydrophobic dioctyl phthalate (DOP) and DOP particles coated SOA. We demonstrated the stability of both morphologies and that the DOP surface coating evaporates fast, while SOA-coated DOP does not. These results explain the observations that some organic compounds like oleic acid have atmospheric lifetime orders of magnitude longed than predicted.

PN11012/2334

Conversion of Biomass to Jet Fuels

Karthikeyan K. Ramasamy

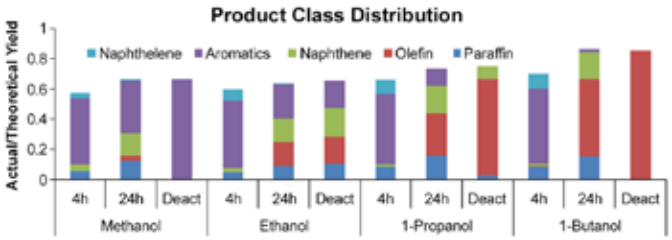
This project provides an understanding of the catalytic conversion of biomass-derived alcohols to aviation fuel range hydrocarbons, utilizing the fundamental knowledge generated to design catalysts that produce full performance jet fuels from a range of biomass-derived low value oxygenates.

Currently, all aviation fuel is derived from fossil sources. For aviation and heavy transportation, liquid fuel is the only viable source for the foreseeable future. Because of high interest in renewable fuel resources, private industries are involved in developing a process to produce jet fuel from natural oil, algae resources, and biomass derived oxygenates. Presently, jet fuel requires roughly 40 wt% of aromatic and cyclic content along with approximately 60 wt% of C_8 to C_{16} paraffinic (straight and branched) hydrocarbon due to stringent specifications such as low freezing point, high flash point, and high volumetric density. In addition to unfavorable cost, these processes produce only paraffinic content and depend on other resources for aromatic content.

In this project, we studied the elementary steps for the alcohol-to-hydrocarbon conversion using a similar concept to the methanol-to-gasoline (MTG) process developed by Mobil. We have probed the nature and design of catalysts with controlled morphology and structure to achieve the target activity and selectivity toward the product slate to meet the stringent jet fuel composition. Successful outcomes from this project not only lead to producing higher hydrocarbons from alcohols but also generate the knowledge base that can be used to develop processes to produce higher hydrocarbons from a wide range of biomass-derived low value oxygenates.

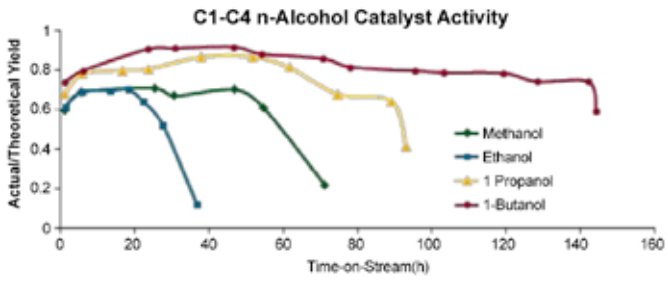
In FY 2011, our focus was converting ethanol to aromatic and the cyclic content of jet fuel fraction. Experiments were designed to understand the nature of catalysts and the primary reaction mechanisms to achieve high jet fuel fraction with 100% carbon efficiency. During FY 2012, the primary focus was improving the straight and branched chain hydrocarbon selectivity and reducing the unsaturated cyclic compounds generated from ethanol to fit the jet fuel composition. It was identified that the distribution between the product class changes with time-on-stream (TOS).

During the past year, we conducted experiments to develop the catalyst activity and deactivation mechanisms of C1-C4 n-alcohols over the HZSM-5 catalyst in producing jet fuel range hydrocarbons. For all alcohols tested, the product class distribution between paraffin, olefin, naphthene, aromatic, and naphthalene compounds for samples collected at 4 h TOS were similar. The biggest difference was in the yield of liquid hydrocarbon. The smaller alcohols tended to generate higher gas fractions in LPG range not condensable at room temperature. This was a primary reason for the difference in liquid hydrocarbon yields between alcohols in the early stages of the experiment.



Product class distribution of the liquid hydrocarbon sample collected at different TOS from the conversion of methanol, ethanol, 1-propanol, and 1-butanol conducted over HZSM-5 at 360°C, 300 psig.

In all cases, distribution between the product class changed for 24 h TOS samples compared with 4 h TOS samples. The main differences are a decrease in aromatic and naphthalene composition and an increase in the olefin, naphthene, and paraffin compounds. The liquid analysis of the samples collected at deactivation time of methanol contains only aromatic compounds (primarily 1,2,4,5-tetramethylbenzene), and the sample collected from the 1-butanol at the deactivation time primarily contains the straight branched chain olefins between C_6 and C_{11}. The catalyst activity and life time were different for alcohols with different carbon numbers: activity for ethanol dropped more quickly, followed by methanol and 1-propanol and 1-butanol, respectively. The amount of coke deposition and the deactivation of active site occurs at faster rate for alcohols with a higher carbon number.



Liquid hydrocarbon yield generated vs. TOS for methanol, ethanol, 1-propanol, and 1-butanol over HZSM-5 at 360°C, 300 psig.

PN11014/2336

For 1-propanol and 1-butanol, the activity for liquid hydrocarbon was longer; however, product distribution was mainly composed of higher olefinic compounds in later stage of the catalyst lifetime. This can be explained that the oligomerization of higher olefins requires comparatively weaker acid sites or low activation energy. The activation energy required for the dimerization of ethylene, parent olefinic compound of ethanol, is ~120 kJ/mol compared to the activation energy requires for the dimerization of propylene, parent olefinic compound of 1-propanol, ~70 kJ/mol. Even after the stronger active sites responsible for dehydrocyclization deactivate, oligomerization continues to form liquid range hydrocarbons for the alcohols above C_2.

Acid site strength plays a major role in product class distribution and catalyst life in the transformation of alcohols to hydrocarbon. By fine tuning the HZSM-5 catalyst acid strength and controlling or maintaining the catalyst deactivation using fluidized bed reactor jet fuel range hydrocarbon, ~40 wt% of cyclic content, along with ~60 wt% of olefinic in C_8 to C_{16} range from the alcohols greater than C_2, can be generated. A mild hydrogenation step is required to convert the olefinic to paraffinic compounds.

Design, Synthesis, and Activity Measurements of Monodispersed Multifunctional Catalysts for CO_2 Reduction

Roger J. Rousseau

We are developing technologies for efficient CO_2 conversion to various forms of energy carriers as a prerequisite to reducing the atmospheric carbon level.

With increasing energy needs and potential long-term environmental effects from CO_2 emissions, carbon-neutral fuels are essential to the environmental sustainability of the energy supply. These cleaner fuels will also contribute to a U.S. energy independence from foreign sources. While sequestration may prove a viable short-term solution, the ultimate utilization of CO_2 as a renewable resource in a carbon-neutral economy is more desirable. As fuel consumption forms the major component of carbon utilization, the reduction of CO_2 to hydrocarbon fuels requires a substantial energy input, with most transformations being endothermic and posing a formidable challenges for designing an efficient catalytic process.

This project aims to move soil biogeochemistry and ecology toward a more informed understanding of the relationships between carbon chemistry, microbial community structure and function, and ecosystem perturbations. The demonstration that our analytical tools can be integrated to identify such signatures in a complex system like soil has the potential to transform classical ecological research. We synthesized two different classes of CO_2 conversion catalysts and tested their activity toward CO_2 reduction in a variety of reaction environments. Our overall hypothesis was that catalysts with a dual hydrogenation and acid functionality are required for effective CO_2 activation and conversion. Our goal was to build catalysts ideal for CO_2 reduction and test their kinetics. From these results, we will systematically determine the type of material that will make a good CO_2 reduction catalyst for different reaction media.

The focus of our work in FY 2013 was to synthesize, characterize and measure the activity of monodispersed model alumina-supported heterogeneous catalysts, which ranged in size from single atom to nm clusters. Catalysts were tested in the reduction of CO_2 with H_2 using a temperature programmed reaction method in a flow through catalytic reactor. The aim of these measurements was to find a correlation between catalyst activity and CO/CH_4 selectivity, active metal phase dispersion, and the nature of the support material. We found that highly dispersed metals (atomically dispersed and small metal aggregates) on the oxide support exhibited high CO selectivity, while metal clusters tended to produce CH_4 with high selectivity. Previous experiments on a C-supported catalyst revealed that an oxide presence is critical to CO_2 activation during reduction. These experiments underscored the need of bifunctionality in CO_2 reduction catalysis.

In addition to the previous tasks, we systematically tuned the nuclearity of metal species in the PolyOligoSileSquioxane (POSS) framework. We identified and developed synthetic protocols to stabilize mononuclear and dinuclear species as a fragment of the POSS framework. We also synthesized metallo POSS frameworks containing Ru and Pd Ni species. Structural characterization of these compounds using solution NMR and mass spectroscopy confirmed the existence of mononuclear and dinuclear compounds formed during the synthesis. During this process, we tailored the finer chemical structure to tune the electron density around the metal. The synthesized catalysts are tested extensively in three different media phases of CO_2 hydrogenation: gas, ionic liquid, and liquid. We are investigating structure property relationships to understand the slower kinetics.

For catalysis in switchable ionic liquids, we probed the relationship between structure activity and environment to determine how our hypothesis associated with bifunctionality need be modified to tailor active catalysts in a strongly ionic media. As a target synthesis, we chose the ambitious goal of hydrogenation of alkylcarbonate (formed when CO_2 desolves in the ionic liquid) allows for base/alcohol-catalyzed conversion of CO_2 to methylformate in one pot. The capture and activation of CO_2 into alkylcarbonate allows for unique reactivity during hydrogenation to become formate, which can then be esterified to form the corresponding formate ester. All three reactions are carried out in one pot and are catalyzed by base and alcohol, with the only byproduct being water. In general, we found that catalysts efficient for gas phase or low dielectric hydrogenations of CO_2 show lower activity than catalysts that prefer a more heterolytic hydrogenation pathway, which may be more active in this media. Collectively, our work during this year has yielded two publications in ACS Catalysts and presentations at three international conferences.

Our work in FY 2014 will focus on kinetic measurements of POSS-based high surface area catalysts for CO_2 reduction; initial screening, characterization, and kinetic testing of catalysts in switchable ionic liquids, and the publication of our results on synthetic protocols.

PN13043/2524

Development of New Soft Ionization Mass Spectrometry Approaches for Spatial Imaging of Complex Chemical and Biological Systems

Julia Laskin

Novel experimental approaches provide unique insights about biochemical processes of interest for environmental cleanup, bioremediation, carbon sequestration, national security, and health sciences. We are developing complementary cutting-edge mass spectrometry (MS) imaging techniques that will enable mapping of chemical compounds produced by biological systems with unprecedented spatial and mass resolution.

MS imaging is a powerful technique for obtaining a molecular-level understanding of chemical and biological systems. It offers a number of unique advantages for characterization of complex systems, including high sensitivity, speed, and unprecedented chemical specificity. The limitations of current state-of-the-art MS imaging techniques include sample pre-treatment prior to analysis, unwanted fragmentation of analyte ions, and limited spatial resolution as compared to optical and electron microscopy methods. The objective is to bring soft ionization, sensitivity, and unsurpassed chemical specificity of MS to the nanoscale and apply these novel tools for the characterization of microbial biofilms. We are using *Shewanella oneidensis* and cyanobacteria biofilms as model systems for obtaining molecular level understanding of metal reduction and CO_2 sequestration by microbial communities.

Shewanella has been extensively studied because of its potential use for cleanup of uranium and toxic metals. However, the mechanism of the extracellular electron transport in *Shewanella* biofilms responsible for metal reduction remains largely unknown. *Synechococcus* sp. PCC7002 is a photosynthetic bacterium that in the presence of light converts CO_2 into sugars and other organic molecules. During the course of this project, we developed approaches for comprehensive MS characterization of extracellular material in biofilms. These tools have been used for detailed characterization of chemical gradients generated at interfaces between biofilms and mineral surfaces or interfaces between different microbial communities grown on agar plates.

MS imaging experiments of chemical gradients generated at interfaces revealed the identity of several extracellular redox molecules that facilitate electron transfer between cells and mineral surfaces essential for metal reduction by the biofilm. An integrated approach that combines complementary MS-based imaging approaches with optical and electron microscopy holds great promise to answer conclusively how microorganisms cycle metals such as Fe and Mn, catalyze acidification of metal-rich acid mine waste streams, and participate in the transformation of toxic metals such as U, Cr, Tc, and Pu.

To achieve these goals, we developed two novel platforms for imaging biological samples. The first platform utilizes ambient nanospray desorption electrospray ionization (nanoDESI) MS. This method provides unique advantages for highly sensitive spatially resolved chemical characterization of biological samples in their native state. We demonstrated highly sensitive ambient nanoDESI imaging with spatial resolution > 10 µm, 10× better than that obtained using more traditional ambient pressure surface ionization techniques. The second system combines the C_{60} secondary ion mass spectrometry (SIMS) system with a high resolution Fourier transform ion cyclotron resonance (FT-ICR) MS imaging apparatus. The one-of-a-kind C_{60} SIMS FT-ICR MS apparatus is characterized by high mass accuracy (< 1 ppm), high mass resolving power ($m/\Delta m_{50\%}$ > 200,000), and tandem MS capabilities. The ability to resolve isobaric peaks in SIMS of biological samples is key to characterizing chemical gradients present in such complex systems. Recent instrumental upgrades drastically improved platform performance; specifically, the simplification of ion transfer optics significantly increased sensitivity and mass range-enabling detection of intact lipid species with good sensitivity.

In FY 2013, we demonstrated the unique utility of nanoDESI for direct sampling of living bacterial colonies and for imaging chemical gradients generated by microbes on agar plates. In these experiments, bacterial biofilms grown on agar are analyzed directly from the growing medium without any sample preparation. Spatial profiling of the chemical signals generated by cyanobacteria on agar plates enabled first characterization of the molecules secreted by the bacterial colony demonstrating how the chemical gradients change as a function of the age of the colony. This presents a unique capability for understanding metabolic exchange between microbial communities and between biofilms and mineral surfaces. We also developed novel experimental approaches for accurate quantification and for multimodal imaging of the observed metabolites. The combination of different imaging modalities enables detection of a significantly broader range of metabolites and lipids, thereby providing chemical characterization of microbial communities with an unprecedented level of detail.

PN11028/2350

Development of Preparative Mass Spectrometry for the Creation of Novel Catalyst Materials

Grant E. Johnson

This project will use soft landing of mass-selected ions onto surfaces coupled with* in situ *characterization techniques to prepare well-defined catalytic materials and establish relationships between structure and function that will facilitate the directed design of improved future catalysts.

The chemical and physical properties of nanoparticles, clusters, and complex molecules supported on surfaces are strongly influenced by size, structure, charge state, and composition. The most commonly employed solution-phase method for preparing metal nanoparticles and clusters involves chemical reduction of cationic metal precursors in solutions containing organic capping ligands which arrest the particle growth at a certain size. Although this approach has been shown to produce nanoparticles with controlled size distributions and shapes, it is unable to access the sub-nanometer size regime with atomic-level precision. Therefore, complex and costly separation methods such as gel electrophoresis and fractional distillation are often required to prepare clusters containing exactly a certain number of metal atoms. Soft landing of mass-selected ions is a powerful alternative approach for the preparation and modification of surfaces and enables unprecedented control over the composition and surface coverage of deposited materials. In addition, non-thermal physical synthesis techniques such as laser vaporization combined with gas aggregation may be used to prepare novel clusters that cannot be synthesized using traditional thermal approaches in solution.

The size and ionic charge states of metal clusters are known to exert a strong influence on their optical, electronic, and reactive properties. Chemical reduction synthesis in solution was combined with high-resolution mass spectrometry to examine how size, charge state, and ligand-substitution influences the fragmentation pathways of diphosphine-ligated gold clusters substituted with cyclohexene units. It is shown that in addition to the length of the alkyl-chain spacer between phosphine groups in diphosphine ligands, the chemical substitution of the phosphine centers has a large influence on the size and charge state of gold clusters synthesized in solution. Moreover, it is observed that a small change in substitution at the phosphine centers from C_6H_5 to C_6H_{11} may result in different gas-phase fragmentation pathways for otherwise identical gold clusters. These results indicate that the common practice of substituting simple hydrogen atoms in place of more computationally demanding functional groups may not accurately describe the properties of diphosphine ligated gold clusters. These outcomes are the subject of an invited publication in *ChemPlusChem*.

In addition to metal clusters and nanoparticles, metal oxides find widespread use in commercial catalysis, for instance, in the production of sulfuric acid. The size, charge, and oxidation state of metal oxide clusters have been shown to exert a strong influence on their catalytic reactivity. Solution-phase synthesis was combined with electrospray ionization, in-source fragmentation, and tandem mass spectrometry to generate a wide range of anionic polyoxovanadate clusters for subsequent gas-phase experiments and soft landing onto surfaces. Pure vanadium oxide as well as chlorine-, proton-, and ammonium-ligand containing species were observed. Tandem mass spectrometry experiments revealed surprisingly simple fragmentation pathways for ions generated by in-source fragmentation, while ions originating directly from solution were observed to have more complex fragmentation pathways. These results are the subject of a publication in the *Journal of the American Society for Mass Spectrometry*.

To access a wider range of vanadium oxide clusters for gas-phase studies as well as soft landing experiments, polyoxovanadates were synthesized in solution using two different size ammonium cations and heteroanions. Employing high-resolution mass spectrometry, it is shown that the smaller heteroanion (Cl^-) results in the preferential formation of larger multiply charged polyoxovanadates containing 14 vanadium atoms. By contrast, the larger acetate heteroanion (CH_3COO^-) promotes the formation of smaller singly and doubly charged polyoxovanadates containing six vanadium atoms. The ammonium cation is observed to have a much weaker influence on the size of the clusters with the larger ammonium cation favoring slightly larger polyoxovanadates. These results, which provide insight into the interactions between metal oxides, heteroanions, and ammonium cations are the subject of an invited publication in the *International Journal of Mass Spectrometry*.

Future areas of inquiry will include the following two objectives: the preparation of atomically monodisperse nanoparticles and clusters on surfaces and electrochemical examination of their size- and shape-dependent reactivity toward O_2 reduction; and a structural and thermodynamic characterization of subnanometer ligated gold clusters in the gas phase.

PN10026/2268

Exploitation of Kinetic Processes in Gas Separations

Paul H. Humble

We are identifying promising materials and developing adsorption-based gas separation techniques to collect and purify trace atmospheric gases with low energy requirements, few or no consumables, and small equipment sizes.

The ability to collect and measure trace gases in the atmosphere is important for treaty monitoring and other national security applications. Current techniques use liquid nitrogen-cooled adsorption or large ambient temperature traps. These processes are resource intensive, often requiring large amounts of energy to heat and desorb large ambient temperature adsorption traps. Rapid-pressure-swing-adsorption (R-PSA) separation processes that exploit kinetics of the adsorbent-adsorbate interaction have potential to reduce energy requirements and equipment size and remove or reduce the need for consumables such as liquid nitrogen. This project combines novel finite-element simulations with experimental measurements of adsorption properties to advance concepts for next generation whole air collection and purification systems.

Data were collected using two instruments. The Hyden Dynamic Adsorption Analyzer was used to collect the bulk of the equilibrium and kinetic adsorption data. An Adsorption Breakthrough Instrument (ABI) was assembled and used to collect data for promising adsorbent-gas pairs in a realistic adsorption bed configuration. Process modeling was accomplished using Aspen Adsorption and COMSOL Multiphysics software. The adsorption data collected using the Hiden and ABI were used to fit adsorption isotherm models and other parameter required for the simulations. The process simulations were used to develop R-PSA cycles for separations of interest.

Hyden Dynamic Adsorption Analyzer used to collect adsorption kinetics and equilibrium.

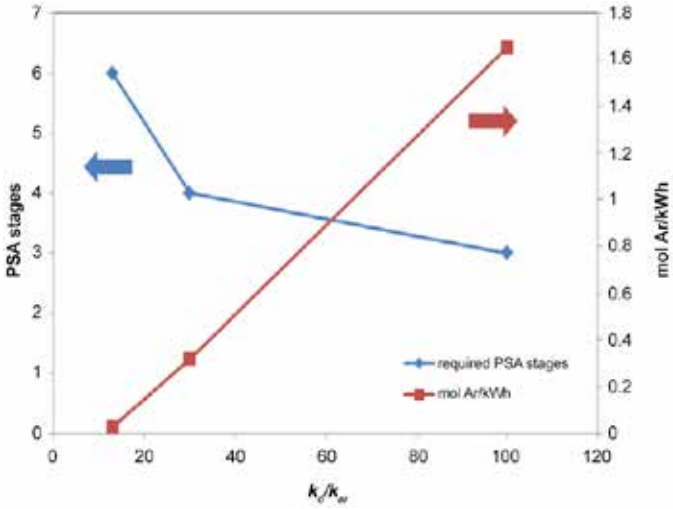
During FY 2012, the kinetic and equilibrium adsorption properties of a number of adsorbents were measured, including coconut shell activated charcoal (as a baseline), Carboxen 1000 and 569, MS 5A, MS 13X, and Ca Chabazite. Gasses tested included N_2, O_2, CO_2, Ar, Kr, Xe, and CH_4. Adsorption process models (R-PSA simulations) that included adsorption kinetics were also developed and tested. Out of the materials investigated during the first year of the project, Ca Chabazite showed the largest kinetic effects and formed the basis for the early modeling effort, which along with the experimental results led us to focus on Ar-O_2 and Ar-N_2 separations. Materials that showed kinetic selectivity for Ar-N_2 and Ar-O_2 showed the same for the larger and higher molecular weight species. However, high molecular weight species showed higher equilibrium uptake for the adsorbent-gas pairs we examined. In many cases, the desired selectivity (slow adsorption of the species of interest) was countered by high equilibrium uptake. The Ar-O_2 separation was identified as a good base case; if a material showed good selectivity for this separation it could also be used for larger gases such as Xe. In these cases, the process would produce a product stream that was mostly Ar but enriched with the heavier specie. Process simulations using Ca Chabazite showed that the degree of Ar-O_2 separation for a single stage was modest, and adsorbents with higher kinetic selectivity were desirable.

During FY 2013, we continued the search for adsorbent materials with enhanced kinetic selectivity. Samples of Ionex Ag 400, Ionex AG 900, MS 3A, and MOF-5 were tested. Samples of LiX and carbon molecular sieves (CMS-200 and CMS-250) were also procured and tested. At the end of this year, all the adsorbent-gas pairs identified at the start of the project were tested as well as several that were identified during literature reviews and process modeling activities.

Our experimental and modeling results show the carbon molecular sieve materials CMS-200 and CMS-250 showed the largest kinetic selectivity for the adsorbent-gas pairs we tested. A selectivity of 15 for Ar-O_2 on CMS-200 was measured. This allows for a single stage to produce a 40% Ar

PN12029/2430

product. Multiple stages can be used to boost the product purity. As expected, there is a trade-off between product recovery and product purity, and higher purities are possible at the cost of lower recovery. Although we measured a kinetic selectivity of 15, higher selectivity (up to 100) for CMS materials have been reported, which has a beneficial impact on purity and recovery and would make a one- or two-stage R-PSA process feasible. Our modeling work shows that to achieve 99.99% product purity requires more than three stages.



Trade-off between PSA stages and mols Ar per kWH as a function of kinetic selectivity for producing 99.99% Ar.

However, a gas separation process used for treaty monitoring or on-site inspection would likely also include a chromatography step or a small temperature swing step. In these cases, a single R-PSA process could produce the required purity for the first stage separation processes. A promising configuration would include an R-PSA process using MS-5A to remove nitrogen followed by an R-PSA process to produce a crude Ar product. This product would then be refined using gas chromatography and temperature swing adsorption.

In this second and final year of the project, the focus has been on developing and maturing our R-PSA process simulation capabilities. A large number of studies have been done examining the impact of kinetic selectivity, cycle timing, and cycle steps on separation performance. Promising separation cycles have been identified for Ar collection from air, and this will form a basis and starting point for future work on collection systems and Ar-37 detection for treaty monitoring. Our ability to measure and make use of adsorption data has been increased, and we have developed a modeling capability that will benefit future projects.

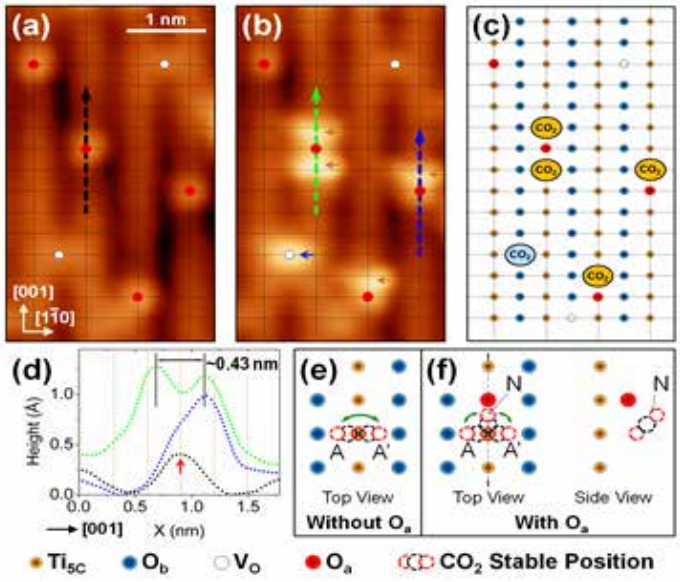
Fundamentals of Carbonate Formation: Interactions of Carbon Dioxide with Supported Metal Oxide Clusters

Xiao Lin

The goal of this project is to obtain a detailed, fundamental understanding of how to control catalytic conversion of CO_2 (and relevant molecules) on various oxide catalysts.

As industries rely on fossil fuels for energy and the global warming problem mounts, there is an immediate need for improving the fundamental understanding of CO_2 activation, which is instrumental for carbon sequestration and CO_2 conversion to fuels. Using TiO_2(110) and other well-characterized oxides surfaces as models provide an opportunity to understand the fundamental aspects of elemental processes such as CO_2 adsorption, diffusion, and reaction.

From FY 2010 to FY 2012, we studied the adsorption and dynamics of propane-1,3-diol and CO_2 molecules on partially reduced rutile TiO_2(110)-1×1 surfaces. In FY 2013, we studied the adsorption of CO_2 on TiO_2(110) with pre-dosed oxygen by means of scanning tunneling microscopy.



Identical area STM images of TiO_2(110) (a) before and (b) after adsorption of 0.06 monolayer of CO_2. The grids indicate positions of the Ti_{5c} ions. V_0 and O_a sites are indicated by white and red dots, respectively; (b) CO_2 molecules on V_0 and next to O_a are indicated by blue and orange arrows, respectively; (c) Schematic view of (b); (d) Line profiles of O_a (black), CO_2-O_a (blue) and CO_2-O_a-CO_2 (green) complexes depicted by the dashed-lines in (a) and (b). Dashed vertical orange lines indicate the positions of the Ti_{5c} ions; (e) Energetically preferred binding configurations for CO_2 on Ti_{5c} sites on TiO_2(110); (f) Binding configurations for CO_2 on Ti_{5c} sites next to O_a proposed based on the STM images.

Pre-dosing oxygen at low temperature followed by annealing to 220 K gives rise to oxygen adatoms (O_a) binding on five coordinated Ti sites (Ti_{5c}) on the TiO_2 surface. At 50 K, adsorbed CO_2 molecules are found preferentially on Ti_{5c} next to O_a rather than on oxygen vacancies (V_0), the most stable adsorption sites for CO_2 on TiO_2(110) without the O_a. Only a smaller fraction of V_0 sites is populated by CO_2 at 50 K. Temperature dependent studies show that subsequent annealing to 100–150 K leads to the CO_2 redistribution with a preferential occupation of V_0 over Ti_{5c} next to O_a, indicating that at 50 K, the adsorption of CO_2 onto V_0 is kinetically hindered. The equilibrated fractional populations obtained after thermal annealing suggest that the CO_2 binding energy on V_0 is only 9 ~ 25 meV higher than on Ti_{5c} next to O_a. The diffusion barrier of CO_2 away from the Ti_{5c}-O_a sites is estimated to be ~ 0.17 eV based on the CO_2 hopping rate at 50 K.

The detailed structure of CO_2 features on Ti_{5c} sites next to O_a was observed. The ideally symmetric appearance of crescent-like CO_2 is affected by the presence of nearby V_0 and other CO_2 molecules. Assuming that the brightness of each configuration in the time average observed reflects the statistical population of binding configurations, we conclude that CO_2 tilts preferentially away from the O_b row containing the V_0 defect but tilts toward the O_b row that has another CO_2 molecule directly across on the neighboring Ti_{5c} row. We reported these results in two articles that appeared in and were highlighted on the cover of *Journal of Physical Chemistry C* and *Physical Chemistry Chemical Physics*.

We also studied the preparation of other oxides, specifically RuO_2(110) grown on Ru(0001) crystal surfaces by in situ oxidation. The surface structure of RuO_2(110) is the same as rutile TiO_2(110) and has been shown to activate CO_2 readily to chemisorbed CO_2 and form negatively charged dimerized species CO_2-CO_2. To date, no high resolution site-specific research had been done on the CO_2/RuO_2 system. In future studies, we will investigate the structure and dynamics of CO_2 on RuO_2 surfaces and compare the results with that of CO_2/TiO_2. This study will provide valuable information about structure, activity relationships on such model systems.

PN10038/2280

In Situ Molecular-Scale Investigations of Reactions between Supercritical CO_2 and Minerals Relevant to Geological Carbon Storage

John S. Loring

This project will result in unique new data for molecular mineral transformation processes involved in carbon capture and sequestration, including processes from mineral dissolution/nucleation/precipitation and hydration/dehydration to* in situ *sorption/desorption studies in supercritical carbon dioxide ($scCO_2$).

One approach to slow the flux of greenhouse gases into the atmosphere is geologic carbon sequestration (GCS), whereby CO_2 is stored deep underground, such as in saline basalt formations or exhausted oil wells. At targeted injection depths, carbon dioxide exists as a supercritical fluid ($scCO_2$). With repeated injections, the $scCO_2$ will displace formation water, and the pore space adjacent to overlying caprocks will eventually be dominated by dry to water-saturated $scCO_2$. Water-bearing $scCO_2$ is highly reactive and capable of carbonating and hydrating certain minerals, whereas anhydrous $scCO_2$ can dehydrate water-containing minerals. Because these geochemical processes affect solid volume and thus porosity and permeability, they have the potential to affect the integrity of the caprock seal. To predict the viability and risks of GCS, it is vital that we understand the reactions at reservoir conditions between wet $scCO_2$ and host and caprock minerals.

The chief goals of this project are to integrate a powerful suite of *in situ* experimental capabilities recently developed at PNNL to investigate reactions between minerals and $scCO_2$ containing variable amounts of water, and to focus these capabilities on a range of geologically relevant host and caprock materials. These in situ probes include infrared (IR) spectroscopy, magic angle sample spinning nuclear magnetic resonance (NMR) spectroscopy, atomic force microscopy, and x-ray diffraction (XRD). This project's success will lead to an increased understanding of geologically relevant mineral transformation processes in low water $scCO_2$, which will help to inform and guide development of reactive transport simulations for subsurface CO_2 reservoirs.

The work performed during FY 2012 set the stage for progress in FY 2013. Our focus during FY 2012 included analyzing results and writing a manuscript on the reactions of olivine and serpentine with wet $scCO_2$, completing experiments on hydration-dehydration of the expandable smectite clay, montmorillonite, and initiating studies of a natural rock sample, Eau Clarie shale. Additionally, two new *in situ* high-pressure capabilities were brought online: a quartz crystal microbalance (QCM) and an IR titration system housed in a temperature controlled laboratory.

There were four main focuses in FY 2013. The first was to write, submit, and publish several articles. In summary, FY 2013 ended with six articles in print and three submitted manuscripts. Further, results were disseminated through a total of seven conference presentations.

The second focus was to analyze and interpret IR-titration, NMR, QCM, and XRD data for the reaction of smectite clays with variably wet $scCO_2$ at 50°C and 90 bar. Results on Na-, Ca-, and Mg-saturated Wyoming bentonite (a model smectite) reveal correlations between interlayer spacing, water content, adsorbed CO_2 concentrations, and interlayer cation. The data demonstrate that the interlayer spacing is a complex function of partitioning of water and CO_2 between the clay and the supercritical fluid. Further, molecular scale information from the IR results provide structural and environmental information about adsorbed CO_2 molecules that are helping to constrain ongoing computational modeling efforts by molecular dynamics simulations.

The third focus during FY 2013 was the study of the reaction of an Eau Clarie shale caprock with wet $scCO_2$. The XRD of this natural sample revealed that the major components were smectite, illite and kaolinite clays. An IR spectroscopic titration with water at 50°C and 90 bar $scCO_2$ demonstrated water uptake, presumably by the smectite fraction. *In situ* high-pressure XRD experiments revealed that hydration-dehydration processes lead to volume changes. These results suggest that the dehydration of the expandable clay fraction in a shale caprock by dry $scCO_2$ could decrease solid volume and potentially increase caprock permeability and porosity.

Finally, the fourth main focus was an investigation of adsorbed water film thickness on forsterite carbonation chemistry using *in situ* IR spectdroscopic titrations and *ex situ* scanning electron microscopy, transmission electron microscopy, thermogravemetric analysis, XRD, and NMR. From these cumulative measurements, we conclude that the reactivity of forsterite with wet-$scCO_2$ can be divided into three adsorbed water concentration threshold regimes: 1) up to ~ 2 monolayers of adsorbed water, only a highly structured and low mobility chemisorbed water and bicarbonate film is present; 2) between ~ 2 and ~ 5 monolayers, limited carbonation occurs and then nearly stops; and 3) above ~ 5 monolayers, continuous carbonation occurs, and magnesite is the dominant product.

PN11040/2362

Increased Sensitivity and Improved Quantification of Th and U in Particles by SIMS

Albert J. Fahey

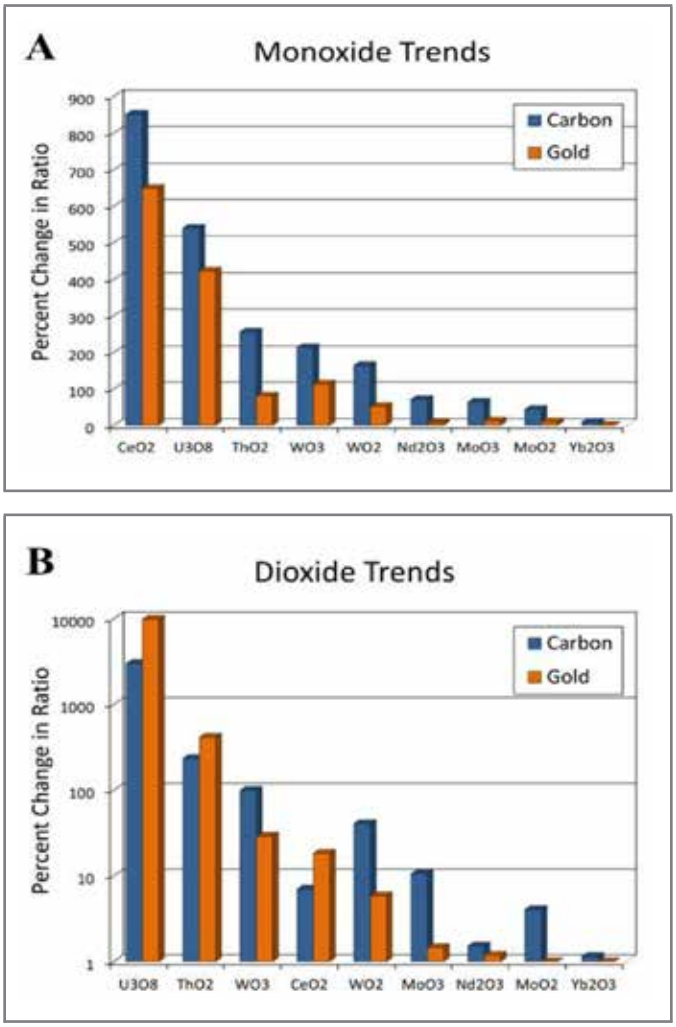
This research studies how oxide and elemental ions form on various substrates and under different conditions through elemental oxide measurements and dioxide signals from single particles under prescribed conditions. This may yield an over-determined system amenable to least squares analysis, providing redundancy and consistency checks on the validity of the resultant data.

A specific high-value method involving the measurement of thorium and uranium was developed during the late 1990s. Advancing the method from its current state, ionization efficiency of elemental and oxide ions will be explored for various substrate materials and under conditions of oxygen flooding. The primary focus of this project is to increase the ion signal intensity of the oxide and dioxide ions of thorium and uranium in secondary ion mass spectrometry (SIMS). Once the magnitude of the effect is measured and confirmed on the PNNL IMS-4f, samples will be investigated on an IMS-1280. Data and methods pertaining to the application of method and fundamental mechanisms of sputtering ionization will be obtained and analyzed.

We are determining the optimal protocol for uranium, focusing on ionization the efficiency, identification, and reduction of interferences, increased signal stability, and lower statistical variability. Additionally, we are focusing on understanding the basic mechanism of oxide and ion formation. Other compounds will be characterized on the various substrates used. Finally, we will attempt to understand and quantify these results, specifically relating them to basic chemical properties of the element and oxide ions formed.

During FY 2011, we obtained sample and substrate material, brought the laboratory up to operational status, and developed a cursory method to produce SIMS substrates for silicon and gold materials. In FY 2012, we verified that U isotopes measurements can be made as the oxide species on various substrates (Si and Au) without isobaric interferences.

For FY 2013, experiments were performed by monitoring the secondary ion yield as a function of ion dose (number of ions/s/cm^2). A typical SIMS response is a convolution of



Plots of the percent change in oxide to element ratio for (A) monoxide (linear scale) and (B) dioxide (log scale) from data obtained during oxygen flooding on and off. Substrates indicated as carbon (blue) and gold (orange).

two phenomena: an exponential degradation of the ion signal as a function of ion dose known as the erosion dynamics model; and an asymptotic enhancement of the ion signal as a function of ion dose due to increase ionization efficiency. Background subtracted data show secondary ion response for the elemental, monoxide, and dioxide signals for analysis of a dispersion of MoO_2 particles on a gold substrate. In this case, the element signal was unaffected, whereas the monoxide signal had a strong increase and the dioxide signal responded had a weak increase in signal. Once oxygen flooding was terminated, the signals returned to baseline intensities. This experiment was run for various metal oxide dispersions on both gold and carbon substrates.

PN11041/2363

The comprehensive data are shown in the figure, which plots a comparison of the element to monoxide and dioxide ion signal ratios for each sample with and without oxygen flooding. The data indicate several general trends: 1) the relative signal intensity for oxides almost always increases during oxygen flooding; 2) the oxide signal almost always responds greater to oxygen flooding for particles dispersed on carbon than for those on gold substrates (UO_2^+, ThO_2^+, and CeO_2^+ are exceptions); 3) metal oxide particles in higher oxidation states respond stronger to oxygen flooding than those in lower oxidation states; and 4) the relative sensitivity of the metal oxide signals is inversely proportional to the electron affinity of the corresponding element.

Although the relative ionization yields show enhancements for oxide molecules over atomic species for several elements the question of absolute ionization efficiency for the atomic and molecular species must be investigated for the situation where particles are mounted on different substrates. This is a key piece of information when deciding the experimental protocols, especially when the highest possible sensitivity is required.

Particles of U_3O_8 of U reference materials were placed on gold and carbon substrates and measured with constant ion current density. Particle locations were marked by fiducials for re-location between the SEM and SIMS instruments. Images of each particle were acquired, and the apparent presented areas of the backscattered electron images were measured with the GNU image manipulation program (GIMP). The particles were re-located in the SIMS instrument and consumed to exhaustion. The total counts of U, UO, and UO_2 were computed from the acquired data. A plot of the total counts vs. the 3/2 power of the apparent presented area was fitted to a line (force through the origin) to yield the absolute counts per unit volume for each ion signal on gold or carbon.

Overall, we were successful in increasing the relative sensitivity for elemental oxide ion signals for SIMS analysis with the greatest effect observed for UO_2^+ and ThO_2^+. The data show that while the magnitude of the increase can be characterized within a few generalized trends, it is still strongly dependent on the combined analyte and substrate surface chemistry. Nevertheless, Th and U are good candidates for increasing their ionization efficiencies in this manner. By combining the FY 2013 results with the FY 2012 isotopic results, we believe this is a viable method for improving measurement statistics through decreased signal variability and potentially providing a path toward improving quantitative relative abundance determinations by SIMS.

Although the relative ion signals for oxide ions are higher than the atomic ion signals in most cases of SIMS from metal oxides on substrates (either gold or carbon), the absolute ion yields are higher for atomic ions, at least of uranium, than for the oxides. This leads to the conclusion that atomic ions of uranium from oxides on carbon afford the greatest sensitivity possible given the conjecture that sputtered carbon getters away much of the oxygen that would normally bond to the uranium.

Light Source Photocathode Performance and Development

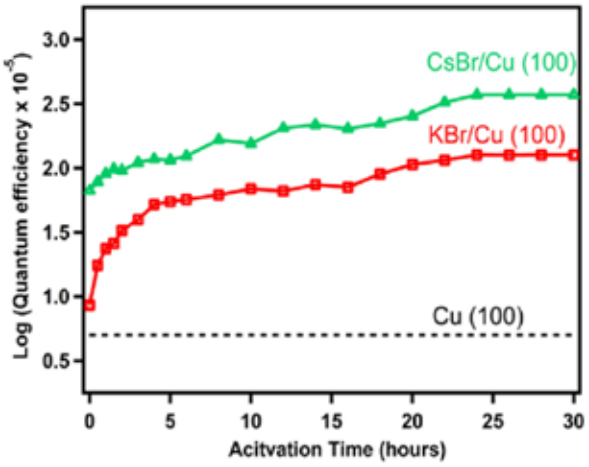
Wayne P. Hess

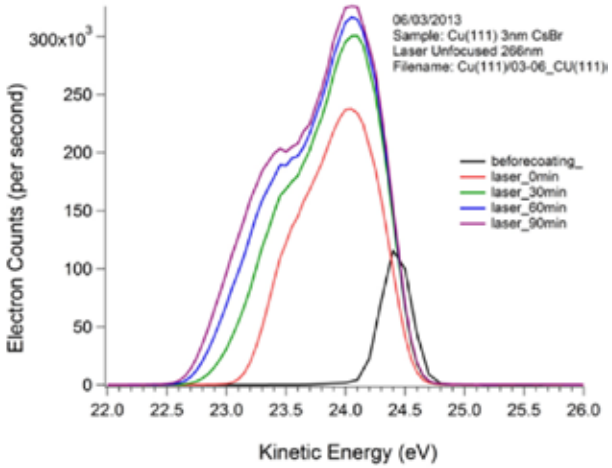
This project will develop new capabilities for developing and testing existing photocathode materials and designs that measuring photophysical properties, including work function, photon energy dependent electron yield (bunch intensity), and emittance (electron angular distribution).

Whether based on free electron laser (FEL) or storage ring designs, output characteristics and cost of new ultraviolet or X-ray sources are dependent on photocathode brightness and emittance characteristics. Novel photocathode designs could potentially reduce light source construction costs enormously—by a factor of two or more—by significantly simplifying downstream accelerator or FEL design. The international light source community recognizes the need for a more scientific approach to new photocathode development and is in initial stages of addressing this issue.

Our initial goals were to increase the photocathode longevity by understanding and modeling photocathode failure mechanisms. We conducted a post-mortem analysis of used (expired) photocathodes from the Thomas Jefferson National Accelerator Facility Laboratory (JLab) and determined that the photocathode emission yield degrades under operating photo-injector conditions consistent with a model in which the surface Cs layer is removed from the surface by ion back bombardment sputtering. Next, we constructed a coupled photocathode preparation chamber and ultraviolet and X-ray photoelectron spectrometer to create and test photocathode designs.

For FY 2012, we published a manuscript in *Physical Review Special Topics – Accelerators and Beams* and completed our post-mortem analysis of expired JLab photocathodes. Helium ion microscopy results showed that the photocathode activation procedure induced significant surface roughening that likely increases emittance. The observed roughening led us to propose a novel method for activation of GaAs photocathodes of the type used at the JLab. A similar project was undertaken at Cornell University such that our efforts were redirected away from semiconductor photocathode development and toward metal photocathode development with application to dynamic transmission electron microscopy (DTEM). Toward that end, we contracted for the procurement of a 100 kV electron gun as a component in a test stand for DTEM photocathode development to be constructed for a follow-on project during FY 2014.





Activation curves (top panel) show quantum efficiency of Cu(100) (dashed black line), CsBr/Cu(100) (triangles, green line) and KBr/Cu(100) (squares, red line). The measured quantum efficiency of Cu(100) is $5x10^{-5}$ at 4.7 eV. Electron energy distribution determined by ultraviolet photoemission spectroscopy showing two contributions to quantum efficiency enhancement.

In FY 2013, we tested the utility of thin films of KBr, in comparison with CsBr, to enhance quantum yield from Cu. We were able to identify two initial states for enhanced electron emission for both materials. The first resulted from the work function lowering due to growth of the CsBr or KBr layer on Cu. The second state was due to formation of F centers within the alkali halide thin films, which are the sources of photoelectrons that are rapidly re-filled following photoemission by electron transfer from the Cu substrate. Upon an initial coating with CsBr and KBr, the increase in quantum efficiency is displayed in the figure. CsBr shows a greater capacity for enhanced QE both before as well as after activation. The results of this study were published jointly with Stanford University in *Applied Physics Letters*.

PN10050/2292

Localized Surface Plasmon Resonance Spectroscopy, Microscopy, and Sensing

Patrick Z. El-Khoury

This project combines spectroscopic (surface- and tip-enhanced Raman) and theoretical (static and dynamic quantum chemical simulations) tools to probe the interplay between a single molecule and its local environment. This knowledge will enable the construction of novel plasmonic devices operating at the ultimate limits of space and time.

Efforts to reach the ultimate detection limit – the sensitivity required to probe 1.66 yoctomoles (1/NA) of a substance – has rendered the art and science of designing plasmonic junctions increasingly popular over the past few years. This is a consequence of the extreme localization of electromagnetic fields between plasmonic nanostructures, whereby the detection of the feeble optical response nascent from a single scatterer is feasible. The effect is most evident in surface- and tip-enhanced Raman scattering (SERS and TERS, respectively) from molecular reporters coaxed into hot spots. Through this project, a combination of experiment and theory allows us to understand the governing principles in single molecule SERS and TERS, and to apply the acquired knowledge towards designing plasmonic devices with desired functionalities.

Our recent PNNL works revealed that inelastic scattering at plasmonic junctions not only probes molecular polarizabilities but also broadcasts intimate details about both the environments in which the scatterers reside and the interplay between molecule and plasmon. In the *Journal of Physical Chemistry*, we illustrated that it is possible to design a plasmonic junction – in this case between a silver nanoparticle on a silver surface – which exclusively operates at the quantum limit of plasmons. Namely, when the separation between the two metals defining the junction is dictated by a flat adsorbate, the broad response characteristic of charge transfer plasmon-enhanced Raman scattering completely overshadows the familiar dipole plasmon-enhanced molecular line spectra. In a subsequent work, and through TERS trajectories recorded at a nanojunction formed by a gold AFM tip in contact with a silver surface coated with BPDT, we further established that junction plasmons can tunnel through conductive molecules which bridge the nanogap through chemical bonds. In the aforementioned study, the change in scattering activity associated with shorting the junction plasmon served as a method for monitoring the breaking and making of single chemical bonds.

In FY 2013, our combined experimental and theoretical investigation of Raman scattering from 1,3-propanedithiol at a hot spot resulted in a peer-reviewed publication in *Chemical Physics Letters*. In this study, we employed tools of density functional theory to compute the Raman spectra of 1,3-propanedithiol (PDT) isolated in the gas phase, solvated in methanol, tethered either to the face or vertex of a tetrahedral Ag20 cluster, and bridging two Ag20 clusters. The derived molecular polarizability derivative tensors were used to simulate molecular orientation-dependent Raman scattering, achieved by rotating the polarizability tensors relative to vector components of the incident/scattered radiation fields. Our framework was weighed against our AFM/SERS experiments, which probed the optical response at a hotspot formed by an Ag surface coated with PDT and a single 60-nm Ag nanosphere. A remarkable agreement between experiment and theory not only allowed for reliable spectral assignments to be made but also provided a well-founded theoretical basis to account for commonly observed phenomena in the single molecule realm (e.g., spectral fluctuations or blinking). In FY 2014, we will build on these findings to refine and generalize our theoretical treatment of single molecule SERS/TERS and to construct novel plasmonic devices based on the particle on a mirror design.

Our discoveries from FY 2013 prompted ongoing combined experimental and theoretical endeavors, where the scattering from SERS/TERS reporters chemisorbed (e.g., benzenthiol) or physisorbed (e.g., *meso*-tetraphenylporphyrin) on roughened silver substrates are currently under investigation. Our experimental findings prompted the expansion of our quasi-static theoretical treatment of finite metal cluster-bound π-conjugated organic complexes to the time domain. This is done using *ab initio* molecular dynamics simulations, carried out either by propagating the density matrix or the orbital coefficients; performed either using atom centered functions or plane waves. An allocation of computing time (500k SUs) and a supplementary allocation (500k SUs) were awarded from the San Diego Supercomputer Center to support the theoretical effort.

We presented our research and findings at the 2013 Gordon Conference on Photochemistry and at the 2013 PNNL Postdoctoral Research Symposium, which attracted several external collaboration opportunities with that will be pursued in FY 2014.

PN13092/2573

Molecular Structure and Interaction at Aqueous, Non-Aqueous Liquid Interfaces and Catalytic Solid Surfaces

Hongfei Wang

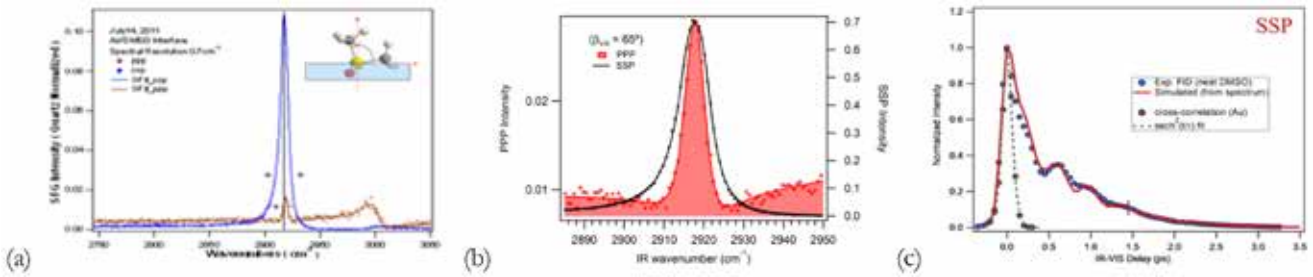
Using cutting-edge instrumentation, this project will lead to new discoveries in surface science and have broad applications on the characterization of structure and interactions at various molecular interfaces that are critical but elusive in many scientific and engineering fields.

The interface selectivity and sub-monolayer sensitivity of the second order nonlinear optical spectroscopy make it a unique tool to study the structure and interaction of molecules at various surfaces and interfaces relevant to energy as well as environmental and biological processes. In the second order optical method, two photons with the same frequency (second harmonic generation) or different frequencies (sum-frequency generation) interact simultaneously with the same set of atoms or molecules to generate a new photon at the sum of the frequencies. The symmetry requirement for these processes forbids the processes from the bulk liquid or amorphous solid with centrosymmetry but allows them to be generated from the surface or interfaces that always have asymmetry. The new photon generated is also background free from the large number of incoming photons, so it allows sensitive detection such as photon counting for submonolayer level vibrational and electronic spectral measurement.

For the past two decades, scientists have been using this tool to understand and characterize the details of the structure and interactions at various molecular surfaces and interfaces. This has been a growing field and is ready for exploring applications related to material and biological surfaces and interfaces. PNNL played leading roles in developing the experimental methodology and theoretical treatment to push forward the advancement in this field from the simplest air/liquid interface to more complex liquid/solid, liquid/liquid, liquid/membrane, and nanoparticle interfaces.

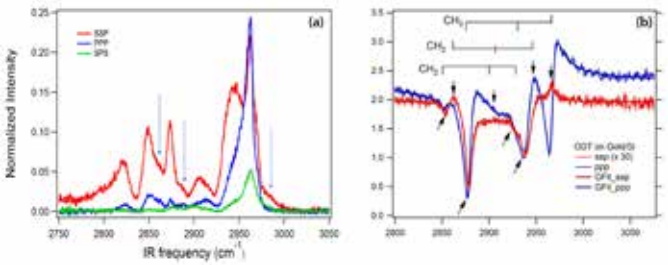
Specifically, this project aims to employ the recent development of ultrafast laser technology to push the surface nonlinear spectroscopy and dynamics unprecedentedly to higher spectral resolution, time resolution, and detection sensitivity to solve problems with complex molecular surfaces and interfaces. It is a unique capability and a game changer that will impact in understanding molecular surface science, especially to complex surfaces that ubiquitously exist in energy, environmental, and biological systems that cannot be accessed or studied in situ and at ambient or extreme conditions using other x-ray, electron, or ion probe tools and instrumentations.

We commenced work in 2010 to support experimental and theoretical studies applicable to various aqueous, non-aqueous liquid interfaces, and catalytic solid surfaces. The key to successful completion rests on the successful design, construction, and testing of the complete laser



The HR-BR-SFG spectrometer enables measurement with unprecedentedly accurate line shapes in different polarization combinations, and the coherent vibrational dynamics can be accurately predicted using the spectral data. (a) and (b) For the air/dimethyl sulfoxide (DMSO) interface SFG spectra, line shapes in different polarization combinations enable resolving the peaks of the two methyl groups of the interfacial DMSO molecule, which appear to be one peak in each spectrum in the ssp and ppp polarization combinations, respectively. The separation between peaks is 2.8 cm-1, compared with full width at half maximum ~8-9 cm-1. In addition, the phases and oscillator strengths of the ssp and ppp spectra can accurately determine the tilt angle of the two methyl groups from the interface normal. (c) The coherent vibrational dynamics as probed with ultrafast SFG free-induction decay (SFG-FID) is quantitatively reproduced with the high-resolution spectral data. This demonstrated that the HR-SFG is not only a tool for spectroscopic and structural studies, but a powerful tool for dynamics studies. Such information has not been expected, and this development is expected to change the related fields significantly.

PN10058/2300



HR-BB-SFG-VS as game-changing tool for complex surface spectral and structure analysis: (a) The HR-BR-SFG spectral of cholesterol at air/water interface. More than 10 overlapping peaks can be well resolved from the congested spectra. The polarization dependence can also allow unique assignment of the peaks; (b) Unique fitting in both ssp and ppp spectra of ODT self assembled monolayer on gold surfaces can identify 8 peaks, and their intensity and phase relationships allow determination of the conformation of the ODT molecule. SFG spectrum on metal surface is usually dominated by ppp polarization, and the ssp spectral line shape is usually unclear. With the high resolution SFG, the ssp spectral line shape is with excellent signal/noise ratio, allowing quantitative polarization and structural analysis for SFG on metal surfaces.

system; the nonlinear optical experimental setup; and the various optical cells to be used for various surface measurements. All design and construction were underway, setup, and tested by the end of the year. In FY 2011, we showed that the high resolution and excellent line shape in high resolution broadband surface sum-frequency generation vibrational spectrometer (HR-BB-SFG-VS) enables resolving spectral and structural details of molecular interface, which was not attainable before. We successfully built the world's first HR-BR-SFG-VS.

Spectral resolution is ~0.6 cm^{-1}, about 10–20 × better than current BR-SFG spectrometers. High resolution SFG-VS spectra allow accurate measurement of the surface SFG spectral line shape, allowing spectral line shape analysis and detailed SFG spectral phase analysis through accurate SFG intensity spectra measurement. The ability for surface SFG-VS to resolve detailed chemical shifts and structure of the molecular interface shall make SFG-VS a unique analytical tool. For non-aqueous liquid and solid/liquid interfaces using high resolution and scanning SFG-VS spectrometer, we observed catalytic reaction product of from acetic acid on CeO_2 nanoparticle surface with the established procedure for using SFG-VS in in situ nanoparticle surface characterization; measured and analyzed the hydrogen bonded water spectra at the aqueous/SiO_2 and aqueous/CaF_2 surfaces; and measured the surface keto-enol structures of acetylacetone, a model ligand in liquid-liquid extraction of heavy elements under different conditions.

In FY 2012 and FY 2013, we continued the research on various molecular interfaces that focused on the following aspects:

Pushing the limit for resolving spectral and structure details of relatively simple liquid and Langmuir monolayer surfaces. We showed that high-resolution SFG-VS can clearly capture spectral interference and phase relationship in closely overlapping vibrational peaks. Examples are demonstrated on the long-chain nonadecanenitrile (C18CN) and the 4-pentyl-40-cyanoterphenyl (5CT) Langmuir monolayers.

Applying what have been learned to study the spectral and structure details of complex and inhomogeneous surfaces and interfaces, such as nanoparticle surfaces, aqueous/oxides interfaces, and catalyst surfaces. For example, we resolved 8 peaks (vs. 3 or 4 in the literature; see figure at left) in the C-H stretching vibration region for the octadecanethiol (ODT)/gold surface, and demonstrated how to perform polarization analysis on SFG spectra of molecules on metal surfaces. We also used cholesterol molecule as the example to show that high-resolution SFG can clearly identify spectral details in complex vibrational spectra. These point to the possibility to use the C-H vibrations as the new vibrational fingerprint region.

Exploring the dynamic information of surface molecules in the line shape of the high-resolution SFG-VS. We demonstrated that the peak intensity (position and width obtained from high-resolution SFG-VS line shape) can quantitatively predict the coherent vibrational dynamics of surface molecules, including homogeneous and inhomogeneous broadening and interactions. Unlike the direct dynamics measurement that is limited to well separated vibrational peaks, the frequency domain high resolution SFG-VS measurement can be used for obtain such information for molecules with closely overlapped peaks. This shall greatly broaden the scope of dynamics analysis of molecular surface interactions.

These studies significantly moved the field of surface nonlinear spectroscopy forward, leading to a deeper fundamental understanding of the molecular interface. Most importantly, these results have shown that high-resolution SFG is a powerful tool that has opened new opportunities in surface/interface studies.

Molecular-level Analysis and Microscopy Studies of Brown Carbon

Alexander Laskin

We are focused on laboratory studies that investigate the relationship between chemical composition of brown carbon (BrC) particles and their optical properties.

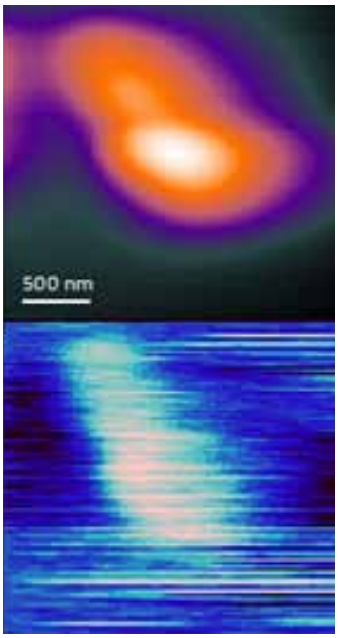
Our present understanding of atmospheric aerosol chemistry is limited to the extent that its impact on the environment and climate cannot be accurately predicted and mitigated. Aerosols can be predominantly scattering (e.g., stratospheric sulfuric acid aerosol), predominantly absorbing (e.g., black-carbon or soot), or have intermediate optical properties (e.g., light-absorbing organic aerosol, BrC). Although the first two types of aerosols have been extensively studied, BrC aerosol has been only recently recognized as a significant environmental factor. Little is known about the formation, composition, and physical properties of BrC.

Understanding and mitigating the environmental impacts of energy production and use have been identified as one of the important priorities of DOE. The production and use of energy often leads to the formation of aerosol particles that affect atmospheric chemistry and climate change. To help achieve DOE's vision of environmentally responsible production and use of energy, it is necessary to improve our understanding of the origin, fate, chemical, and physical properties of fine atmospheric particles.

Atmospheric aerosols affect climate in two major ways. The indirect effect arises from the nucleation of cloud droplets and ice crystals by aerosol particles. The direct effect arises from absorption or scattering of sunlight radiation by particles. Most atmospheric aerosols are characterized as "white" because they efficiently scatter visible radiation. However, a significant and highly variable fraction of aerosols absorbs radiation including soot, mineral dust, and BrC organic aerosols. Recent studies have shown that BrC aerosols are widespread in urban environments where they may dominate the total aerosol absorption. These observations challenge present understanding of the absorption properties of organic aerosols, and call for systematic studies to unravel the chemical composition and absorption properties of BrC.

Our research maximizes the novel analytical platforms developed in the user facility at PNNL, the Environmental Molecular Sciences Laboratory (EMSL), applied to complementary molecular-level and microscopic analysis of BrC particles. We used nano-DESI/HR-MS, SFG, and IR SNOM spectro-microscopy capabilities to provide unique insights on the mechanisms of atmospheric formation and evolution of BrC.

During the first year of the project (FY 2013) and in collaboration with the University of California-Irvine and the University of Colorado, we identified one type of BrC chromophores as highly conjugated nitrogen containing compounds (NOC) produced as a result of chemical reactions between the carbonyl species in SOA and ammonia. We concluded that ketolimononaldehyde (KLA, $C_9H_{14}O_3$), a second-generation ozonolysis product of limonene, is an important BrC precursor that produces a range of visible light-absorbing chromophores through its reactions with ammonia and amino acids. Our efforts focused on the molecular speciation of chemical composition and optical properties of these NOC chromophores and are carried out along a couple of major direction. Chromophores are further characterized using a combined LC–UV/Vis–HR-MS platform for SOA analysis, which allows us to focus on an LC-separated subset of compounds that absorb radiation at a specific wavelength measured by an inline UV/Vis detector. In addition, double resonance sum frequency generation (DR SFG) spectroscopy was applied to identify characteristic vibrational frequencies of BrC chromophores in aerosol samples. In this approach, SFG provides amplified signal only when both the IR and UV/Vis characteristic frequencies are in resonance with the characteristic vibrational and electronic states of the targeted chromophores allowing spectroscopic fingerprint of BrC. Finally, results of the DR SFG measurements will guide IR SNOM spectro-microscopy measurements at nm-scale with potential to enable characterization of the lateral distribution of BrC chromophores in individual particles and their evolution during atmospheric life cycle.



First time demonstration of the SNOM mapping of an organic aerosol particle. Upper panel: particle topography; lower panel : IR SNOM image at 1720 cm^{-1} (absorption by carbonyl groups).

In FY 2014, we will continue to apply these techniques to probe plausible molecular structures of NOC chromophores, identify their characteristic IR vibrational frequencies, and explore their possible lateral distribution in individual particles collected in laboratory-based studies of model organic aerosol and field-collected ambient particles from real-world environment.

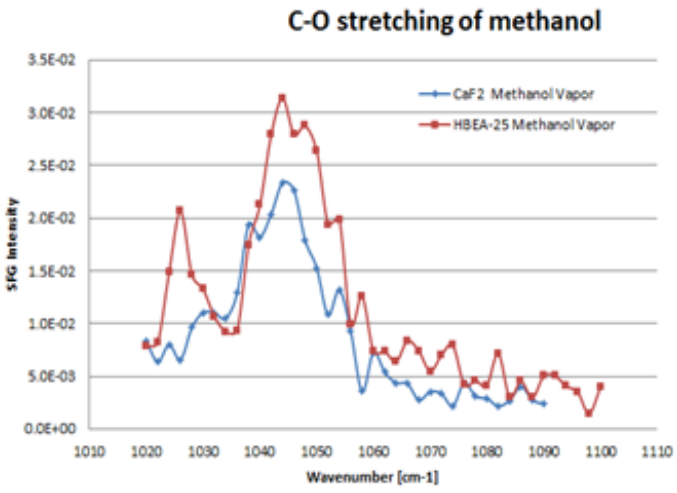
PN13080/2561

Monitoring Sorbed and Reacting Species of Lignin Fragments in Zeolites in an Aqueous Environment

Hongfei Wang

This study is developing both a molecular-level understanding of and predictive models for critical interface properties related to energy production and storage.

The central role and critical importance of catalysis in an energy future based on sustainability, combined with the insight that developments must be knowledge-based, have motivated significant efforts to understand catalyzed processes and develop new catalytic routes from this information. While research has yielded impressive insight into gas-solid catalyzed reactions, the information about surfaces in contact with liquids (particularly water) are far less established. The investigative chemistry of liquid-solid interfaces at the molecular level has not been easy as many powerful surface techniques that have been developed over the years are suitable only for exposed surfaces studies. In general, optical methods can access the buried solid-liquid interfaces. The chemistry of buried interfaces, however, depends on the ability to discriminate selectively between the few molecule layers at that interface and the much larger number of molecules that exist in the two bulk phases involved. This limits the application of common optical methods such as FTIR and Raman spectroscopy.



C-O stretching at the CaF_2/methanol vapor and HBEA-25/methanol vapor interface. The 1030cm-1 peak is unique for the zeolite surface and needs further investigation.

We are developing in situ sum frequency generation (SFG) spectroscopy to monitor adsorbed complexes in zeolites submersed in an aqueous solution of reactants. Specifically, we are exploring the interactions of target molecules by probing the vibrational spectroscopy (VS) of the adsorbed and solvent molecules on the zeolite surface using the SFG-VS. Previously, C-H stretching vibrational bands in the 2800–3000 cm^{-1} region of adsorbed molecules on catalyst surface in solution have been used extensively. By extending the detection to 1000–4000 cm^{-1}, SFG can probe the solvent molecules and the fingerprint region vibrational spectra of the groups for molecules adsorbed or interact with the zeolite surface. The competitive adsorption of solutes and mixed solvents at the zeolite interface can be explored, enabling molecular vibrational signatures to characterize and monitor the molecular adsorption and interactions at the solid-liquid interfaces with chemical specificity.

A late FY 2013 project start, the project obtained the following results in establishing protocols for solid-liquid interface measurement, characterizing the baseline SFG spectra for zeolite material, and observing SFG spectra in C-H, C-O, and O-H region with and without adsorbed molecules. These results provided benchmark and directions for future development.

Protocols were established for making SFG measurement on powdered material such as zeolite in contact with liquid. To gain insight into the surface chemistry and adsorption phenomena of bio oil derived compounds on zeolites, H-β-25 purchased from zeolite was chosen. For close contact between the sample and optical window (calcium fluoride, CaF_2) that couples SFG signal out for detection in the 1000–4000 cm^{-1} spectra range, pressed pellet, and more sophisticated techniques like spin coating and drop casting were developed. For further studies, drop casting was applied for simplicity and reproducibility, for which a slurry with a solid concentration of 4 mg/mL in water was prepared. A droplet of 100 µl was casted on the CaF_2 substrate, and the window was heated to 350 K to accelerate drying and remove excess surface water. A film thickness of ~1–10 µm was achieved.

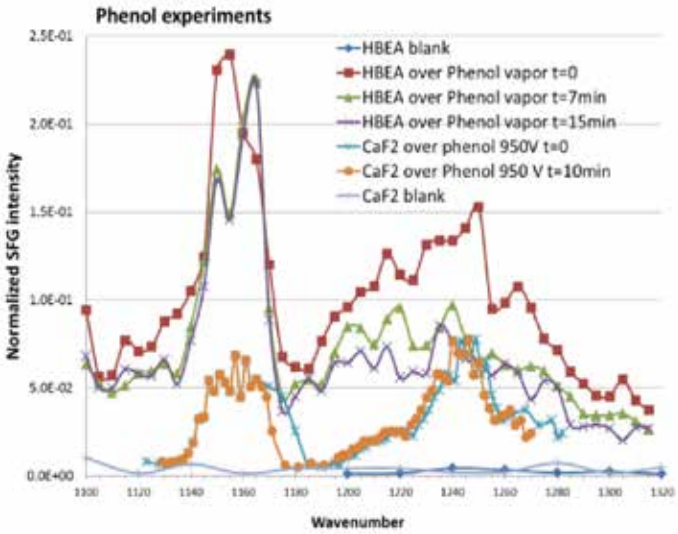
Initial experiments were addressed to observe signals rising from the zeolite. According to IR studies, framework vibrational modes should be indicated at frequencies

PN13094/2575

between 1000 and 1400 cm^{-1}; additionally, O-H stretching vibrations between 3500 and 3800 cm^{-1} were observed. In SFG measurement, these features are not present due to centro-symmetric structure of the zeolite unit cells. This provided a zero background to observe adsorbed species on the zeolite surface. It was found that even for dry zeolite powders, SFG spectral features in the O-H stretching region are observed (3200–3600 cm^{-1} for hydrogen bonded O-H and 3700 cm^{-1} for the dangling O-H of water) indicating predominant affinity between water molecule and the zeolite surface.

To study the adsorption of bio oil derived molecules, both the gas-solid and liquid-solid interfaces were studied. Probe molecules are small alcohols such as methanol and ethanol; more complex alcohols such as isopropanol and glycerol; and cyclic alcohols such as phenol and cyclohexanol. For all molecules, distinct peaks in the C-H stretching region (2800–3100 cm^{-1}) could be observed for the air-liquid, solid-liquid, and the solid-gas interface. Alternatively, C-D stretching modes were successfully obtained at solid-liquid interface around 2000 cm^{-1}. The signal for the C-H stretching on the solid-liquid interface is usually much weaker as it is for the solid-vapor interface. This can be understood by the difference of the refractive indexes of the two adjacent phases.

Unlike C-H signals, functional groups such as C-O stretching mode around 1000-1300 cm-1 are of high importance to study different adsorption phenomena. We recently discovered C-O stretching modes for methanol at the solid-vapor interface of H-β-25 (~1040–1050 cm^{-1}), though such a feature becomes non-detectable for the solid-liquid interface. One reason might be that the signal intensity is too low at low wave numbers. Phenol C-O stretching is easily obtained at 1250 cm^{-1} at air solution (usually 7 wt%), solid solution, and solid-gas interfaces.



C-O stretching at the CaF_2/phenol vapor and HBEA-25/phenol solution interface. The apparent spectral feature difference indicating different chemistry at the solid/solution interface.

SFG effects on powdered catalyst materials are still preliminary. However, results have shown that it is possible to monitor adsorbed molecules and the water solvent in the broad spectra range 1000–4000 cm^{-1}. The results seem indicate that at room temperature, the zeolite surface predominantly adsorbs water molecules and the competitive adsorption between the solvents, and target molecules at different temperature might be a key step for the catalytic reaction. Once molecules are detected, further experiments under controlled temperature, pressure, concentration, and pH may help understand the processes and mechanisms on the surface interaction and reactivity of the zeolite catalyst.

NMR Spectroscopy of Pu-Containing Materials

Herman Cho

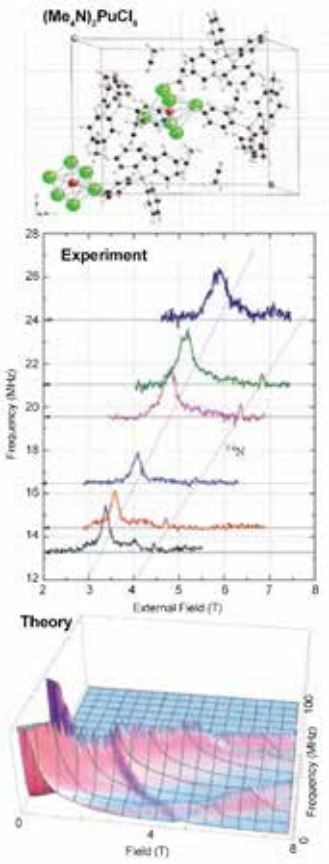
We are in an unprecedented position to demonstrate how nuclear magnetic resonance (NMR) spectroscopy may be used to elucidate structure and bonding of plutonium (Pu)-containing compounds.

Plutonium commands intense attention for its promise as a carbon-neutral energy source and the enduring danger its stockpiles present to mankind and the environment. Nevertheless, research on the remarkable chemistry and physics of plutonium has been severely constrained by the inability to perform routine but critical experimental measurements on samples containing this highly hazardous element. In addition, the absence of experimental tools adapted for investigation of Pu-containing samples is exemplified by NMR spectroscopy, which despite its importance in almost 70 years of chemical, physical, and materials science research has contributed to only a handful of published studies of Pu compounds. Although the spin-1/2 ^{239}Pu isotope has nuclear parameters favorable for NMR spectroscopy, there has been no concerted effort to observe its NMR signal, and the few previous NMR experiments on Pu-bearing samples have instead been directed towards detection of surrounding nuclides.

The oxides of Pu are at the center of its beneficial uses as well as its most vexing difficulties. The simple binary formula and cubic crystal structure of PuO_2 belie an amazingly complex behavior, and are of particular interest for scientific studies. Along with more complex systems, NMR determination of the magnetic and electronic environment at lattice sites of PuO_2 will provide definitive insights on decades-long questions on the nature of bonding in Pu compounds.

To demonstrate feasibility and lay the groundwork for a long-term program on NMR spectroscopy of Pu systems, the following tasks were undertaken. Initially, a collaboration was formed with the Seaborg Institute and the Condensed Matter and Thermal Physics Group at Los Alamos National Laboratory (LANL). A week-long visit to LANL to participate in experiments and discussions of research directions marked the start of this early FY 2013 collaboration. This arrangement will allow LANL's world-class capabilities in actinide synthesis and characterization to be combined with PNNL's state-of-the-art resources and expertise in magnetic resonance studies of radioactive materials. Ultimately, a program for future cooperative research was developed.

Also in FY 2013, analyses of NMR data on PuO_2 and $(Me_4N)_2PuCl_6$ were initiated to assign and interpret the shapes and positions of observed lines in terms of local nuclear environments. In particular, a second moment calculation of the presumptive ^{239}Pu line was performed for the PuO_2 face-centered-cubic lattice, and variable-field powder lineshapes of the ^{35}Cl/^{37}Cl resonances from $(Me_4N)_2PuC_{l6}$ as a function of the applied frequency were computed by numerical diagonalization of the nuclear Hamiltonian. In both cases, it appears basic models that include only homonuclear dipolar couplings and electric field gradient interactions are not sufficient to account for experimental lineshapes, and factors such as defect sites and hyperfine couplings must be added to the description of the nuclear environments.



Crystal structure of $(Me_4N)_2PuC_{l6}$ (top); NMR spectra obtained by sweeping the magnetic field at the indicated constant frequencies (middle); spectra computed by assuming the observed resonance is from the ^{35}Cl/^{37}Cl NMR.

Further, procedures were formulated to ship Pu samples synthesized by LANL to Richland, Washington for high resolution measurements on state-of-the-art NMR instruments at PNNL. Preliminary results and a prospectus of NMR studies of Pu-containing compounds were presented at an international NMR meeting.

PN13077/2558

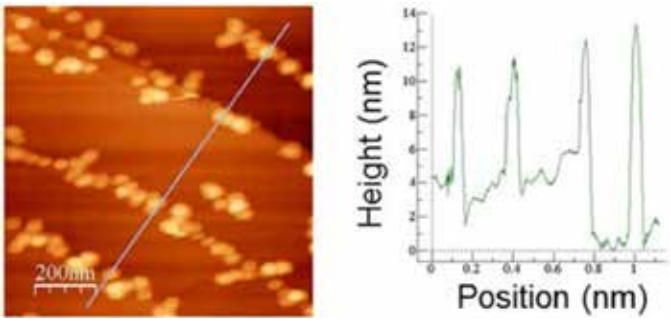
Novel Alloy Nanoparticle Materials for Catalysis and Energy Storage

Grant E. Johnson

This project is developing one-of-a-kind capabilities for the controlled preparation and characterization of alloy nanoparticles (NPs) of precise composition for applications in fuel cell catalysis and energy storage. The emerging capabilities are being applied to establish structure-reactivity relationships for alloy NPs.

We are focused on understanding how the size, elemental composition, surface density, and support-interaction of alloy NPs may be tuned to improve catalytic activity, selectivity, and durability toward the oxygen reduction reaction (ORR). Custom modifications are underway on a unique nanocluster deposition system that will enable atomically-precise control of the size and elemental composition of alloy NPs soft landed onto surfaces. A non-thermal physical synthesis technique is being developed to produce binary and ternary alloy NPs across a range of sizes and compositions that cannot be prepared using traditional solution-phase methods. New nanomaterials with promising catalytic properties are being discovered using this approach. A highly sensitive surface analysis technique is also being developed to screen the catalytic activity of NPs supported on surfaces at ambient conditions with minimal sample preparation. Achieving precise control of NP surface coverage and size is necessary to enable catalytic activity to be measured and calculated accurately as a function of metal loading and exposed surface area.

Employing simultaneous direct current magnetron sputtering of up to three independent metal targets in a single region of inert gas-aggregation, we produced anionic binary and ternary alloy NPs that contain a variety of transition metals (e.g., Fe, Co, Ni, V, Cu, and Pt). Mass-selected NPs were soft landed onto conductive glassy carbon (GC), highly-oriented-pyrolitic graphite (HOPG), and indium-tin-oxide (ITO) coated glass surfaces that are compatible with electrochemical analysis in solution using cyclic voltammetry (CV). We controlled NP coverage soft-landed onto substrates by measuring the ion current at the surface and adjusting the deposition length. The NP surface coverage was determined by screening the surfaces with atomic force microscopy (AFM) following deposition. Using this approach, we demonstrated that on flat substrates such as GC, NPs locate randomly on the surface at low coverage. By comparison, stepped surfaces such as HOPG NPs nucleate preferentially along step surface edges, forming extended linear chains and leaving the terraces unoccupied. With increasing coverage, we observed with AFM that terrace sites fill until a dense layer of NPs is present across the HOPG surface. We also determined the size and size distribution of deposited NPs from the height measured frequency distribution. To validate the AFM observations, we analyzed the surfaces using scanning electron microscopy (SEM), which confirmed that at low coverage, the soft landed NPs are separated and located randomly on the flat GC surface. SEM results also verified that NPs line up at step edges at low coverage on stepped HOPG surfaces.



Left: AFM image of 8 nm diameter Cu NPs soft landed onto graphite; right: height profile across rows of Cu NPs.

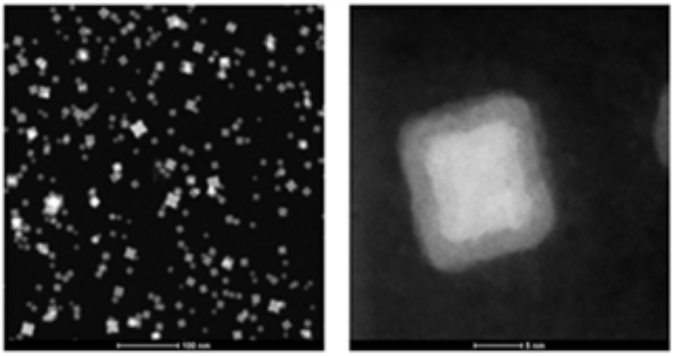
In addition to controlling the size and surface coverage of soft landed NPs, it is desirable to tune their morphology and elemental composition. In a similar vein, elemental composition may be adjusted to reduce loading expensive precious metals (e.g., Pt and Au) and bring about physical lattice compression and strain as well as electronic structure effects that improve catalytic activity and selectivity. We employed systematic adjustments to the sputtering power applied to each of the three metal targets, the flow of sputtering (Ar) and buffer (He) gas, and the length of the gas-aggregation region to tune the morphology and elemental composition of NPs formed by magnetron sputtering and deposited them onto surfaces.

While the majority of NPs produced using this physical synthesis approach are spherical, we demonstrated that in the case if vanadium (V), it is possible to produce highly monodisperse nanocubes and soft land them onto carbon surfaces for subsequent atomic-level characterization using scanning transmission electron microscopy (STEM) at EMSL. As shown, the wide-field STEM images established the cubic shape and similar size of the individual NPs on the surface.

PN13005/2486

Atomically resolved near-field STEM images revealed sharp differences in contrast of the center and edges of the V NPs, indicating the presence of metallic V cores surrounded by thin V_xO_y shells. In addition, we achieved control of the elemental composition of alloy NPs. X-ray photoelectron spectroscopy (XPS) was used to confirm the incorporation of multiple metals into size-selected alloy NPs soft landed onto surfaces. A multimodal approach involving analysis by medium energy ion scattering (TOF-MEIS) and STEM confirmed the controlled creation of Pt-M (M = V and Cu) NPs with homogeneous alloy compositions and core-shell morphologies. In particular, STEM combined with elemental line scans using electron energy loss spectroscopy (EELS) demonstrated the formation of V-core Pt-shell NPs. These particles will make an efficient use of a limited loading precious Pt metal by concentrating it in the NP shell, not in the inaccessible NP core. The catalytic activity and stability of alloy NPs on GC was evaluated using CV under acidic solution conditions. A promising trend of increasing activity toward the ORR with incorporation of early 3D transition metals into the Pt-alloy was observed for NPs of similar size and surface coverage.

In FY 2013, a paper on the synthesis and characterization of sub-nanometer gold clusters that are promising candidates for application in oxidation catalysis was published in *ChemPlusChem*. A second manuscript that describes the preparation of alloy NPs and their characterization using a multimodal combination is under review at the *Journal of Vacuum Science and Technology A*. A third manuscript that



STEM images of 8 nm V NPs soft landed onto carbon.

will illustrate our abilities to synthesize and control NP size, morphology, and composition on surfaces is in preparation. In addition, the technical and scientific accomplishments summarized in the preceding paragraphs were presented at a national conference. We also established an international collaboration with the DGIST group in Korea, where they are analyzing the supported NPs using a unique ion scattering technique (TOF-MEIS). A second collaboration is underway with the University of Southern California-Los Angeles on the controlled synthesis and characterization of titania NPs, which are widely used in photovoltaic and catalytic applications. We designed, constructed, and installed the high-resolution mass-filter that will enable stoichiometrically selected deposition of alloy NPs onto surfaces.

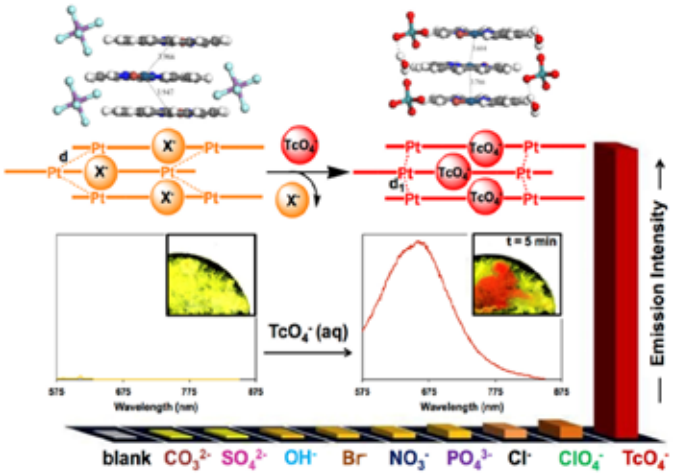
Novel Inorganic Complexes for Tc Management in the Tank Waste

Tatiana Levitskaia

This project is designing and testing a new class of inorganic sorbents and sensor materials with high selectivity for technetium (Tc-99), guided in part by models obtained from computational chemistry simulations.

One of DOE's current challenges is meeting regulatory requirements for Tc-99 disposal. The radioactive component Tc-99 is problematic in high level waste (HLW) at the Hanford and Savannah River tank farms due to its long half-life, complex redox chemistry, high solubility, and volatility at elevated temperatures creating the potential for leaching in the environment. Removal of Tc-99 from low activity waste (LAW) streams will eliminate vitrification and storage issues associated with volatility. In strongly alkaline environments typical of Hanford tank waste, Tc-99 is found both as Tc(VII) in the form of pertechnetate TcO_4^- as well as low oxidation-state Tc(I) tricarbonyl $Tc(CO)_3^+$ species. This research contributes to enhancing knowledge of Tc-99 speciation in liquid waste and investigates new approaches to its removal by controlling redox chemistry of Tc-99 species. The availability of the detection and sorption capabilities will enhance our understanding of the Tc-99 sorption mechanism and behavior in the tank waste, leading to the design of new separation technologies.

We are targeting Tc-99 selective materials that possess inorganic framework compatible with the vitrification process. Novel inexpensive, readily available transition metal-based materials promise new capabilities for the select uptake of TcO_4^- and potentially non-TcO_4^- species from LAW streams. The oxide/hydroxide structure of these materials is similar to hydrotalcite naturally occurring minerals and clays compatible with highly alkaline matrices. Composite oxide/hydroxide materials possess superior characteristics, including large surface area, high sorption capacity, and chemical/mechanical/thermal stability to offer a simple, efficient, economic option for the selective removal of the TcO_4^- anion. Specifically, we demonstrated that solid salts of square-planar transition metals such as Pt(II), Pd(II), Ni(II) or Cu(II) complexes are capable of preferential TcO_4^- uptake by means of exchange with their parent counter anions. Similar approaches can be used for the removal and detection of non-pertechnetate species.



Selective detection and quantification of TcO_4^- in the presence of interfering anions.

The focus for FY 2013 was designing new inorganic materials for TcO_4^- detection and sorption. To investigate the TcO_4^- uptake capacity, we synthesized and tested several inorganic composite materials of the formulae $[M_{1-x}^{2+}M_x^{m+}(OH)_2]^{x+}(A^{n-})_{x/n}.yH_2O$, where M^{2+} is a divalent cation (Ni^{2+}, Sn^{2+}, Zn^{2+}, Fe^{2+}, Co^{2+}, Cu^{2+}, Ca^{2+}) and M^{m+} is a trivalent or tetravalent cation (Al^{3+}, Cr^{3+}, Zr^{4+}). The value of x is equal to the molar ratio of $M^{m+}/(M^{2+} + M^{m+})$, where A is the interlayer anion of valence n (e.g., hydroxide, oxide, and phosphate). Some commercially available reagents (bonechar and hydroxyapatite) were chemically modified and tested for pertechnetate uptake. The TcO_4^- uptake capacity of these composites was tested both from neutral solutions and highly alkaline solutions consistent with tank-waste supernatant conditions. It was observed that the highest capacity was exhibited by composites where Sn^{2+} was present.

To quantify TcO_4^- in liquid matrices, we designed a novel colorimetric and luminescence detection technique capable of detecting TcO_4^- nano-molar quantities from a mixture containing interfering anionic species (ClO_4^-, NO_3^-, Br^-, Cl^-, I^-, PO_4^{3-}, BrO^{3-}, OH^-, SO_4^{2-}). Quantitative measurement of aqueous TcO_4^- anion in a 10^{-8} to 10^{-3} M concentration range was achieved using solid $[Pt(tpy)Br]SbF_6$ (tpy=2,2′:6′2″ -terpyridine) material. This novel detection method relies on the color change of the Pt(II) complex and intense luminescence response upon SbF_6^- exchange with TcO_4^- due to concomitant enhancement of Pt•••Pt interactions. This spectroscopic response was highly selective for TcO_4^-.

Computational modeling and simulation were done to provide insight into the molecular structure and elucidate the sensing mechanism of the sensor materials. We constructed

PN13036/2517

a novel density functional theory model for $[Pt(tpy)Br]SbF_6$ sensor material that has helped us better understand the sensing mechanism. Absorption spectra of the parent and $[Pt(tpy)Br]TcO_4$ compounds were simulated. The model illustrates that presence of an aqueous environment is key to the sensing module. Upon addition of water molecules to the $[Pt(tpy)Br]TcO_4$ system, Pt•••Pt distances become shorter and more aligned generating the unique spectroscopic response. These modeling studies have a critical role in designing highly selective sensing medium.

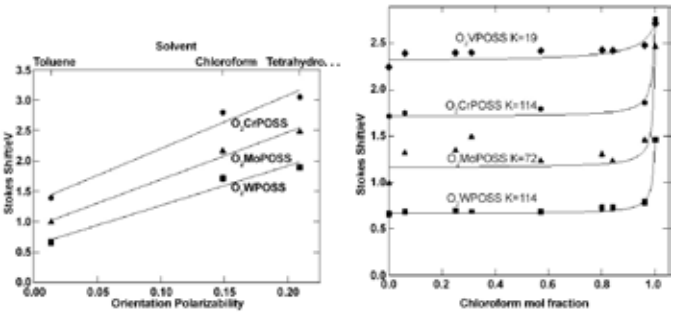
While the chemistry of Tc-99 in higher oxidation states such as Tc(IV) and Tc(VII) is sufficiently understood, Tc(I) and particularly $Tc(CO)_3^+$ species properties are unknown. In FY 2014, our approach is to investigate the potential to control the redox chemistry of non-TcO_4^- species and design hybrid redox-active inorganic sorbents that will remove it from the waste supernatant and simultaneously oxidize Tc(I) compounds to TcO_2. Identification of the redox behavior and nature of non-TcO_4^- species will enable rational design of the methods to remove these species from the alkaline matrices.

Probing Composition and Structure of Polarizable Reaction Mixtures Inside the Pores of Supported Metal-oxo Catalysts

Robert S. Weber

We are combining PNNL's advanced measurement, computational, and chemical synthesis capabilities to describe and understand the title catalysts and their interaction with the reactants and products of biomass conversion.

The process of converting biomass into liquid fuel requires the use of catalysts to accelerate and direct the chemical reactions that lead from the original plant material to transportation fuels (or their precursors) that can be further processed in existing refineries. This project provides a way to quantify the detailed (molecular scale) composition and structure of liquid reactants and products in contact with a catalyst. The research will therefore help us understand one aspect of how the catalysts work and identify ways in which they can be improved. While there has been considerable research into the molecular structure of the catalysts themselves, much less is known about their interaction within their immediate environment. Because rates of chemical reactions depend on the concentrations of reactants and products, it is necessary to measure those concentrations as locally as possible, preferably right at the catalytic sites so we can learn how to guide molecules to and from the sites. This project will yield those compositions and information about the orientation of molecules that can affect or be involved in the catalyzed conversion of the biomass into fuels.



Left: A calibration curve showing that the spectra of the model compounds can be used to identify the polarizablity of the solvent in which the compounds had been dissolved. Right: The compounds are preferentially solvated by just one of the components of a binary solution (toluene in the case of a toluene/chloroform mixture).

We adapted the Stokes shift spectroscopy (measurement of light absorbed and emitted from catalysts) to study the immediate solution environment of the catalyst site. The technique is quite generalizable because even apparently clear materials absorb light at least in the ultraviolet and infrared ranges. Light emission is familiar to anyone who has pasted glow-in-the-dark "stars" to the ceiling of a child's bedroom. We measure the changes in spectra, particularly shifts in energies of spectral features (color) and changes in relaxation times on a scale comparable to molecule movements (millionths to billionths of seconds; glow-in-the-dark stickers relax more slowly, typically over many minutes).

During FY 2013, we prepared and characterized a series of model compounds that resemble isolated sites of metal oxides supported on silica. The model compounds were used to construct a calibration curve from which the composition of the liquid phase around the dissolved compoounds could be inferred from their absorption and emission spectra. Next, we used that calibration to explore the surrounding of the model compounds in mixtures of the solvents (shown in the left side of the figure). The spectra of the compounds changed remarkably less than we expected when the model compounds were dissolved in mixtures of toluene and chloroform (two organic solvents that totally miscible) for compostions ranging from 100% toluene to 95% chloroform. We interpret the constancy of the spectra (horizontal regions shown in the right side figure) as evidence for the preferential solvation of the compounds by the toluene, a surprising result given that all three species—toluene, chloroform and the model compounds—are completely miscible.

Additionally and as promised in FY 2013, we extended our work to a more realistic set of catalyst samples, namely materials in which the catalytic sites were supported on solid, highly porous silica. We determined that the model compounds can be activated with light to act as hydrogenation catalysts.

In FY 2014, we will extend the photocatalytic experiments to include the upgrading of a range of molecules found in feedstocks for biofuels. We will also assay the interactions of the chromophores when they are in contact with "1-D" and "2-D" reaction media in the pockets and pores of solid supports. We will then be in a position to capture the effects of preferential solvation on the rates of the reactions photocatalyzed by practicable catalysts.

PN12062/2463

Probing Structural Dynamics with High Spatial and Temporal Resolution

Nigel Browning

We are developing a fundamental understanding of materials dynamics (from µs to ns) in systems where the required combination of spatial and temporal resolution can be reached only by the dynamic transmission electron microscope (DTEM).

This research will develop the DTEM capability to perform single-shot *in situ* measurements in liquid and gas stages with a combined spatio-temporal resolution at least four orders of magnitude faster than any other competing technique (including aberration corrected scanning transmission electron microscope [STEM]/TEM). In this temporal regime, the DTEM is expected to have atomic spatial resolution, providing an *in situ* TEM capable of studying nanoscale dynamic phenomena with several orders of magnitude time resolution advantage over any existing *in situ* TEM. In addition to providing unique insights into the inorganic and organic systems studied, this research will test the limits of temporal resolution and define electron optics and alignment for the next generation of truly ultrafast TEM (i.e., sub-ns). For nanoclusters in fluid around an electrode, interactions between the particles leading to the nucleation and growth of thin films will be examined for the first time. The ability to observe "live" biological systems will provide unique insights into biological systems.

The primary goals accomplished in the first year of the project (FY 2012) were relocating the DTEM from the University of California (UC)-Davis and installing the microscope within EMSL. In FY 2013, the two laser systems needed to convert the microscope from a field emission microscope to a DTEM were installed. The first laser system creates a ultraviolet pulse used to create the photoemission pulse of electrons (this defines the DTEM resolution) while the second laser system creates an optical pulse used to stimulate the specimen (i.e., creates the reaction in the DTEM). In addition to the laser system, the final part of the DTEM is the modification of the electron gun. Working with JEOL Ltd., Energy Beam Sciences, and consultants from the University of Michigan and University of Illinois at Chicago, the gun was designed, assembled, and is currently under testing on the DTEM column. Now complete, all that remains on the DTEM is the final testing stages of the instrument, which are expected to be completed within the early part of FY 2014.

Our work continues to investigate the propagation of the electron beam by building a complete model of the electron optical system within the electron microscope. The overall aim of this work remains generating an understanding that will enable an increase in the spatio-temporal resolution. During FY 2013, our collaboration with the University of Michigan and CEOS Ltd. (the latter of which is the manufacturer of the aberration correctors within the DTEM) has worked to build models for the key electron optical components. As a result, our agreements provide the key proprietary instrument parameters for the model in place, a completed version of which will be available by FY 2014.

The dynamic transmission electron microscope (DTEM)

In addition to the move and equipment installation, work commenced to develop new technologies for the implementation within the instrument framework. Our initial work involved a collaboration with the National Center for Microscopy and Imaging Research at the University of California, San Diego. The goal was to determine the ability of direct electron detectors to provide an increase in the signal-to-noise over conventional CCDs, thereby increasing the overall spatio-temporal resolution of the DTEM. This work showed that a simple frame shifting device plus detector should be able to generate a frame interval of ~50 µs.

The first direct electron camera (capable of 100Hz continual acquisition movies) will be loaned to PNNL starting in FY 2014. This instrument will be used to demonstrate the imaging capability while allowing for unique movies of dynamic processes to be acquired.

PN12063/2464

Quantitative Imaging of Atomic Scale Chemistry Changes at Interfaces

Nigel Browning

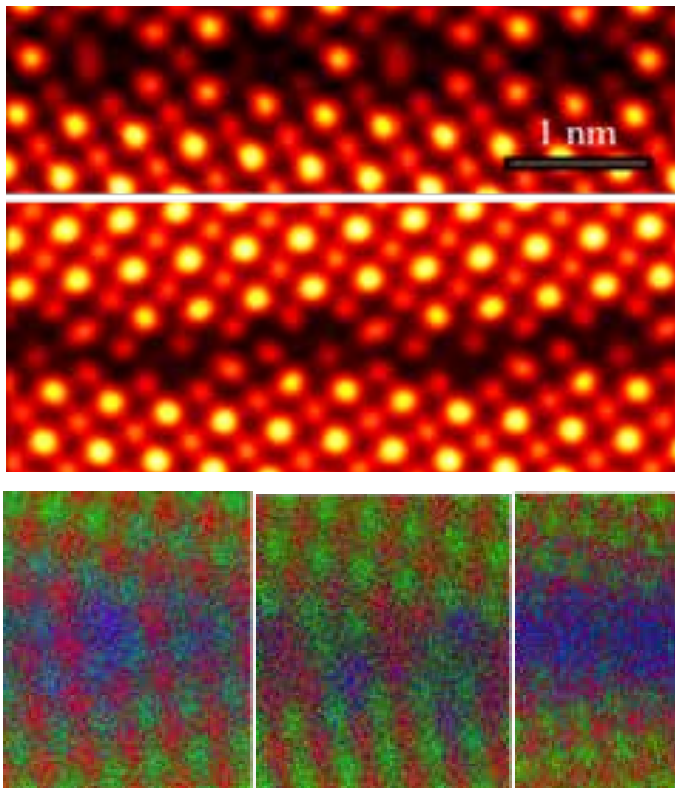
This research will develop the statistical methods necessary to quantify changes in the atomic structure and the chemistry that occurs at interfaces and in nanoparticles.

This project is developing a framework for quantitative chemical imaging of interface/nanoparticle structures under environmental settings, specifically different ambient gas conditions. Quantitative chemical imaging has been applied only ad-hoc and has not been used for detailed comparisons between experimental results and first principles simulations. A quantitative framework for image interpretation is the only way to compare atomic-scale determination of interface structure and composition with larger scale property measurements.

During FY 2012, we studied a range of grain boundary structures in $SrTiO_3$ and CeO_2 and compared them to face-centered-cubic (FCC) metal structures to determine a set of rules for the formation of grain boundary structures. Such analysis has the potential to create a set of crystallographic rules for interfaces that will allow the understanding of defect properties, and therefore the synthesis of new materials for specific applications.

For FY 2013, we investigated the effect of cation doping on the grain boundary plane structure. Initial results from $SrTiO_3$ grain boundaries prepared to be intentionally doped with Pr and Eu suggest that dopants affect more than just chemistry. The presence of Pr in the structure seems to limit the grain boundary plane to one particular structural variant, a symmetric structure with no rigid body shift. This means that composition fluctuations in the nominally undoped grain boundary plane caused two equal energy structures to form, while the intentional doping of Pr on the other hand can shift the energy balance to the symmetric boundary plane. The Eu doping had the opposite effect and pushes structural stability in the other direction. In addition, the structure is much more variable along the boundary plane.

To clarify this effect, additional work in FY 2013 used aberration corrected (scanning) transmission electron microscopy (STEM) techniques to evaluate the distribution of dopants inside a specific grain boundary plane. To study the fundamental principles of doping behavior in $SrTiO_3$ GBs, bicrystals are used as model structures. A thin layer



Atomic resolution Z-contrast images (top two) show distinct boundary structures (symmetric and rigid body shift). Energy dispersive X-ray maps show the location of Pr, Eu/W, and Mn (respectively, left to right) dopants (blue) in the boundary plane.

(~5 nm) of the dopant was coated onto the surface of one crystal prior to diffusion bonding to create the impurity concentration at the boundary plane. The diffusion bonding was carried out at the highest temperature (1350°C for over 20 h) for sufficient diffusion. TEM samples were then prepared by mechanical polishing and ion-milling. Atomic resolution Z-contrast imaging and electron energy loss spectroscopy (EELS) characterizations of the GB atomic structures were performed. For GB chemical analysis, atomic resolution EDS analysis was performed using a 200kV FEI ChemiSTEM installed with an aberration corrector and silicon drift detector (SSD) at Sandia National Laboratory.

Using statistical analysis of the grain boundary plane, we see that the type of dopant for the structure changes the prevalence of the two undoped structures in the boundary plane (i.e., atomic scale doping effects control the meso-scale energetics of the boundary plane). Using energy dispersive spectroscopy (EDS), we can investigate how the

PN12067/2468

location of the dopants causes this effect. For Pr, dopants segregate to the Sr and Ti sites in the boundary plane; segregation for Eu/W dopants occurs primarily on Ti, whose specificity changes the prevalence of the symmetric structure from 35% (undoped grain boundary plane) to 81% (Pr) to 41% (Eu/W). Intuitively, we may expect that these statistical prevalences (energetics) follow dopant distribution: Eu/W seems to segregate to specific Ti locations and has a small effect on the mesoscale structure. On the other hand, Pr segregates to many sites (both Sr and Ti) in the grain boundary plane and has a large effect on the mesoscale structure.

One of the main results obtained in FY 2013 is the analysis of Mn-doped grain $SrTiO_3$ grain boundaries. Mn-doped $SrTiO_3$ shows promising magnetic and electrical properties, but until this time, the doping mechanism has been unclear. During this project and using EDS and EELS techniques, Mn^{4+} was found to substitute for Ti in bulk $SrTiO_3$, but it is Mn^{2+} that segregates to the grain boundaries, substituting for Sr^{2+} and being present at interstitial sites. Mn interstitial doping has never been previously reported but is found possible here with the formation of Sr vacancies. This finding is significantly different from the amphoteric doping of Mn^{2+} substituting Sr and Mn^{4+} substituting Ti sites, therefore leading to a completely different understanding of the defect mediated electrical and magnetic properties of these transition metal doped perovskites.

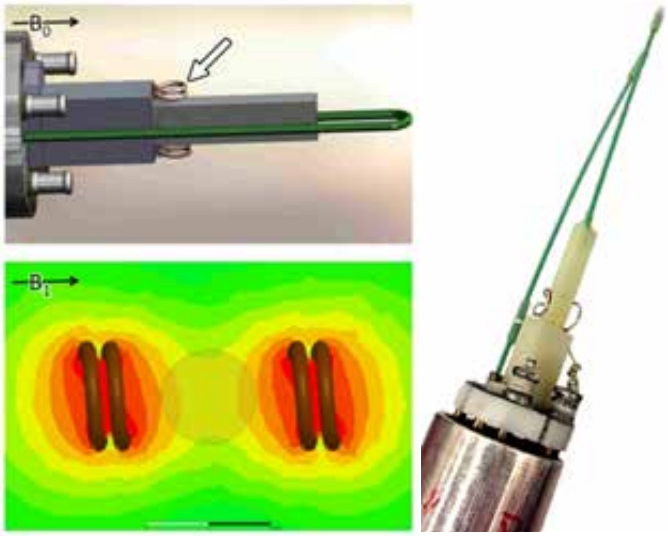
This FY 2013 result builds on our main success from the previous year: development of a classification scheme for grain boundaries in metals and ceramics based on the crystallography of parent compounds. This classification scheme is important for future analysis of grain boundary properties as it allows the "base structure" of all boundaries to be predicted without the need for extensive experiments and simulations (these structures were used to understand the doping effects seen this year). In the materials used in energy technologies, these types of interfaces are key components in the overall properties. The properties themselves are typically characterized by ionic/atomic diffusion into and out of the interface structure. This characterization framework for interfaces means that future experiments that will be performed on individual grain boundaries can be related back to a more extensive set of systems based on the similarities between the structures and chemistries of parent compounds. For our work in FY 2014, we plan to complete the analysis of this set of interfaces by investigating the effect of oxygen vacancies on the grain boundary plane using *in-situ* methods. In all practical applications, it is the oxygen vacancies that are the mobile species. Building on the framework of the base structures and effects of cation doping already analyzed (in FY 2013), this should allow an atomic scale model to be developed to produce an understanding of ionic conductivity at oxide interfaces.

Simultaneous Electrochemical and Nuclear Magnetic Resonance Techniques for the Study of Electrochemically Active Biofilms

Ryan S. Renslow

Understanding the microscale chemical gradients in electricity-producing biofilms and the ability of these microorganisms to use extracellular electron transfer can lead to novel applications in bioremediation, alternative energy production, and electrochemically stimulated synthetic biology to produce high value products.

Previous studies demonstrated that some bacteria are capable of transferring electrons derived from respiration to solid extracellular electron-accepting materials. Termed extracellular electron transfer, this ability allows bacterial biofilms to use solid conductive substrata as a sole electron sink. The electrochemically active biofilms have numerous implications in biogeochemistry and microbial ecology, and they have also been utilized in practical applications such as in marine-based microbial fuel cells to power oceanographic sensors and monitoring devices to extract energy and enhance chemical oxygen demand removal during wastewater treatment; to act as biosensors for nitrate/nitrite, glucose, and other chemicals; to desalinate water while simultaneously generating electricity; and for producing hydrogen gas via microbial-driven electrolysis. Understanding EET mechanisms will allow us to exploit and engineer this process for many beneficial applications, several of which are of interest to DOE, such as heavy metal and radionuclide bioremediation and electrochemically stimulated synthetic biology.



Optimized EC-NMR probe with disposable (3D-printed) biofilm reactors. Top left: Simulations of the EC-NMR probe to optimize the magnetic field generated by the radiofrequency coil. Bottom left: The fully constructed EC-NMR reactor. Right: The new probe coil offers an induced magnetic field that is 4× stronger than previous coils, enabling higher resolution and lower detection limits.

It not known how microenvironments inside electrochemically active biofilms affect electron transfer mechanisms or if it is possible for a biofilm to use multiple electron transfer strategies simultaneously. The goal of this proposed research to address the role of microenvironments on electron transfer at a fundamental level, ultimately leading to enhanced applications of electricity producing biofilms. Combining nuclear magnetic resonance (NMR) with electrochemical techniques, we can explore electricity-producing microorganisms non-invasively at the microscale. To accomplish these challenges, a novel NMR biofilm reactor will be created to act as a potentiostat-controlled 3-electrode electrochemical cell, allowing for simultaneous electrochemical and NMR techniques.

To reach our goal, an electrochemical-nuclear magnetic resonance biofilm reactor was designed and optimized to allow for simultaneous NMR and electrochemical techniques at high resolution and quick scan rates. Three out of the four subgoals for optimization have been completed, including creating a reactor with built-in reference and counter electrodes, improving the radio frequency coil, and incorporating electromagnetically shielded wire to minimize noise on the potentiostat lines. We also generated a new reactor design using 3D printing technology. Compared to the previous design, the new reactor eliminates the optical window, has a cross-sectional area ~60% smaller (~40% smaller perimeter), and has a 50% less internal volume. Finally, a custom Faraday cage was built to enclose the potentiostats used for electrochemical techniques, improving the signal to noise ratio. These improvements are the focus of an article under review and was the topic of an international biofilm reactors conference presentation awarded the “best in section.” Two papers based on our preliminary work with the EMSL simultaneous EC-NMR reactor setup were published in *Energy and Environmental Science*; two additional papers are in peer review.

PN13035/2516

As proposed, one aim of this research was to develop and verify a mathematical model to determine the role of microenvironments on kinetic calculations and to validate current theories regarding electron transfer in biofilms. Toward this goal, we worked with Chinook supercomputer support to install the Comsol modeling software package. Several draft models were tested on the cluster. The simulations were run to completion, and the model scaled reasonably well with increasing nodes. Modeling work will begin in earnest after obtaining the first experimental data set to guide parameter values and biofilm geometry.

Over the next year, we will begin using the EC-NMR setup to measure microenvironment gradients in Geobacter biofilms. This will include significant time on the 500-MHz Bruker and some time on the 600-MHz Varian and use of the FEI Helios SEM. Modeling work will also continue on Chinook. Electrochemically active biofilm research is receiving increased attention in the literature, and our novel EC-NMR system has the unique capability to address several questions in this field. After completing our current aims, in FY 2014 we plan to apply the EC-NMR system beyond what was originally proposed, including coupling an electrochemical quarts microbalance to the reactor to allow for detection of single cell attachment to the working electrode. Further, we will test simultaneous electrochemical impedance spectroscopy and NMR to determine individual rate limiting steps for extracellular electron transfer. All experimental data will be enhanced with theoretical computations.

The Statistical Mechanics of Complex Process in Bulk and Interfacial Environments

Marcel D. Baer

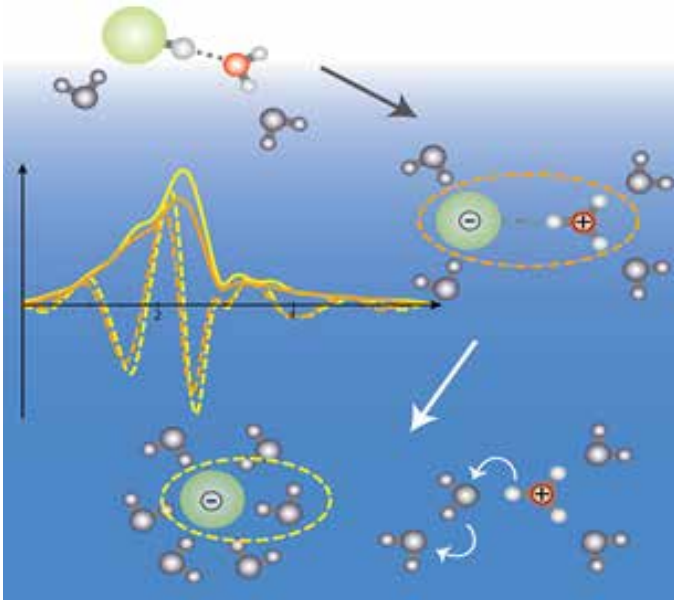
This project is developing new research capabilities for simulation of reactivity and structure in complex heterogeneous and homogeneous environments.

A significant issue in basic energy sciences is understanding and characterizing the novel chemistry that takes place at interfaces. This effort requires the use of molecular simulation with interaction potentials that contain charge transfer, polarization, and the ability to make and break chemical bonds (i.e., chemistry). The development of novel simulation protocols to probe differences between bulk and interfacial environments is a crucial component for making quantitative predictions about reaction thermodynamics, selectivity, and activity. The nature of ion solvation at interfaces can change how we conduct basic energy research; however, fundamentals of the propensity of an ion to be present at an interface and the important interactions are not completely understood.

In recent years, theoretical and experimental studies have presented a picture of the aqueous interface, wherein hard and/or multiply charged ions are excluded from the interface but large polarizable anions show interfacial enhancement relative to the bulk. Thus, a fundamental understanding of ions in solution is an important step toward controlling matter. Surface sensitive spectroscopy and current and next-generation X-ray absorption experiments require high quality simulations in order to interpret observation. There is thus a unique opportunity to use and develop molecular simulation methodology based in quantum mechanics to advance our understanding of these important systems and provide sorely needed data to aid in the interpretation of experiments. To this end, the air-water interface serves as a model interface where much of the important phenomena regarding ions and proton transfer are an active area of research.

During FY 2011, a joint experimental and theoretical study revealed remarkable differences in the solvation structure around molecular HNO_3 under bulk and interfacial solvation, suggesting that hydrogen bonds between HNO_3 and water molecules at the solution surface stabilize the molecular form at low concentration. In FY 2012, the first solvation shell structure about the polyoxyanion, iodate (IO_3^-), under bulk conditions as measured by extended X-ray absorption spectroscopy (EXAFS), was investigated. The region of IO_3^- behaves like cation, as the iodine atom has a formally positive charge while the other region behaves like an anion wherein the iodate oxygen atoms have formally negative charge. We further expanded the concept of the first solvation shell in determining/predicting interfacial propensity.

During FY 2013, we investigated three ions (iodide, iodate, and thiocyante) that have similar polarizabilities yet significantly different surface propensities. Their hydration structure under both bulk and interfacial conditions suggest that it is indeed the characterization of the hydration structure, rather than the underlying polarizablility, which is a strong indicator of surface propensity. These results were presented in two publications, *Journal of Physical Chemistry Letters and Faraday Discussions*.



Schematic representations of the solvation and dissociation of HCl.

Through the use of EXAFS measurements in conjunction with state-of-the-art density functional theory simulations, we can obtain an unprecedented view into the molecular structure of medium to high concentrated electrolytes. Most especially, we find a $Cl—H_3O^+$ contact ion pair that persists throughout the concentration range studied, which depicts a significantly different structure from moieties studied in micro-solvated hydrochloric acid clusters. Similar structures are also found as stable structures at the air-water interface, highlighting that the interface resembles structures found in high concentrated bulk solution as already seen for the initial HNO_3 studies.

We will continue to pursue a better understanding of complex chemistry at interfaces using molecular simulation. This includes but is not limited to the continued development of interaction potentials and sampling methods toward characterizing experimental observables in concentrated acid and electrolyte solutions in bulk and interfacial environments.

PN11075/2397

Upgrading of Hydroprocessed Pyrolysis Bio-Oils by Selective Ring Opening of the Aromatics and Cycloparaffins to Paraffins for Diesel and Jet Markets

Vanessa Dagle

We are developing a two-step process to have better control and maximize the yield of high molecular weight paraffins products. This effort will lead to the efficient upgrading of hydroprocessed pyrolysis bio-oils and be applicable to upgrading any bio-oil or fuel containing cyclic compounds through ring opening catalysis.

Bio-diesel represents less than 1% of the total U.S. diesel market, and all aviation fuels are currently derived from fossil fuel sources. Fast pyrolysis and subsequent catalytic hydroprocessing is one of the most viable technologies to produce liquid biofuels. However, bio-fuels obtained from direct liquefaction routes have limited use due to the high cyclic nature of the carbon chains produced. Efficient upgrading of these aromatics and cycloparaffins into (iso)-paraffinic compounds which are highly valuable for diesel and jet markets will support the U.S. economy. Heavy transportation consuming 70 billion gal/yr of diesel + jet fuels is a future probable target industry for biomass-derived diesel and jet fuels.

Hydrotreated bio-oils contain up to 90% of cyclic hydrocarbons that could be converted into high molecular weight paraffins via ring-opening technology. This process is conventionally conducted in a single step over a bifunctional catalyst containing a metal and an acid function. The efficiency is limited for C_6 rings due to the production of a large amount of cracked products. When metal and acid are mixed together, multiple reactions take place simultaneously leading to undesired products. Our hypothesis is that to solve this issue, acid and metal functions must be separated into two catalysts loaded in two separate reactors. Hence, we are developing a two-step process consisting of a ring contraction catalyst reactor for selective production of C_5-ring compounds and a hydrogenolysis catalyst reactor active for C_5 ring-opening to maximize the yield of high molecular weight paraffins.

First, we studied process conditions and screened catalysts for the 2-step ring-opening of methylcyclohexane (MCH), which is one of the major compounds of hydrotreated pyrolysis bio-oils. For the "ring-contraction" step, parametric tests have shown that the selectivity toward the desired alkylcyclopentanes (C_5-ring compounds) decreases with the increase of the conversion. Thus, lower temperature (250–320°C) and higher space velocity (GHSV=10,000-23,000 h^{-1}) are preferred. The pressure (15–45 bars) has little effect on catalytic performance. At moderate temperature and pressure, a remarkably high yield of alkylcyclopentanes (65%) can be obtained.

We also demonstrated that the nature of zeolite support significantly affects catalytic performance. Indeed, the 1% Pt/H-ZSM22 catalyst was selective to the desired alkylcyclopentanes compared with the 1% Pt/H-ZSM5 catalyst, especially at high conversion. For example, for conversion equal to 70%, the alkylcyclopentanes selectivity was high and equal to 95% for 1%Pt/HZSM22 and only 60% for 1%Pt/HZSM5. The undesired cracking reactions leading to the formation of unwanted C_1-C_4 products are more facile over the 1%Pt/HZSM5 due to higher concentration of acid sites. Note that for the 1%Pt/HZSM22 catalyst, exceptionally high selectivity (> 90%) toward desired products was obtained over a wide range of conversions studied (0% conversion ≤ 70%). Because the 1% Pt/HZSM22 catalyst appeared promising, we investigated the effect of metal loading (0–3% Pt) for the supported HZSM22. The results have shown that Pt is necessary to produce olefins that are readily transformed into ionic intermediates leading to C_5 ring formation. However, Pt dispersion does not affect significantly the activity and only small quantities of Pt (≤ 1%) are required.

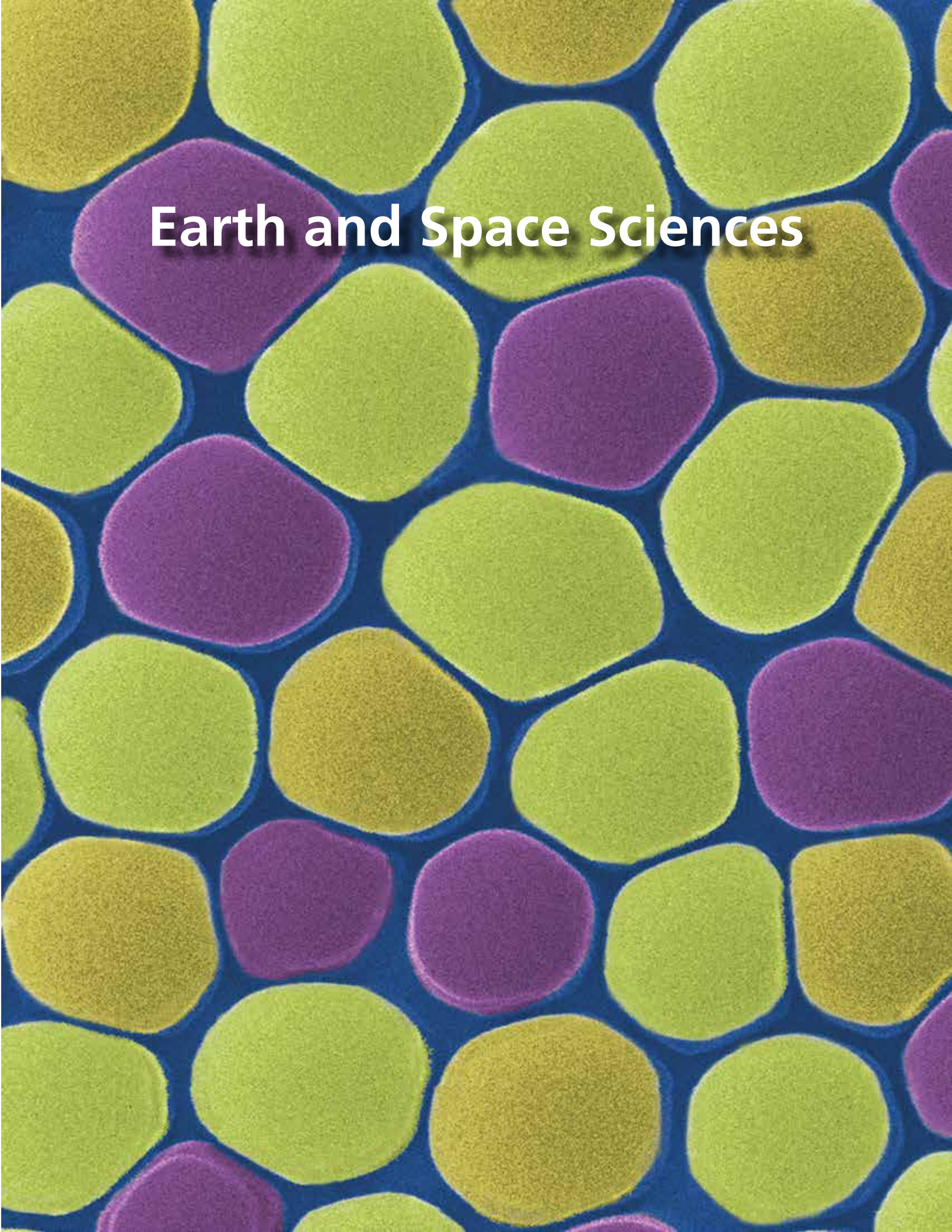
For the hydrogenolysis step, we examined the effect of the process conditions using a 1% Ir/Al_2O_3, as this catalyst is known to be efficient for C_5 ring-opening/hydrogenolysis. The products and unconverted MCH of the ring-contraction step were used as feedstock. The results have shown that the selectivity toward the n and iso C7 paraffins is higher at lower temperature (300°C) and pressure (15–20 bars). Cracking reactions were indeed favored when temperature

PN13054/2535

and pressure increased from 300 to 360°C and 15 to 30 bars, respectively. For T=320°C and P=21 bars, up to 67% selectivity of n and iso C_7 paraffins was obtained.

Next, we compared the efficiency of this novel two-step process with the one-step. For the one-step, we studied two configurations: a "mixed bed" configuration where the 1%Pt/ZSM22 catalyst was mixed with the 1% Ir/Al_2O_3 catalyst and a "cigarette bed" configuration with the 1%Pt/ZSM22 catalyst loaded on top of the 1% Ir/Al_2O_3 catalyst. For comparison purposes, the catalytic performance of the "mixed bed" and the "cigarette-bed" was evaluated at 320°C and 21 bars. For conversion=30%, the selectivity toward n & iso C_7 paraffins was equal to 26% for the mixed bed configuration, 35% for the cigarette bed configuration and 41% for the two-step process. These results demonstrate that is it preferred to separate the ring-contraction catalyst and the hydrogenolysis catalyst to avoid multiple reactions leading to carbon loss to take place simultaneously. Another important finding is that optimal n and iso C_7 paraffins selectivity is obtained when the two catalytic functions are loaded into two different reactors rather than on top of each other in one reactor. For the two-step process, the hydrogenolysis step can be conducted at higher throughput, which limits cracking and thus favors formation of the desired products.

We successfully demonstrated proof-of-concept for a novel two-step ring-opening process minimizing cracking while operating at mild conditions. Upgrading of hydroprocessed pyrolysis bio-oils to produce quality fuels can be achieved more efficiently with this two-step process compared with a more traditional one-step process.


Earth and Space Sciences

A New Non-Aqueous Fracturing Technology for Eliminating Water Requirements for Fracturing Liquid-Sensitive and Non-Liquid-Sensitive Reservoirs

Kirk J. Cantrell

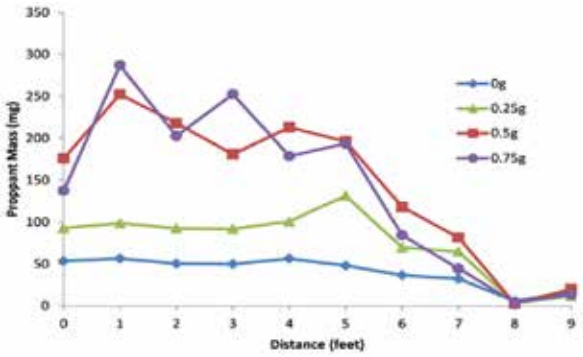
We are developing and demonstrating a high viscosity CO_2/polymer formulation to transport proppant more effectively into source rock fractures to enhance the recovery of gas and oil from unconventional (shale) reservoirs.

Current hydraulic technologies require large quantities of water to conduct fracturing operations, which produce large quantities of flowback water that must be properly disposed or treated. In addition, aqueous-based fracturing fluids can cause formation damage and loss of productivity in shale source rock formations as a result of clay mineral swelling and pore plugging. An alternative to aqueous based fracturing fluids is the CO_2/Sand Dry –Frac technology. Major benefits to the CO_2/Sand Dry –Frac approach over aqueous fracturing fluids are the potential for formation damage in the shale source rock is virtually eliminated along with the water requirements, and the CO_2 flows back from the fractures to the wellbore, leaving them open for hydrocarbon transport. A shortcoming of the existing technology is that unaltered liquid CO_2 makes a poor fracking fluid due to its low viscosity which, as a result, lowers the CO_2 for suspending and transporting proppant through fractures during injection.

In aqueous systems, polymers are added to increase viscosity and enhance proppant emplacement. While CO_2 is good for most nonpolar and some polar molecules of low molar mass, it is a poor solvent for most high molar mass polymers under mild conditions. An alternative approach is using polymer-CO_2 suspensions to increase viscosity. Polymer particles can be stabilized in supercritical CO_2 with surfactants that have CO_2-philic and anchor block(s) that adsorbs to the polymer.

Polymethacrylate latex dispersions in CO_2 were successfully synthesized and tested at a variety of concentrations in CO_2. Viscosity measurements completed on these formulations indicated that no significant increase in viscosity resulted from these formulations. As a result of this setback, we developed a new strategy for improving proppant emplacement with liquid CO_2. In this new approach, polymeric fibers are used to increase greatly the transport and emplacement of proppant. The use of synthetic fibers has been shown to provide excellent proppant transport in low viscosity aqueous systems (slickwater). We demonstrated that this approach can also increase the efficiency of proppant transport using liquid CO_2 as the primary fluid during hydraulic fracturing operations.

Experimental work with CO_2/proppant/fiber formulations was conducted using a specially designed apparatus that consisted of a Parr pressure reactor in which proppant, liquid CO_2, and fibers are mixed into a slurry; a stainless steel tube to simulate a fracture; a CO_2 reservoir to limit pressure drop after slurry is released from the Paar; an ISCO pump to transport CO_2 into the reservoir and reactor; and various valves to control CO_2 movement and slurry. Each test was conducted by allowing the slurry to enter the tubing by rapidly opening a valve at the entrance to simulate flow into a fracture. Ultra-light weight proppants (ULWPs) were used with a density of 1.04 g/cm^3 and readily suspended in liquid CO_2. The fibers are polylactic acid (PLA), which are non-toxic and readily degradable in the subsurface.



Proppant emplacement as a function of quantity of PLA fiber added.

Results collected with this system indicate that the addition of PLA fiber to the CO_2 proppant slurry dramatically increases transport into the tubing. The image shows the results using various concentrations of PLA fibers. With no fiber added, ~50 mg/ft of ULWP was emplaced within the tube, fairly evenly distributed through the first 5 ft. Addition of PLA fibers significantly increased ULWP emplacement. The addition 0.5 gm and 0.75 gm of PLA fiber increased the emplacement of ULWP by a factor of four within the first 5 ft of tubing. Higher concentrations of fibers (1.0 and 1.5 gm) did not result in any further increase in proppant emplacement. Overall, polymetha-crylate latex dispersions in CO_2 were not effective for increasing the viscosity of the liquid CO_2 formulations.

PN13042/2523

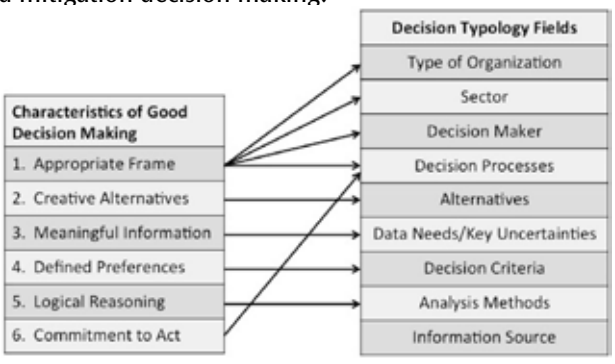
Decision Support Research for Integrated Regional Earth System Modeling

Richard H. Moss

This project is researching what uncertainties matter most in different decisions, what information users want regional models to provide, and how information can be presented in the context of uncertainty.

The output of regional integrated models such as the Platform for Regional Integrated Modeling and Analysis (PRIMA) is relevant to a growing set of management and investment decisions. Despite advances in model development itself and considerable interest from decision makers, incorporation of information on potential climate exposures, impacts, and the effectiveness of response options into decision making remains largely an aspiration. A "usability gap" exists between information provided by the science community and what decision makers appear to need, stemming from differences in perspective between producers and users as well as the inherent difficulties of applying climate information in many decision making processes.

This focus area targeted understanding stakeholder needs to validate planned model capabilities, identify desirable new capabilities, and guide model demonstration and uncertainty characterization, all in support of ensuring that PRIMA outputs and insights are usable and used. Through interviews with potential users in the energy, water resources, and agriculture sectors as well as review of published and grey literature on adaptation and mitigation decision making in the region, the project identified a number of stakeholder information needs. In addition, drawing on tenets of effective decision making in decision science, we have analyzed these data using a typology of decision criteria and decision making processes. The results have been used to plan a pilot decision support process and contribute to literature on scenarios in adaptation and mitigation decision making.



Decision support tools and information must be flexible to meet the needs of different decision frames and provide a range of capabilities for modeling and presenting output results.

In one dimension of our research, we focused on state-level decision making processes. Based on this research, we identified 8 of the 14 states in the region that have active processes addressing mitigation and/or adaptation. At least a dozen different decision making and decision support processes are proposed or in use, with each state employing multiple processes. Processes range from legislation or executive orders to advisory groups, a wide range of models, and technical working groups. There is strong interest in identifying co-benefits of adaptation and mitigation. These results confirm that decision support tools and information must have flexibility to meet the needs of different decision frames and provide a range of capabilities for modeling and presenting output results. In contrast to the breadth of decision process findings, the typology identified only three types of analysis methods in use: cost-benefit, independent factor, and scenario analysis. Methods to address uncertainty and not widely applied, in spite of the degree of uncertainty in all aspects of the phenomena and effects that PRIMA addresses. Regarding decision criteria, the typology reveals a wide range in use from standard quantitative metrics (economic benefits) to more qualitative factors (the desire for "no surprises").

Building on our results, we are seeking funding to continue to examine the patterns of decision making and information needs in municipalities, private sector firms, and non-governmental organizations that are concerned with regional climate change and its implications. We are including research on processes in which climate change considerations are being "mainstreamed" into ongoing investment and management decisions to identify the particular needs of this specific category of users.

This is the final year of research focused on stakeholder information needs – the related project on uncertainty characterization and communication concluded at the end of FY 2012. The results of both projects are being applied in planning extensions of the platform and conducting the demonstration experiments in FY 2014. The project has established the importance of incorporating a clear understanding of decision context, stakeholder information needs, decision processes, and analysis methods in developing and applying modeling platforms intended to provide insights to decision makers. It has also extended stakeholder research in several respects, including developing the typology approach to provide structure and consistency in recording and analyzing what can often be very "messy" data resulting from a literature review or from stakeholder consultations.

PN12019/2420

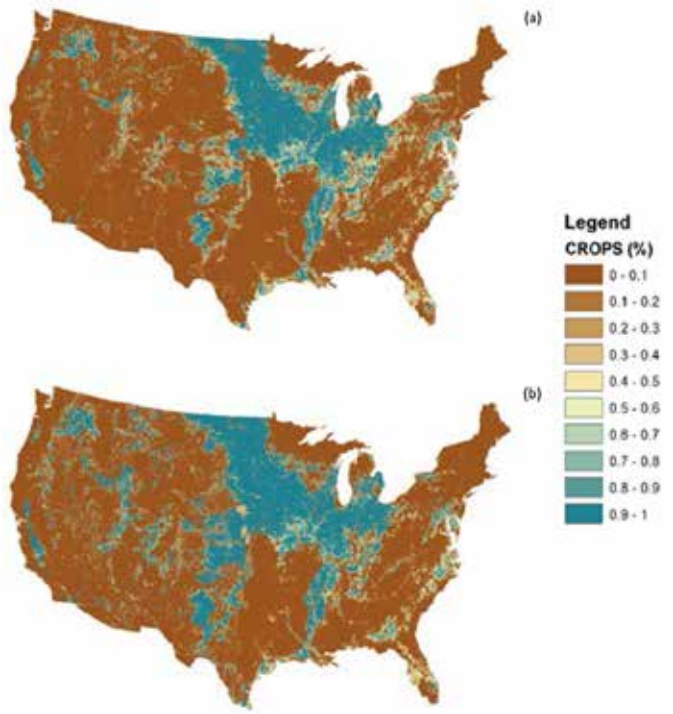
Developing a Dynamic LULCC Model for Spatially-Explicit Future Realizations of Projected Land Use Under an Integrated Regional Earth System Modeling Framework

Tristram O. West

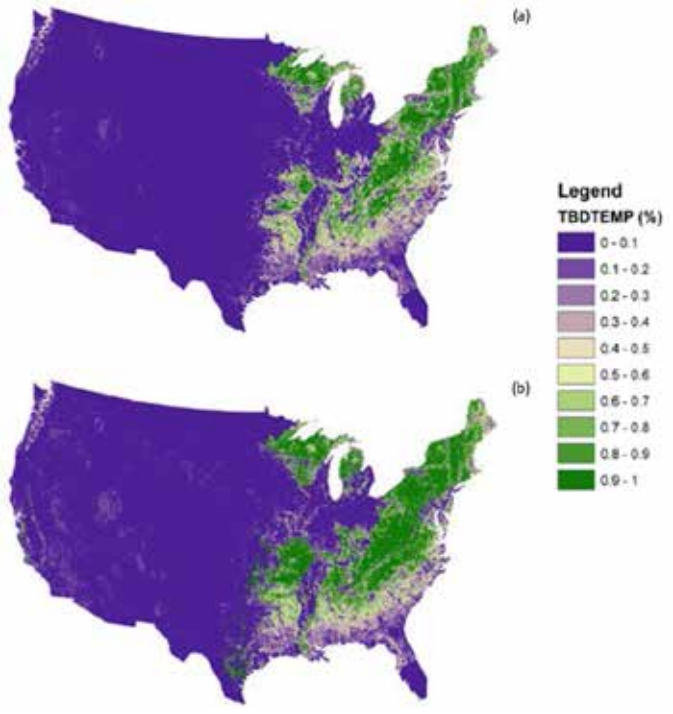
We are developing a geospatial distribution model that distributes projected land cover over the landscape based on current land distribution estimated by satellite data. The resulting data can be used in earth system models to analyze climate feedbacks, greenhouse gas emissions, and numerous other environmental impacts.

Changes in land cover and land use are often projected within geopolitical boundaries using integrated assessment models. Because these projections are not geospatially distributed over the landscape, it is difficult to analyze how future changes in land cover will impacts environmental variables such as water erosion, nutrient runoff, greenhouse gas emissions, climate feedbacks, and water and carbon cycles. Similar work has been conducted in the past but was insufficient in scope or in use of best available data and models. Our research was specifically intended to downscale and spatially distribute GCAM integrated assessment model results for RCP scenarios to 0.05 degree resolution for use in the Community Land Model (CLM) that is operated by PNNL colleagues. This enables direct feedback between projected changes in land cover from the GCAM model to possible climate feedbacks in CLM. Additionally, the resulting product may also be used by other Earth system and biogeochemical modelers in the broader research arena. The research fills a key data gap between future land cover projections and climate modeling. Our intent is to fill this data gap, publish papers on the method and how it alters current climate models, and have the dataset and papers used and cited by the research community.

The method was developed and used with a 38-region area in the midwestern United States because it had an increased spatial resolution of GCAM output over what is normally generated at the more coarse AgroEcologocical Zone (AEZ) level. This output corresponded with a related



GCAM spatial distribution of U.S. crops in (a) 2005 and (b) 2090 for the RCP4.5 scenario.



GCAM spatial distribution of U.S. temperate broadleaf deciduous forests in (a) 2005 and (b) 2090 for the RCP4.5 scenario.

PN13002/2483

project that generated the ability of GCAM to develop output for the midwestern region at higher spatial resolution. It was later determined that this region did not fully encompass the target region in CLM. The method was therefore expanded to the entire United States area using GCAM output for AEZs. During expansion, we improved the computer code and moved from R to Python code. We also included climate region rules and rules for changing area between land classes based on geospatial coordinates. The results are promising and coincide with our initial project objectives.

This project was focused more on capacity building than on research analysis. Therefore, the most important findings will likely be found when our data product is used within CLM. It was always the intent to conduct CLM simulations using these data in a follow-on project. The most important immediate finding from this research is that the method we developed to distribute integrated assessment model output spatially commensurate with satellite-derived MODIS land cover works well and will likely prove a useful data commodity within PNNL, DOE, and the global research community as a whole.

Our project is not continuing in FY 2014. However, future work should include a 50-state version of the agricultural land use (AgLU) module within GCAM. This will help constrain where land cover and land use change occurs. A 50-state version of the AgLu module should be spatially distributed using the Cropland Data Layer at 30m resolution, as opposed to the 500m MODIS data that was used in this project. In all cases, the data is scaled up to 0.05 degree resolution for climate modelers. However, the climate modeling community appears to be rapidly moving to 5-minute resolution global coverage, which is the equivalent to 0.08 degree resolution.

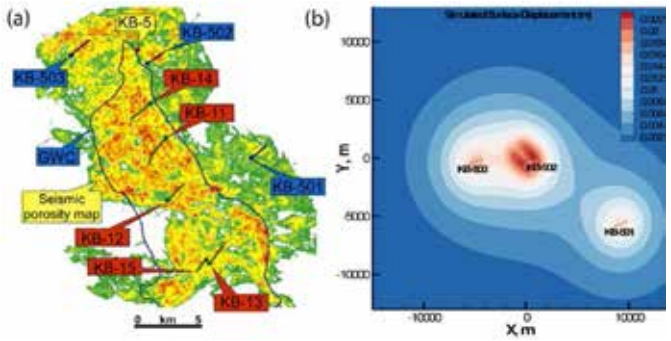
Future work should include finer resolution GCAM output and baseline land cover data for the United States as well as potential changes in final product resolution. These issues would complement existing global spatial distribution of land cover using MODIS satellite data and GCAM output at the AEZ level but would improve resolution in the United States for decision-making purposes.

Development of Coupled Flow, Thermal and Geomechanical Capability for Carbon Sequestration

Yilin Fang

We are developing critically needed geomechanical capabilities to address key research issues in CO_2 geological sequestration. This project will enhance PNNL capabilities in large-scale subsurface simulation.

Geologic sequestration of CO_2 is an attractive option for reducing greenhouse gas emissions without adversely influencing energy use. However, high pressure during the injection phase can result in large displacements of pore fluids and large stress changes on natural fractures and faults. A major risk associated with CO_2 sequestration is potential CO_2 leakage through the cap rock and overburden, which can lead to CO_2 release into shallow potable aquifers or the atmosphere. Developed at PNNL, subsurface nonisothermal, multiphase flow, and reactive transport code (STOMP) does not have the capability of geomechanical simulation, which is critical to evaluate the overall suitability of the geological reservoir for safe CO_2 injection and CO_2's long-term subsurface containment.



(a) Field layout and the three injection well locations in the Krechba gas field at In Salah (Algeria); (b) Simulated surface vertical displacement after 2 years of CO_2 injection. Solid black lines are the injection well locations, dotted rectangle area is the location of the fault.

We are developing a simulation tool that couples nonisothermal fluid flow and geomechanical deformation processes to address key research issues in CO_2 geological sequestration. A particular interest is to develop a large-scale subsurface simulation that will eventually be used to evaluate the long-term sealing ability of caprock and fracture networks above target CO_2 storage reservoirs during CO_2 injection and storage. Geomechanical risks such as opening of pre-existing fractures, rock fracturing, weakening of the strength of faults, and induced seismicity increase, along with pressure increases in the reservoir during CO_2 injection. One methodology for our approach to reducing greenhouse gas emitted by coal-fired power plants is to capture the gaseous CO_2 effluent and inject it into deep geologic reservoirs as supercritical (sc) CO_2.

Initial analyses rely on predicting the evolution of effective stresses in rocks and faults during CO_2 injection. Potentially hundreds of $scCO_2$ injection wells will operate in the Illinois Basin, and analysis will become necessary to determine the collective impact of these long-term injections and the aggregation of $scCO_2$ beneath the caprock on structural integrity.

During FY 2012, we tested the geomechanics code, Rigid-body Interface Element Method (RIEM), which was incorporated into the existing parallel subsurface nonisothermal, multiphase flow and reactive transport code (eSTOMP-CO_2) for linear and nonlinear problems. We explored the coupling between a discrete element method (DEM) and a continuum scale model to study a fracture propagation problem within a faulted domain during CO_2 injection.

For FY 2013, we used the RIEM coupled with eSTOMP-CO_2 to simulate the double-lobe uplift pattern observed in the Krechba gas field at In Salah (Algeria), a site that demonstrated the success of a CO_2 sequestration effort into a deep saline formation. We used our models to study the impacts of caprock and reservoir properties and the injection rate fluctuation on the geomechanical response for CO_2 sequestration uncertainty quantification. We implemented fracture and plastic model in RIEM, and the model is under test. We submitted several manuscripts documenting our work to professional journals.

Our work incorporated a geomechanical model from the linear elastic theory into eSTOMP-CO_2 to provide coupled geomechanical and multiphase flow capabilities for the suite of simulation tools. Enhanced DEM at the mesoscale examined the structure-property relationship within and close to individual fractures, captured the relevant dominating micro-mechanisms that could not be accurately treated at larger scales, and provided reliable and robust submodels for model integration. Additional methods will be developed to couple models at the mesoscale and continuum scale.

PN11025/2347

Enhanced Sediment Geochronology Achieved Using Ultra-Low Background Materials and Ultra-Sensitive Detection Capability

Gary A. Gill

New and enhanced capabilities developed by this project will have a significant impact on environmental geochronology and other low-background radiation measurement applications, significantly advancing PNNL as a national leader in new geochronology capability as well ultra-sensitive detection of radioactive signals.

A gap exists in age dating of environmental systems (e.g., sediment, trees, groundwater) between approximately 100 and 1,000 years, limiting our ability to understand ecosystem changes as a function of climate and human drivers. Thus, the objectives of this project are to:

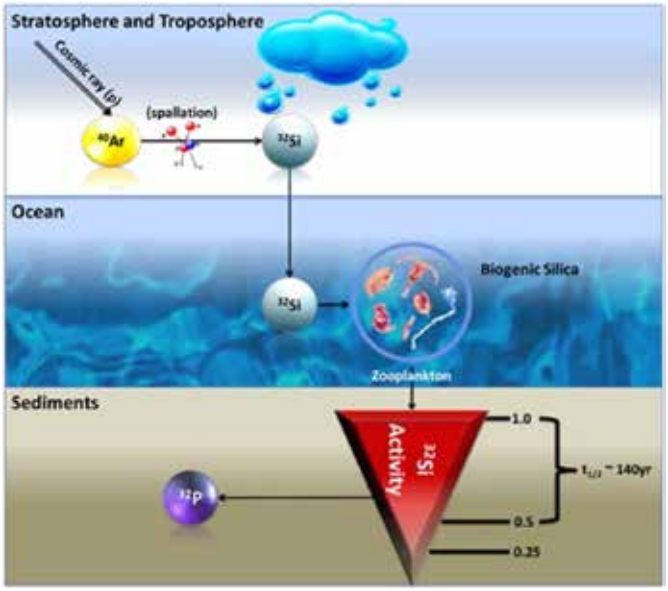
- produce a new sediment geochronology tool based on highly sensitive radiometric dating using 32-Si (half-life ~ 140 yrs)
- develop the world's lowest-background beta detectors utilizing new low-background plastics to enable the 32-Si age dating method
- demonstrate 32-Si age dating with Puget Sound sediment cores in the 100-1000 year age bracket
- take a new, high-precision measurement of the 32-Si half-life.

The project began in FY 2013 and has been organized into three components: a geochemical task involving isolation of 32-Si (32-P) from sediments for radiometric detection; development of a highly sensitive and low-background beta detector; and identification of low background plastics for use with detectors.

Sediment Geochronology Tool. This project is producing a new sediment geochronology tool based on 32-Si (half-life of ~140 yrs) that will be useful in 100–1,000 yrs, a time period not currently covered adequately by other radioisotopic geochronology methods. The tool will be achieved through the development of new ultra-sensitive beta detectors and identification and manufacturing of detector component materials with low background signals. This capability will be demonstrated by conducting a geochronology (age dating) of sediment cores collected from Puget Sound through comparison to radiochemical geochronology using 210-Pb.

Sediment Geochemistry. A major effort in FY 2013 was to develop and optimize procedures for quantitatively isolating highly pure biogenic silica (containing 32-Si) from bulk sediments. An extraction and purification scheme was developed and continues to undergo optimization. The scheme involves solubilization of biogenic silica from sediments with strong base, with clean-up approaches using cation and anion exchange and highly selective phosphomolybdenum precipitation.

The major challenge has been to develop procedures that are quantitative with high yields while working with sufficient sediment masses (100–1000 g) needed to isolate sufficient 32-Si for radiometric detection. The next steps for FY 2014 will be to develop and optimize the 32-P "milking" process with the ultimate goal of providing a very small mass sample of 32-P for radiometric detection.



32-Si is produced in the troposphere by cosmic radiation, removed from the atmosphere in wet and dry deposition, incorporated into plankton, and removed to sediments (biogenic silica). 32-Si activity decreases with depth in sediments, providing a means to age date layers. As the daughter of 32-Si, 32-P is quantified, as it has a higher activity and has secular equilibrium with 32-Si.

Ultra-sensitive BIDS Detector Development. Based on a literature review and technology down-select, the design chosen for the development of a prototype low-background counter was that of a variable-length pressurizeable single wire counter that can be paired with an identical counter to provide 4 Pi solid angle coverage

PN13018/2499

of the source. The detector design was completed at the end of April, and drawings for fabrication were sent to the shop in early May. Detector parts were fabricated and cleaned over the summer. Presently, all detector components and custom assembly tools have been received from the shop, cleaned, and are ready for assembly.

The small parts (wiring, electrical connectors, Teflon seals, and commercially available assembly tools) have been identified, purchased, and cleaned. A bench-top laminar flow hood was identified, purchased, and received to provide a clean above-ground space for detector testing. In late Q4 and early FY 2014 Q1, the detectors will be assembled and tested. A single counter will be operated/characterized in the existing shield in 3425/B124, and the double counter will be operated/characterized above ground while we pursue new shielding configurations for B124.

Low Background Materials. The major thrust of the work conducted in FY 2013 revolves around developing and refining of methods to assay potential materials for background contaminants (uranium [U] and thorium [Th]). The analytical approach is based on high temperature ashing of plastics in crucibles and analysis by inductively coupled plasma mass spectrometry.

	Polymer		µBq ^{232}Th/kg	µBq ^{238}U/kg
Vendor A	PEEK	Sample A	67	1689
		Sample B	67	4878
	UHMWPE	Sample A	22	5852
		Sample B	31	10489
		Sample C	61	6137
Vendor B	UHMWPE	Sample A	68	12266
		Sample B	163	19181
		Sample C	153	29093
	HDPE	Sample A	29	354
Detection Limit (based on 1 gram sample)			0.4	1.2

Assay of selected polymer materials for U and Th contamination.

Fluorinated materials have proven very difficult to assay with standard crucibles yielding high backgrounds or damage to the container after use which resulted in high backgrounds. A variety of crucibles were subsequently investigated including quartz crucibles (three vendors), morlite, alumina (best grade), Zr, Ni, vitreous glassy carbon, sapphire, pyrolitic graphite with CVD (chemical vapor deposition) diamond coating, and low background copper. In addition, we switched from a muffle to a tube furnace, which has a controlled nitrogen cover gas, to reduce potential contaminant transfer during ashing. All these new materials exhibited etching particularly with the PEKK and fluorinated polymers. Only the copper crucibles did not show elevated background resulting from the newly exposed surface area.

An alternative approach to assaying plastics was initiated in Q4 FY 2013 based on similar work being conducted for a related project. The method uses isotopic dilution analysis (IDA) of solid samples using (laser ablation-ICP/MS) to preconcentrate trace analytes of interest in plastic materials for subsequent analysis by laser ablation mass spectrometry (LAMS) over a 10-sec analysis window. This approach should eliminate most if not all of the process blank contamination levels currently measured by aqueous procedures and provide for analyte pre-concentration factors of > 500.

Left: The X2 detector mount plate, replaced with the ultra-pure copper implant. Right: The machined 1mm diameter ultra-pure copper implant.

Under this portion of the project, we determined a 1.3 mg/hr laser ablation rate for plastic samples, measured the LA-ICP/MS ion efficiency (0.1%), built a micro implant device, and evaluated the copper implant for blank levels of U and Th.

Enhancing EMSL Mass Spectrometry Capabilities for Characterization of Soil Organic Matter

Nancy J. Hess

We are developing experimental capabilities to allow molecular characterization of organic molecules in soil, leading to harnessing chemical reactions that impact C cycle dynamics in response to climate change, land use and agricultural practices, and extreme weather events.

Our research will develop high throughput, high resolution mass spectrometry (MS) capabilities for the characterization of organic matter contained in mineral- and organic-rich soils. In standard practice, analytical capabilities to characterize soil C provide only broad description of classes of organic components typically based on their chemical solubility in extractants under laboratory conditions or total elemental composition from pyrolysis. Such descriptions are inadequate for development of mechanistic functional models of carbon allocation in soils in terrestrial ecosystems by plant, microbe, and fungal metabolism. Additionally, the lack of chemical and molecular detail makes assessing the vulnerability of soil C to global climate change and its incorporation of potential feedbacks in global climate models equally unobtainable. This technical gap can be addressed by leveraging EMSL's recognized expertise and capabilities in high resolution MS, imaging, and automated data analysis capabilities and develop novel separations and ionization approaches for the characterization of soil C, thereby providing the chemical and structural detail needed to develop mechanistic descriptions of soil carbon flow processes.

In FY 2013, we focused on three issues: addressing the data analysis bottleneck for high-throughput SOM characterization; laser ablation (LA) sampling for solvent-free characterization of intact soils; and development of workflows for LC-FT ICR/MS characterization of SOM extracts. Progress in these areas is described below.

Workflows for data analysis of HR-MS data from SOM. We are creating user friendly data analysis tools for rapid assignment of thousands of peaks in SOM spectra. To this end, an application toolkit is being developed that enables user-defined workflows to align, classify, and identify large amounts of peaks in high-resolution MS data. The unique aspect of our approach is that peaks classification and assignment relies on the high-order mass defect algorithm, which is used in the toolkit for both classification and identification of molecules based on their mass defects. As a classification technique, the algorithm reduces the number of features by clustering peaks with similar first- and second-order mass defects into classes of molecules that vary only by the number of functional groups for the mass defect calculation. This powerful approach efficiently reduces the number of features for identification and enables assignment of elemental formulas to higher *m/z* species that cannot be unambiguously assigned based on the accurate mass alone.

Compared to original macros, we achieved an appreciable improvement in efficiency and speed. For example, a classification process that once took ~ 3 min for 10,000 peaks can be accomplished in ~ 1 sec. Mass alignment and high-order mass defect classification were tested on one of the largest data sets that includes > 90,000 high-resolution mass spectra containing 1,700,000 distinct *m/z* features. Although the data processing speed is a strong function of CPU speed, we found that on an average modern computer, mass alignment and mass defect classification on the large data set could be completed in 10–20 min, demonstrating the toolkit's utility for high-throughput data analysis.

In FY 2014, we will focus on making the second- and third-order mass defect classification expandable to any user-defined bases. Several improvements including error handling, project management features, and increased responsiveness will be added to the user interface.

Laser-ablation sampling for solvent-free characterization of intact soils. LA-aerosol mass spectrometry (AMS) is a state-of-the-art method recently developed at PNNL to allow spatial imaging of molecular and isotopic markers of chemical processes in biological systems at the microscale. This work demonstrated that LA sampling of biological samples preserves isotopic and molecular information as determined by coupling LA with isotope ratio mass spectrometry (IRMS) and AMS, which in itself provides a wealth of molecular information. The preservation of isotopic data during the LA process is a key indicator that the material is largely removed as intact aerosol particles during the laser ablation process. Our goal is to LA-AMS or LA coupled with other MS with higher mass resolution for analysis of intact soil C samples without the use of solvents or extraction protocols.

PN13071/2552

In FY 2013, we ablated a variety of organic molecule targets and subsequently analyzed aerosol particles using FT-ICR MS to test the degree of molecular fragmentation that occurs during the ablation process. The samples included Shewanella bacteria that have been previously well characterized by high-resolution mass spectroscopy in EMSL and elsewhere. Several protein samples, carbonic anhydrase, cytochrome C, and G3PDH were included because their fragmentation is well known under MALDI, which uses laser desorption to remove and ionize sample molecules from a matrix. MALDI results in little or no fragmentation so the comparison of mass spectra from laser-ablated particulates to MALDI spectra are expected to be significant. The polymer sample PEG was also ablated because of the well-known mass spectra for comparison with laser-ablated material. Collection of the high-resolution mass spectrum had not been completed by the time of this report. If little fragmentation is observed, we will analyze the dune sand and soil samples prepared for LC-FT ICR/MS in FY 2014.

Workflows for LC-FT ICR/MS characterization of SOM extracts. A standard polyaromatic hydrocarbon (PAH) mixture with 16 subcomponents was introduced to dune sand matrix. This sample was used to determine the outcomes of three extraction protocols: chloroform/MeOH, DCM/Acetone, and DCM/MeOH. An atmospheric pressure photoionization (APPI) ionization source was tested on the MS with several standards to ensure proper function for PAH analysis. From the results, it was determined that the PAH mixture could be successfully extracted and detected from the dune sand matrix using the protocols developed.

Next, four analyte groups were chosen (PAH, organic acid, peptide, and herbicide mixture) to test four different LCMS methods (positive and negative mode APPI; positive and negative mode ESI). All four mixtures were spiked into dune sand and garden soil matrices. A dune sand only control, garden soil only control, and spiked H_2O control were also analyzed. After extraction using the chloroform/MeOH method, all samples (organic and aqueous phases) were analyzed by LCMS, which showed detection of each megamix component in different extraction layers: the PAH mixture was in the organic layer using the APPI source in positive mode; the organic acid mixture was in the organic layer using APPI and ESI in the negative mode; the herbicide mixture was in the aqueous layer using ESI in the negative mode; and the peptide mixture was in the aqueous layer using ESI in the positive mode. Optimization of current methods is necessary and underway, but the initial set of results shows development in the deep fractionation and coupling with MS ionization methodologies.

Based on these preliminary results, we will continue developing the graduated extraction, deep fractionation, and MS analysis methodologies and integrate NMR structural data on selected LC subfractions in FY 2014.

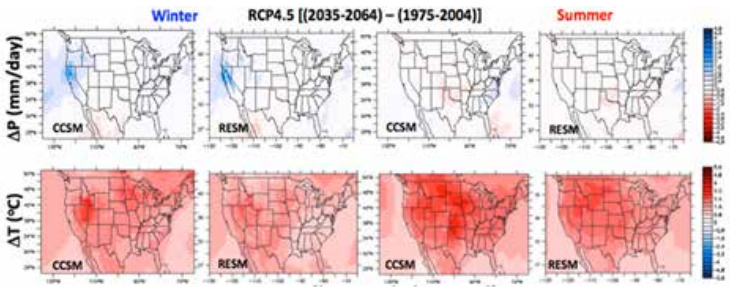
Exploration of Human and Environmental System Interactions due to Renewable Technology Penetration in the Midwest Pilot Region

Jennie S. Rice

We used the novel framework Platform for Regional Integrated Modeling and Analysis (PRIMA) to explore human-earth system interactions under climate change and support adaptation and mitigation decision making.

The PRIMA framework is premised on the hypothesis that climate change, climate policy, and mitigation and adaptation actions taken in response will produce critical interactions between human systems and land, water, and energy resources that can be revealed only through integrated regional-scale modeling. These interactions may be synergistic or antagonistic and as such may indicate the feasibility or infeasibility of certain mitigation or adaptation alternatives. For this project, we demonstrated the PRIMA framework capability through a series of numerical experiments simulating regional climate change, regional mitigation, and adaptation decisions and explored the associated impacts on the land-water-energy nexus, focusing on a 14-state pilot region in the U.S. midwest.

The FY 2013 effort leveraged the research performed in FY 2012 that included developing the scientific integration approaches (algorithms and scripts) between each of PRIMA's component models, testing the coupling of regional climate to each PRIMA model, producing initial results for the impact of climate on state-level building energy demand, and conducting an uncertainty characterization (UC) experiment that demonstrated an efficient fractional factorial design for sensitivity analyses across a wide range of variables influencing the integrated assessment model.



Mid-century climate change signal for global climate model (CCSM) vs. PRIMA's RESM under RCP4.5.

In FY 2013, the project performed numerical experiments involving a range of multi-model couplings:

- Simulation of regional climate in the U.S. from 1975–2100 under Representative Concentration Pathways (RCP) 4.5 using the Regional Earth System Model (RESM), coupling atmospheric, water, and land models
- Simulations coupling climate to hydrology, river routing, and water management models using CASCaDE statistically downscaled climate dataset for A2 and B1 scenarios
- Simulations coupling climate to highly-resolved building energy demand in the pilot region using CASCaDE data
- Simulations and comparison of climate-driven building energy demand from the highly-resolved building energy demand model versus that produced by the coarser integrated assessment model
- Coupling the integrated assessment model to the highly resolved electricity infrastructure modeling.

In addition, FY 2013 numerical experiments extended the UC experiment from FY 2012 to develop a software infrastructure for and propagate uncertainty in the integrated assessment model. As the previous annual report focused on energy-related modeling results, this work summarizes the climate and water management modeling experiment results.

The RESM model simulations for the historical period 1975–2000 demonstrated the capability of the regional scale modeling to capture the geographic variability in extreme temperatures and precipitation as compared to the coarser scale global climate model. The RESM results for the RCP4.5 future scenario show average summer season warming from 1–3°C across the United States for the mid-century (2035–2064) as compared to the historical period. Winter temperature increases are in the range 0.8–2.4°C. On average, the regional model projects slightly cooler temperatures than the global climate model. For precipitation at mid-century, the projected summer season changes range from ~ 0 to –1.0 millimeters per day (mm/day), primarily in the south. In winter, RESM projects increased precipitation, primarily in the western United States, was 2–3 mm/day. Compared to the global climate model, RESM is able to show precipitation changes at higher resolution, and generally predicts dryer summers and wetter winters.

The river routing and water management models were coupled and validated against observations of natural and regulated streamflow and reservoir storage in the Pacific Northwest. Results showed excellent agreement with observations, after which these models were coupled with downscaled water demands from the integrated assessment model to demonstrate the ability to simulate future natural and regulated streamflows.

PN12031/2432

Measuring and Modeling the Climatic Effects of Brown Carbon Atmospheric Aerosols: Developing an Integrated Capability

John E. Shilling

This project will lead to new insights into the conditions under which brown carbon aerosols form; the processes that govern time-evolving chemical, physical, and optical properties; new methods for efficiently representing the processes and properties in climate models; and an evaluation of the impact the particles have on climate.

Aerosols represent one of the largest sources of uncertainty in understanding and simulating future changes in the earth's climate. Aerosols affect the climate system by scattering and absorbing radiation, thereby affecting regional and global energy balances, and indirectly by altering cloud properties and lifetimes. It is commonly assumed that organic aerosols contain only non-absorbing materials and therefore have a predominantly cooling effect on climate. However, increasing evidence suggests aerosol aging may lead to the formation of light-absorbing organic compounds (i.e., brown carbon), which could exert a significant climatic impact currently not captured in climate models.

Unraveling the chemical composition, physical properties, and potential climatic effects of brown carbon aerosols is a multidisciplinary challenge that spans spatial scales ranging from individual molecules to the entire globe. We are developing a fundamental, molecular-level understanding of the chemical and physical processes governing the growth and aging of brown carbon aerosols in the atmosphere; determining how the resulting chemical and physical transformations affect the radiative properties of the species involved; and evaluating the significance of these processes for global and regional climate by developing parameterizations that encapsulate this new understanding and testing these parameterizations in climate models. Incorporating brown carbon formation processes will improve the accuracy of climate models, leading to better predictions of future climate.

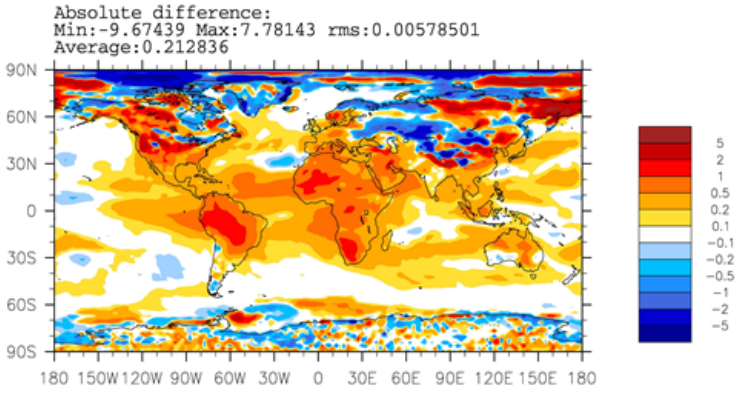
We completed a series of experiments during FY 2013 that investigated the optical properties of organic aerosol produced from the photooxidation of trimethylbenzene as a function of relative humidity, NO_x concentration, and particle aging time. We found that aerosol particles were white (non-absorbing) when produced under low-NO_x conditions and brown (absorbing) under high-NO_x conditions. Under the latter conditions, aerosol absorption increased by a factor of 2–3 when RH increased from < 1% to 30%, further increasing RH to 60% produced no measurable change in the aerosol absorption. Photochemically aging the particles for 24 h in the presence of NO_x increased aerosol absorption by an additional 2–3 factor. Aerosol samples were collected on filters, and the UV-visible spectrum of water-soluble extracts were obtained. Aerosol samples from experiments were also collected for detailed molecular analysis using mass spectrometer facilities at EMSL.

Select SOA samples were analyzed using nanospray desorption electrospray ionization (nano-DESI) coupled to a high-resolution mass spectrometer. Results indicate that SOA obtained under low NO_x conditions is predominately composed of molecules containing C, H, and O_2 but no nitrogen (N_0). By contrast, a significant fraction (> 40%) of molecules in SOA obtained under high NO_x conditions contains one nitrogen atom (N_1) and close to 15% of the assigned peaks correspond to molecules containing two nitrogen atoms (N_2). In addition, the double bond equivalent is higher and the average number of hydrogen atoms is slightly lower in high-NO_x SOA samples while the average number of carbon and oxygen atoms is similar for low-NO_x and high-NO_x samples. These results indicate the incorporation of NO and potentially NO_2 with concomitant loss of one water molecule into the primary oxidation products of TMB.

Several advances in the modeling of brown carbon aerosol have been achieved. On the molecular scale, we have begun performing quantum mechanical calculation of the UV-Vis absorption spectra of several potential brown carbon chemical species. Here we use time-dependent density functional theory (TDDFT) to calculate vertical electronic excitation energies and oscillator strengths (these calculations use a B3LYP/6-311++G** level of theory and basis set, respectively). For the potential brown carbon chemical species to absorb light in the visible region of the spectrum, a large degree of electronic delocalization/conjugation is required. This degree of delocalization occurs when a chain of single-double–single-double– etc. C-C bonds are located in the molecular structure. We are exploring both aliphatic and aromatic compounds with and without carboxyl and nitro groups attached to these species to understand their absorptive properties, based on the preliminary results of laboratory molecular analysis.

PN13073/2554

We performed an exploratory scoping exercise using the CAM5 global climate model to estimate atmospheric and surface forcing that would arise if organic aerosol were much more absorbing than usually assumed by climate models. “Tags” have been implemented that allow us to identify contributions from various precursors of organic aerosol and to assess the forcing associated with aerosol originating from those precursors. We are also exploring the sensitivity of the results to uncertainties in the yields from the precursors. Finally, we made multi-year simulations with a climate model varying the aerosol optical properties and yields and are currently evaluating the model output.



Difference in top of the atmosphere radiative forcing calculated between organic aerosol assumed to be scattering (white) and absorbing (brown) using a modified CAM5 climate model.

In FY 2014, we will extend our laboratory studies to SOA produced from isoprene. Preliminary model results suggest that isoprene SOA has a potential to impact radiative forcing significantly if it absorbs light. As with this fiscal year, we will collect samples and perform detailed molecular characterization of the aerosol to identify chromophores (molecules which absorb light) and the chemical processes generating these chromophores. Quantum molecular calculations of all molecules identified as potential chromophores will be performed to calculate their UV-Vis absorption spectra and the calculated spectra will be compared to the measured spectra. We will add specific chemical reactions in MOSAIC and evaluate the model performance using chamber observations and detailed laboratory analysis of collected aerosol samples. We will apply CAM5-MOSAIC to investigate the effects of the new developments on the global distribution of brown SOA and its radiative forcing.

Microscale Reconstruction of Biogeochemical Substrates Using Combined X-ray Tomography and Scanning Electron Microscopy

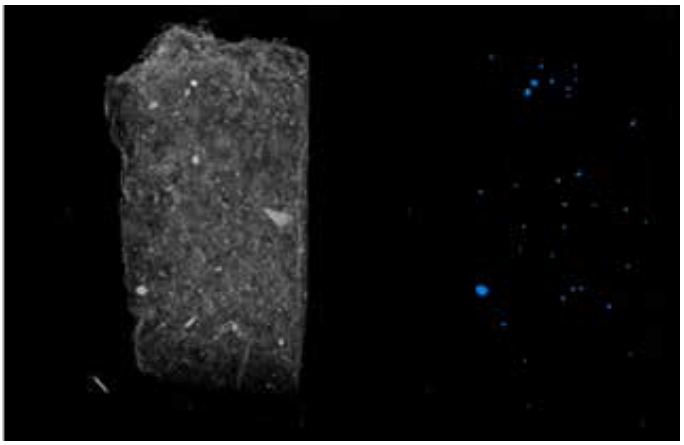
James P. McKinley

The collection of data in the form of sample imagery is a powerful tool used in many scientific investigations, and a family of scientific imaging tools is available to collect complimentary images of solid samples. The ability to combine images across imaging techniques and across imaging scales would represent a scientific tool that is more valuable than its individual parts.

The objective of this research is to combine and examine microscale images across multiple physical and chemical domains. Microscale imaging is ubiquitous in the sciences, and scientific images are common in the popular press such as magazines and newspapers. The best known and most common images are photomicrographs, particularly scanning electron micrographs showing the invisible details of scientific samples; many of us have seen the scanning electron image of a fly's eye, for example. More specialized imagery includes representations of the composition of the sample, including variations in composition; the visualization of 2D spatial relationships among sample parts or subregions; the distribution of different crystalline properties within the sample; and the distribution of void space within a natural or manufactured material. Recently, the ability to produce micro- or nano- scale x-ray tomograms (3D images analogous to CAT scans) has emerged as a significant tool.

Contrast within conventional X-ray tomography is based on X-ray absorption, which in turn depends on a combination of density an elemental composition. Multiple transmission images are collected at different rotation angles that are then reconstructed into a 3D volume. While providing structural material over a large region (~10_9 voxels, typically), X-ray tomography provides little material specificity. Electron microscope images can provide compositional and crystal-phase information but usually in a more limited region. This project and our subsequent work combines different types of electron microscope with X-ray tomography images to provide a detailed 3D image of solid samples that include chemical data.

Our approach is to collect X-ray and chemical data; to assimilate the data using software we develop and adapt; and to progress to more complex volumetric, multi-scale data as our experience and abilities evolve. We initially assembled



600 µm wide 3D reconstruction of natural sediment sample with extracted size and locations of iron sulfide mineralization.

intact sediment and soil samples, and began synthesizing artificial analogues using metal, mineral, and organic substances at different length scales. Tomographic images of the samples were collected, which will be physically sectioned and analyzed via electron microscopy methods. Next, the compositional data will be propagated into the tomographic solid using project-developed algorithms.

For FY 2012, we targeted natural samples with intact pore space and a complex microscale assemblage of crystal phases with varying composition. The first year's work was largely limited to collection of tomographic images and optimizing the resolution of those images across imaging modes (i.e., high and low resolution tomographic images) from the instrument in EMSL as well as the Argonne National Laboratory X-ray beamline, both with and without phase contrast. That work provided the preliminary data needed to begin including compositional imaging from electron microscopy.

In FY 2013, we investigated the combination of compositional and tomographic data using natural samples from both a microbially reduced fine-grained sediment and a hot springs in Yellowstone National Park. We evaluated the results using commercial- and purpose-written software that enabled us to identify specific mineral species within the tomography volume. A 3D reconstruction of fine-grained reduced sediment shows the inclusion of relatively dense iron oxide minerals (e.g., the angular bright clast near the right-hand side edge of the reconstruction). The locations within the volume of biogenic iron sulfides are identified and located based on their density, morphology, and composition.

Our plans for FY 2014 include more disparate samples, such as synthetic materials and anthropogenic glasses.

PN12048/2449

Numerically Robust Climate Simulation Through Improved Interaction between Model Components

Hui Wan

This project will provide new methods for representing the interactions between atmospheric processes in numerical models. It will also help to develop more robust and reliable tools for understanding and predicting climate change.

General circulation models (GCMs) that numerically solve the evolution equations of atmospheric motion have been used for more than five decades as a fundamental tool in climate research. Recent progresses in process studies, the fast growth of computing power, and the decision-makers' need for detailed information of regional climate change have motivated the current emphasis on high-resolution modeling. Complex processes such as aerosol lifecycle and cloud microphysics have been incorporated into global models, which dramatically broadens the spectrum of spatio-temporal scales that are explicitly represented by numerical methods.

Although these processes interact strongly with each other and with the rest of the model, there is not yet an established theory or standard practice regarding the coupling technique. By contrast, a number of studies have shown strong and worrisome sensitivities of model behavior to numerical details. The symptoms have revealed a major weakness in current climate models, which implies large structural uncertainties exist in the simulated climate responses to anthropogenic forcing. The objective of this project is to reduce numerical errors associated with component coupling in climate models. Through development of new coupling methods, we aim to achieve a higher level of numerical and physical consistency between different processes, and constraining artificial sensitivities in climate predictions.

In FY 2012, a review of literature and model codes from various climate research institutions revealed that multi-timescale problems involving compensating and competing processes are a common occurrence in current climate models. A toy problem that describes the temporal evolution of sulfuric acid gas concentration in an aerosol-climate model was investigated. Using a series of numerical simulations, we demonstrated that the widely used sequential operator splitting could introduce severe biases in climate simulations when combined with explicit time integration methods and long-time step. When the strongly interacting processes are treated simultaneously, significantly more accurate results can be obtained without extra computational cost. These results are published in the peer-reviewed journal *Geoscientific Model Development*. Outcomes from the first year enabled us to refine the strategy of this project and concentrate on competing and compensating processes.

In FY 2013, we commenced work on cloud-related processes in the Community Atmosphere Model (CAM) v5. First, the single-column configuration is used to simulate stratocumulus clouds over cold ocean surfaces, a cloud type that has crucial climate impact. Results indicate that in this model, cloud liquid water in stable stratocumulus, is strongly depleted by turbulent diffusion and replenished by large-scale condensation. The process coupling technique currently used in the model can lead to strong unphysical oscillations in the rain production processes. By contrast, if both the turbulent diffusion (sink) and condensation (source) are considered when solving the rain equation, numerical artifacts can be eliminated. This confirms our earlier finding that strongly compensating processes need to handled simultaneously.

Motivated by the single-column model results, a cloud water budget analysis was performed for climate simulations performed with CAM5. Processes in the model that affect the concentrations of water vapor, cloud droplets, and ice crystals were quantified. The main sources and sinks were identified for each of the water species. The analysis led to a clear overview of the relation between moist processes in the model and provided an essential basis for further work on component coupling. In addition to analyzing the default model configuration, we extended the budget analysis to an assessment of numerical convergence with respect to model time step. It was found that although the simulated mean states showed a trend of convergence, a number of process were clearly sensitive to model time step. Initial investigation also suggested that the sensitivities of different processes were caused by different reasons.

In the next fiscal year, we will continue the investigation on cloud water budget in CAM, and develop new process coupling methods based on the budget analysis. We plan to combine global model simulations, single-column tests, and idealized conceptual models to pinpoint the reasons for time step sensitivities in individual processes as well as identify those related to component coupling. Based on the findings, we will develop new coupling methods that provide better convergence properties with respect to model time step. Such methods are expected to help improve the robustness of climate simulations with the CAM5 model, especially in terms of its climate sensitivity and the quality of the future climate predictions at high resolutions.

PN12059/2460

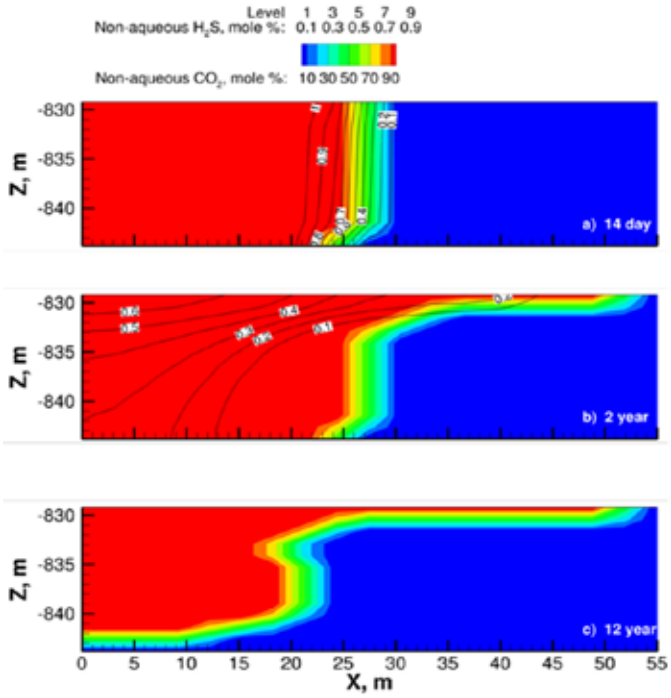
Predicting the Feasibility of Geologic Co-Sequestration of CO_2, SO_x and NO_x Under a Broad Range of Conditions

Diana H. Bacon

We are developing a simulator that will improve predictions of geologic co sequestration impact of CO_2, NO_x, and SO_x on target formation and caprock hydraulic properties under a broad range of mineralogical and phase conditions.

Due to limitations of previous simulators in handling the injection of multi-component and multi-phase mixtures, previous modeling studies of co-injection of SO_2 with supercritical (sc)CO_2 have either simplified hydrogeological systems or have already dissolved in an aqueous phase rather than acted as constituents of the non-aqueous phase. The opposite (and equally unrealistic) extreme case is a scenario in which SO_2 is limited by diffusion through a stationary $scCO_2$ phase. A better method is to simulate the injection of sc/gas mixtures, but a combined component/phase mixture approach is lacking, along with few modeling studies on NO_x co-sequestration impacts.

Our work develops a three-phase variable component nonisothermal simulator capable of mimicking geological sequestration of a mixture of gases in deep saline and depleted oil and natural gas reservoirs. A new simulator that incorporates these complex processes will be used for a survey of the effects of co-sequestration under a wide range of reservoir hydrogeochemical conditions. The results of the modeling survey will help identify the range of reservoir conditions under which co-sequestration is feasible.



Co-injection of CO_2 and H_2S in a basalt formation. H_2S is consumed by mineral reactions more rapidly than CO_2.

To determine the geologic scenarios where cosequestration is practical, we began in FY 2010 by calculating the carbon and co-sequestration potential of 150 minerals to use as a guide for future simulations. Geochemical simulations indicated that co-injection of CO_2 and SO_2 resulted in SO_2 mineral sequestration. In all cases, co-injection of 1 percent SO_2 with CO_2 did not appreciably reduce the amount of CO_2 sequestered nor did it induce a measureable change in porosity vs. injection of CO_2 alone. In FY 2011, we focused on building numerical simulation capabilities for geologically sequestering a mixture of gases in deep saline and natural gas reservoirs. We also developed a kinetic geochemical model of hematite dissolution experiments in CO_2, SO_2 and brine in collaboration with the National Carbon Institute in Spain applicable to co-sequestration in redbeds.

In FY 2012, the component database was expanded to include 10 gases: CO_2, O_2, N_2, Ar, CH_4, C_2H_6, C_3H_8, C_4H_{10}, H_2S and SO_2. The predictions of the new Peng-Robinson equation of state in STOMP for component density, viscosity, enthalpy, and thermal conductivity were compared with the results predicted by higher order equations of state for all these gases in the National Institute of Standards and Testing (NIST) database REFPROP 9. Solubility predictions for each of these gases in pure water and brines were compared to higher-order solubility models and experimental data. The new simulator was extensively tested to ensure that calculations of internal and boundary fluxes are correct for perturbations of all primary variables, including temperature, aqueous, and non-aqueous pressure, non-condensable gas component mole fraction, and salinity under single- and two-phase conditions as well as phase changes. Mass and volumetric source terms were implemented to simulate well injection. We also added the capability to link geochemical reactions with formation minerals to each component to simulate the impact on formation storage capacity due to changes in porosity.

In FY 2013, we implemented a parallel solution scheme using OpenMP, an application programming interface for parallel programming on shared memory machines. Simulations involving geochemical reactions scale almost linearly on multiple processors. As a result of our work, we prepared a journal article on the development and application of the code to co-sequestration of CO_2 and H_2S in basalt. Collectively, our work will guide future studies involving micromodel pore-scale experiments with mineral substrates and mixtures of CO_2, SO_x, and NO_x.

PN10067/2309

Simultaneous ^{14}C and T Dating: A Case Study Using Soil Organic Matter

Jim Moran

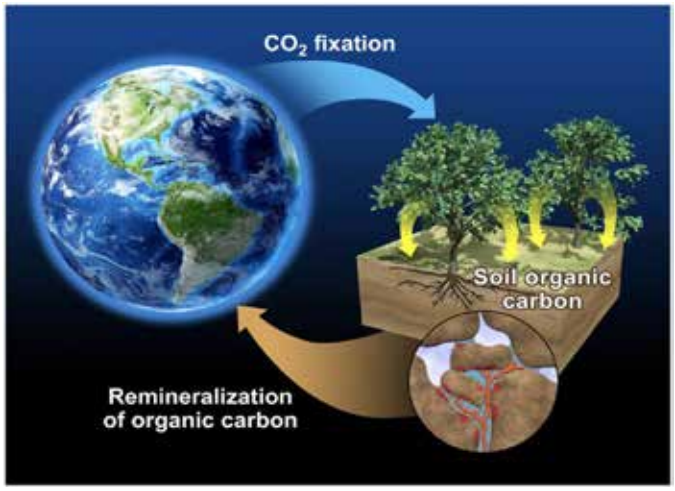
We are developing high sensitivity methods for performing simultaneous carbon (^{14}C) and hydrogen (^{3}H) age dating of organic compounds. Once developed, these unique techniques will be powerful tools for understanding carbon cycling through environmental systems.

Simultaneous ^{14}C and ^{3}H age dating of organic samples is a technique with a great deal of promise for environmental-based applications but also a technique not currently provided in the isotopic community. Current approaches require dividing a sample and processing it independently for each individual isotope, a method neither consistent with small sample sizes nor streamlined for sample preparation. In this project, we seek to develop a robust sample preparation system to allow simultaneous dating of ^{14}C and ^{3}H by proportional counting. By applying a dual-isotope method, we can effectively date specific carbon to hydrogen bonds within a chemical compound rather than determine the ages of specific atoms within the compounds. These measurements will provide us with enhanced resolution for exploring the stability of organic compounds within the environment that will have direct implications on our understanding of global carbon cycling.

We seek to demonstrate this new capability using soil organic carbon as a representative case study. Worldwide, soil stores a tremendous amount of carbon, yet key characteristics governing its stability remain undefined. The dual isotope approach we are developing will offer added resolution to understand what governs stability of soil organic carbon, whether stability is inherent in the type of organic molecule or dependent on its location within the soil. Hydrolysis of bonds in soil organic compounds will likely incorporate hydrogen from younger soil water, creating divergent ^{14}C and ^{3}H dates in the residual structure and indicating compound instability under those conditions. This approach will allow us to assess compound degradation at a very early stage.

The main objective of the project in FY 2013 was to construct the major sample handling infrastructure to convert organic compounds to methane, which is amenable to the dual isotope measurement. We are using a two-staged approach for methane formation: an initial combustion stage and a methane catalysis stage. Sample combustion is through a heated reactor containing oxidation catalysts. Solid or liquid sample is added and the resulting water and carbon dioxide are collected on separate traps. These two products are combined to make methane in the second phase, which uses a high temperature conversion over a metal catalyst. We constructed individual, new reaction lines to accommodate both of these chemical reactions. To date, the newly constructed systems effectively combusted organic compounds and reassembled these products into methane. We also constructed a new vacuum line for synthesized methane purification that is essential for downstream sample measurement via proportional counting.

We assembled two types of soil samples for future analysis. The first soil is an arctic permafrost soil, which contains a disproportionally large amount of soil organic carbon. We seek to use this as one of our initial test samples. We also collected a large amount of highly productive agricultural soil on which we will further test our methods.



Much of the carbon originally fixed by plants is eventually stored in soil. We seek to demonstrate isotope methods by exploring the age of organic compounds stored in soil to provide information about which compounds are more readily decomposed/remineralized to CO_2 and re-enter the atmosphere.

For FY 2014, we will begin making measurements on both test and actual soil samples. Test samples include those measured using traditional isotope approaches, which will confirm the accuracy of isotope measurements. As we refine our techniques, we will transition to analysis of selected soil sample constituents (i.e., humic acids, fulvic acids, water soluble components) to provide a more detailed analysis of carbon dynamics within the soil. Additionally, we will modify our detector housings to reduce background signals, thereby increasing overall sensitivity. This will be coupled with computational modeling to determine the ultimate anticipated measurement sensitivity followed by testing of actual sensitivity with the aim of achieving optimal parameters for our ^{3}H measurements, the effective limiting factor for our overall approach.

PN12071/2472

Tank Residual Waste Stabilization to Reduce Contaminant Release

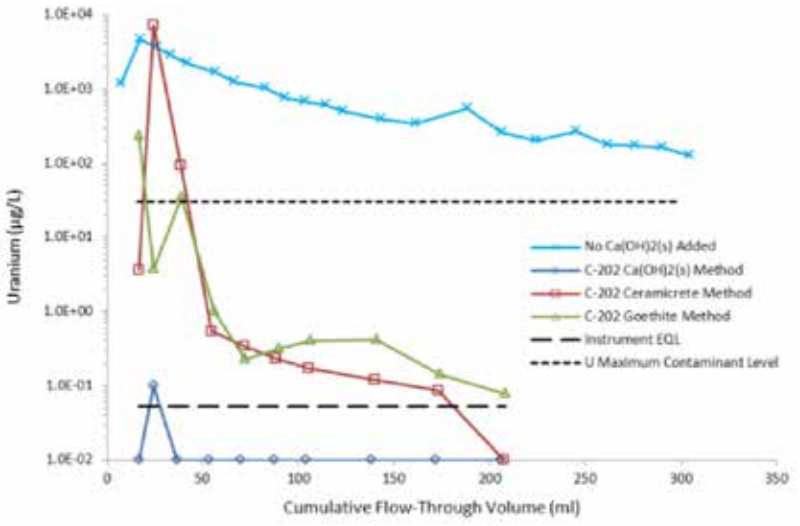
Kirk J. Cantrell

We are developing and demonstrating a science-based approach for stabilizing contaminant release from tank residual waste to determine waste retrieval end-points that are based on risk rather than an arbitrary volume goal.

Recent work at PNNL has demonstrated that the development of the scientifically based mechanistic release model is critical to understanding contaminant release from post-retrieved tank waste residuals. For example, studies about the uranium release from Hanford's C-200 series tanks have shown that in spite of the 99% retrieval goal being achieved, uranium release from the residual waste is still expected to be over three orders of magnitude above the maximum contaminant level (MCL). The focus of this project is to demonstrate that chemical additives can be used to stabilize the residual tank wastes post-retrieval, greatly reducing contaminant release (in particular, technetium [Tc], chromium [Cr], and uranium). This innovative approach has the potential to revolutionize Hanford's tank retrieval process by allowing larger volumes of residual waste to be left in tanks, while providing an acceptably low level of risk with respect to contaminant release that is protective of the environment as well as human health. Such an approach could enable DOE to realize significant cost savings ($10 million or more per tank) through streamlined retrieval and closure operations. With more than 140 single-shell tanks waiting to be retrieved at Hanford alone, total cost savings could be in excess of $1 billion.

With the aid of a thermodynamic solubility release model developed to predict future uranium release from tank residuals, a simple and inexpensive method was designed to render uranium insoluble within the residual waste through the addition of a small layer of lime ($Ca[OH]_2$) on top of the tank residuals prior to closure. In addition, it is possible to immobilize reduced Tc(IV) and Cr(III) in insoluble iron oxide mineral phases through a co-precipitation process that isolates reduced Tc(IV) and Cr(III) from the pore water and prevents their reoxidation to TcO_4^- and CrO_4^{2-}. A third approach was to form stable phosphate bonded ceramics to sequester contaminants. This type of process can be used as part of a risk-based closure strategy to demonstrate that a greater volume of waste can be left in the tanks while reducing actual overall risk levels.

Three treatment methods were tested throughout the course of this project: a $Ca(OH)_2$ addition, a ferrous iron/goethite treatment method, and an *in situ* ceramicrete waste form based on chemically bonded phosphate ceramics. These approaches rely on chemical reduction and the formation of highly insoluble forms of the contaminants of concern. Results demonstrated that for all three methods, the release of the three most significant contaminants of concern from tank residual wastes can be dramatically reduced, compared with the contact with simulated grout porewater. For uranium specifically, all three treatment methods reduced the leachable uranium concentrations well over three orders of magnitude (refer to the figure). For uranium and technetium, release concentrations were well below their respective MCLs for the wastes tested. For two of the three wastes tested, chromium release concentrations were also below the MCL. For the third waste sample, chromium release was above the MCL but was considerably reduced relative to the untreated waste.



Experimental results for uranium, demonstrating the effectiveness of three tank waste stabilization methods relative to grout only (simulated, with no solid $Ca[OH]_2$ added).

For the uranium specifically during FY 2013, solid phase characterization work was conducted to identify the specific mechanisms responsible for reducing contaminant release through various solid phase characterization techniques and geochemical modeling. An article describing the results has been submitted for review to the *Journal of Nuclear Materials*.

PN12073/2474

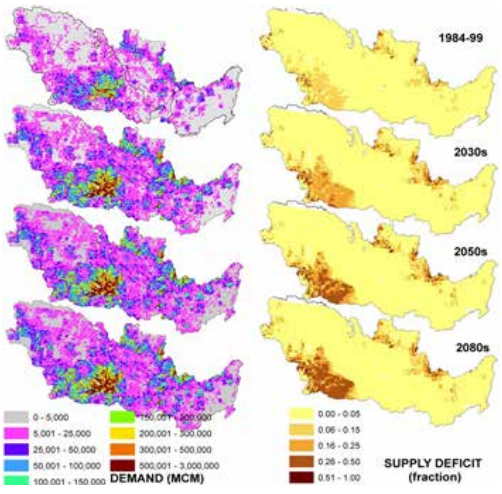
The Integration of Water in PRIMA

Mohamad Hejazi

We are advancing the integrated regional modeling framework in Platform for Regional Integrated Modeling and Analysis (PRIMA) by improving spatial representations of water demands in the Global Change Assessment Model (GCAM) and by reconciling water modeling components to establish inter-linkages that will yield consistent modeling in the water system.

The integration of water representations in integrated assessment and earth system models at the regional scale is an important step toward improving our understanding of the interactions between human activities, terrestrial system and water cycle and evaluating how system interactions are affected by a changing climate. This integration is also critical to enable the PRIMA framework to investigate the implications of climate change impacts and adaptation and mitigation options on water resources (e.g., balances between water demand and supply) in the United States and to determine how water constraints can influence other human decisions and physical processes.

Our project has focused on developing a regional representation of water demands and water allocation framework in the GCAM and creating a linkage between GCAM and the land surface component of Regional Earth System Model (RESM) – including the Community Land Model (CLM), the MOdel for Scale Adaptive River Transport (MOSART), and the Water Management (WM) components – to propagate decisions about water demand per sector and technology from GCAM to RESM at the appropriate temporal and spatial scales. These goals required a careful effort in representing the interaction pathways in the integrated water cycle among GCAM and RESM (i.e., CLM-MOSART-WM) and in addressing any inconsistency in representing the integrated water cycle. This project constitutes a key step for advancing the PRIMA framework toward a consistent, integrated framework of water modeling that is portable, consistent with global modeling and analyses, and provides improvements and insights into the interaction of human decisions and climate changes on regional scales.



Annual total water demand (left) in cubic meters, and fractional water supply deficit (right) for historical and future periods under the B1 emission scenario.

During FY 2013, we successfully implemented a one-way coupling of a GCAM, including a water-demand model with a land surface hydrology – routing – water resources management model (SCLM-MOSART-WM). This coupling required the development of a spatial and temporal disaggregation approach to project the annual regional water demand simulations into a daily time step and subbasin representation. The coupled models demonstrated reasonable ability to represent the historical flow regulation and water supply over the U.S. midwest. Implications for future flow regulation, water supply, and supply deficit were investigated using climate change projections with the B1 and A2 emission scenarios, which affected both natural flow and water demand. The spatial analysis demonstrated the relationship between the supply deficit and change in demand over urban areas not along a main river or with limited storage. This analysis also identified areas upstream of groundwater dependent fields, which exhibited an overestimated surface water demand since groundwater demand is not represented in the current model. This work was published in the peer-reviewed *Hydrology and Earth System Sciences*.

Additionally, we improved the spatial representation of some of the water demand modules in the GCAM framework to track water withdrawals and consumptive use at the state level. Efforts have focused on the energy (electricity generation), industrial (manufacturing), and domestic sectors. Irrigation and livestock water demands are modeled at the 10 U.S. ago-ecological zones. These efforts were complemented with partially accomplishing some of the tasks needed to reconcile water demands and supply in GCAM. At this stage, GCAM can choose whether to irrigate based on economics. This line of work on improving the regional specificity of GCAM has resulted in three additional manuscripts planned for submittal to peer-reviewed journals by the end of this fiscal year.

For FY 2014, the focus will be on implementing a two-way coupling experiment of GCAM and SCLM-MOSART-WM that will require the transfer of information back and forth between the models. This step is also contingent on extending the water allocation mechanism in GCAM to be driven by endogenous water prices that are solved for internally and simultaneously with other markets in GCAM.

PN13001/2482

Uncertainty Quantification and Risk Assessment Pipeline for Carbon Sequestration

Guang Lin

Motivated by challenges in geological carbon sequestration modeling and simulations, we are developing an uncertainty quantification (UQ) and risk assessment pipeline for carbon sequestration to improve the predictive modeling capability for geological carbon sequestration simulators.

With worldwide efficiency improvements underway, fossil fuel usage – a major source of atmospheric emissions of CO_2 – will continue to provide the dominant portion of total energy in the world. CO_2 capture and storage (CCS) in geological formations has become a promising option to achieve the goal of stabilization of atmospheric concentrations of greenhouse gases set at the United Nations Framework Convention on Climate Change. The key issue in the global deployment of geological CCS technology is gaining the acceptance of regulators and the general public, a process that requires scientific risk assessment and cost estimation to understand the full implications. Predictive modeling of multi-scale and multi-physics subsurface systems for CO_2 geological sequestration requires accurate, data-driven characterization of the input uncertainties and understanding how they propagate across scales and alter the final solution.

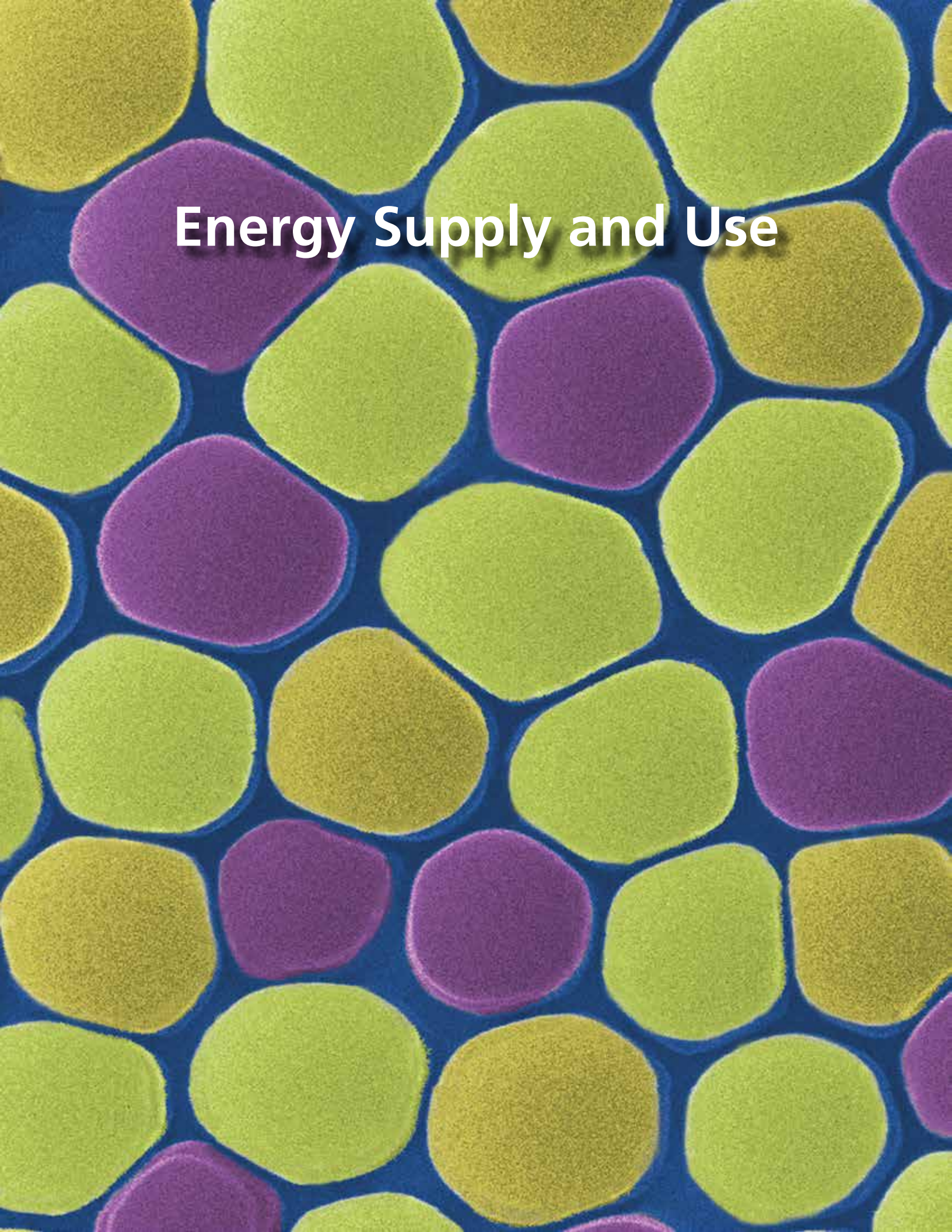
In this project, we leverage existing state-of-the-art tools as much as possible to provide UQ capabilities for the suite of simulation tools and integrate them together as a UQ pipeline. A quantification toolkit and workflow management software involve individual UQ tools such as numerical integration routines, sampling and statistical analysis methods, response surface evaluation, reduced-order model development, parameter calibration/optimization, and workflow management capabilities. The toolkit and software provide reliable means of quantitatively predicting the uncertainty of and assessing the risk to CO_2 geological sequestration, information of significant value to DOE and to our nation.

During FY 2011, we developed a preliminary UQ framework for studying the effect of reservoir spatial heterogeneity on CO_2 sequestration. Continuing into FY 2012, we developed a prototype UQ and risk assessment workflow pipeline based on Velo (a knowledge management framework for modeling and simulation) by coupling input uncertainty characterization, sampling generation, ensemble sampling simulation runs at supercomputer centers, analyzing ensemble results and visualizing the statistical results together. We developed a demonstration package that can automatically generate ensembles given uncertainty distribution, conduct parallel ensemble runs at supercomputer centers, transfer simulation output files back, and visualize uncertainty associated with carbon sequestration to illustrate the capability of our Velo-based workflow for the UQ of carbon sequestration simulations. The developed workflow prototype laid a solid cornerstone for developing a UQ and risk assessment pipeline for the carbon sequestration simulations for testing performed during this current fiscal year.

For a UQ study during FY 2013 on the impact of uncertainty and heterogeneity of spatial permeability field on the CO_2 plume radius propagation, we employed the quasi-Monte Carlo and probabilistic collocation methods from the previous year to generate permeability sampling inputs for eSTOMP and conducted the ensemble runs using eSTOMP on large-scale supercomputer centers. Additionally, we developed a response surface method that can adaptively guide the sampling location and greatly reduce the computational cost in ensemble runs to study the relationship between the uncertainty and heterogeneity of spatial permeability field on the CO_2 plume radius. These results were recorded and submitted to *Mathematical Geosciences*. Further, we explored effects of data quality, data worth, and redundancy of CO_2 saturation data on injection reservoir characterization through PEST inversion, reporting our results to *Greenhouse Gases: Science and Technology*. Further outcomes from this past year included creating research quality codes for building adaptive response surface and developing efficient sampling method to reduce the computational demand for UQ. We worked with another project in which we conducted the model comparison and UQ of the SIMSEQ data sets. We presented these results collectively at both the American Geophysical Union Fall Meeting and the Annual Conference on Carbon Capture Utilization and Sequestration.

To complete our project, we performed the uncertainty analyses of CO_2 plume expansion in reservoirs and wellbore leakage in aquifers. We further developed the UQ and risk assessment workflow pipeline and integrate more UQ components into the pipeline to evaluate the UQ and corresponding risk of uncertainty in hydraulic conductivity on the plume expansion in reservoirs. Finally, we developed several efficient UQ and model calibration methods and software, and submitted several additional manuscripts to international journals.

PN11077/2399



Energy Supply and Use

A Multi-Layer Data-Driven Advanced Reasoning Tool for Smart Grid Integrated Information Systems

Jodi Heintz Obradovich

The objective of this research is to develop a decision support tool – Multi-layer, Data-driven Advanced Reasoning Tool (M-DART) – for improved power system operation.

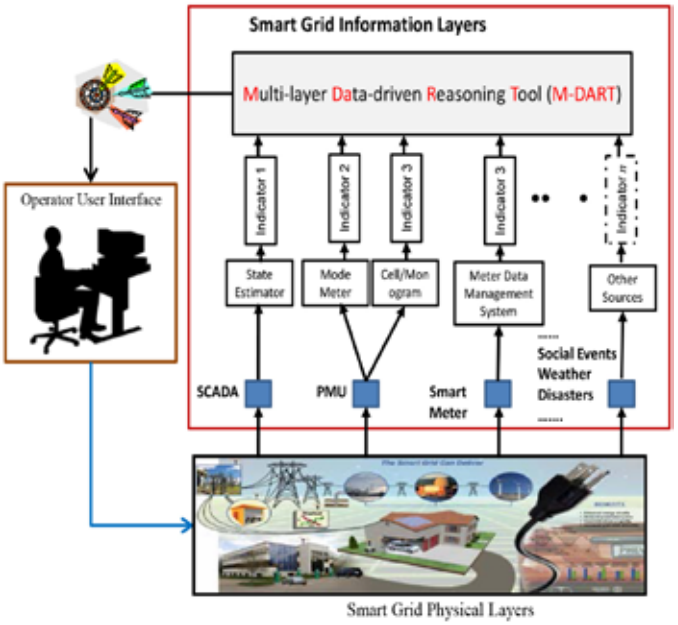
Today's information management in power grid operation tools mainly follow a stovepipe process. Each sensor network produces its own domain-specific data (e.g., phasor measurement unit [PMU], supervisory control and data acquisition, smart meter measurements), and each data source is processed by domain-specific tool sets. Next, outputs are presented to grid operators responsible for making decisions. Isolated data processing mechanisms focusing on individual data domains may not reveal the root cause of failures or accidents and assess their impact to plan for action in a timely manner. As the number of data sources increases, possible action options increase exponentially; therefore, it is a significant cognitive challenge to a stressed human operator to analyze these relationships and/or draw on relevant past experience. Current practice in the power grid operations community is to run massive offline analyses to derive model-based guidelines for online real-time operations. However, because power system components and configurations change in real time, it is hard for offline model-based analyses to maintain relevance and effectiveness.

In view of these challenges, we are using M-DART to investigate existing reasoning tools and apply tools to perform higher level reasoning (cross-correlate and examine data sources to assess anomalies, infer root causes, and anneal data into actionable information) to prioritize problems for the human decision maker. By summarizing results obtained from different reasoning methodologies, M-DART focuses on more accurate early detection of emerging power system events and identifies highest priority actions for the human decision maker to address. M-DART represents a significant advancement over today's grid monitoring technologies that apply offline analyses to derive model-based guidelines for online real-time operations and use isolated data processing mechanisms focusing on individual data domains.

In FY 2012, the efforts were focused on the development of M-DART. Two prototypes were implemented, one of which characterized the PMU data quality. Another M-DART tool is the Columnar Hierarchical Auto-associative Memory Processing in Ontological Networks (CHAMPION), a semantic-graph-based domain-independent reasoning system. CHAMPION can integrate multiple data types and sources. The hierarchy of reasoning components divides the analysis into portions of a potentially large semantic graph to facilitate scalability.

During FY 2013, our efforts were focused on examining the use of multiple data sources for estimating small signal stability. The approach used was to develop test data sets based on a modified mini-WECC model that has as input varying loads and generation characteristics. The test data sets, which consisted of model outputs of simulated PMU and wind and "truth" data (stability estimates) were then used to construct candidate mappings to assess whether the combined information can yield a better estimate of stability than the either of the single data sources. Also this year, the effort commenced to drive the CHAMPION reasoning system.

For FY 2014, in addition to the continuing to develop an M-DART specific CHAMPION configuration, we will develop ontologies representing expert knowledge/higher level reasoning on expertise, including patterns, relationships among events, root causes, and appropriate actions. We will also develop appropriate test scenarios to exercise the reasoning system, integrate them into the GridOPTICS Software framework, and create the required interfaces/displays for the integrated system.



Conceptual illustration of the M-DART architecture

PN12001/2402

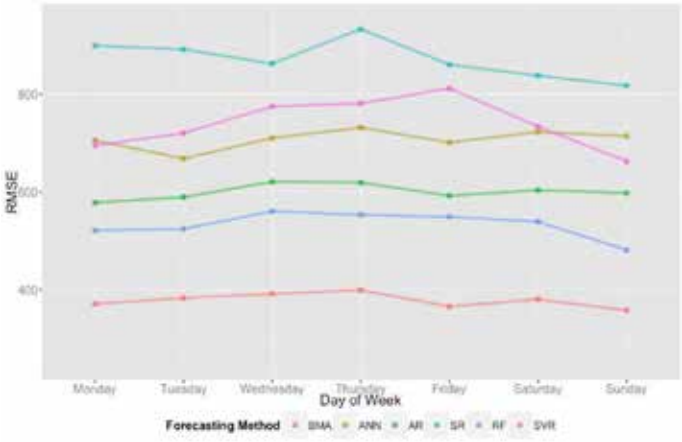
A Statistical State Prediction Methodology to Improve Reliability and Efficiency of Power System Operation

Maria Vlachopoulou

Motivated by increasing uncertainty and variation from the high penetration of renewable generation to power system operations, we are developing a short-term prediction and uncertainty quantification methodology at grid level to enable well informed, proactive operations that will improve power grid operation efficiency and reliability.

A state estimator is an essential tool for power system operation. The estimated states reflect power system status and are the foundation for essential operation decisions. Yet due to the delays from communications and computations in the range of 0.5–5 minutes, current state estimators can only provide power grid status in the past. Operation decisions based on these past states can lower the reliability and efficiency of power grid operation, especially when quick changes and large uncertainty are brought in by the high penetration of renewable generation.

The objective of this study is to enable operators to make well-informed proactive operation decisions by providing forecasted states. To achieve this goal, we are developing a short-term prediction and uncertainty quantification methodology at grid level from three perspectives: the prediction method, which forecasts the values of power grid variables; the prediction error (or uncertainty) quantification method, which gives the confidence interval of the forecast; and the uncertainty propagation method, which calculates the prediction errors associated with the derived variables or states. It is expected that the study results will improve the situational awareness of an operator by providing a comprehensive look-ahead view of power system, which in turn allows time for proactive operations to improve power grid operation efficiency and reliability.



A demonstration case built with integrated prediction and uncertainty quantification methods to show the prediction results using measurement data.

In our previous work, we examined the first field measurement dataset from 2000–2001, during which the preliminary investigation of forecasting methods was performed. We developed a novel efficient uncertainty quantification algorithm and performed an initial study on a sensitivity based adaptive sampling method. In FY 2012, we examined a new set of field data measurement data from 2010–2011 to test the applicability and generalization of prediction methods. We also explored several ensemble approaches (e.g., bagging, stacking, incremental learning) to combine the prediction results from several prediction method to improve the accuracy of the prediction. The sensitivity based adaptive sampling method was refined and compared with Monte Carlo, linearization, and collocation methods based on the IEEE 30-bus model. To achieve same level of accuracy, the sensitivity based adaptive sampling method used 61 sampling points, while the collocation method needs 301 sampling points and the Monte Carlo method requires 10,000 sampling points. Errors using the linearization method were large compared to the other methods. The study shows the significant advantage of the sensitivity-based adaptive sampling method for reducing the computational complexity required to achieve a certain level of accuracy.

In FY 2013, we continued our work on NIS forecasting through improving the performance of the prediction methods and using Bayesian Model Averaging (BMA) as the ensemble approach of choice to combine individual Net Interchange Schedule (NIS) forecasts. The BMA ensemble approach has unique properties that render it superior to other ensemble approaches. The forecasting error for NIS reduced dramatically when we applied our method in comparison to the currently used in industry forecasting algorithm. We received two additional field measurement datasets from 2012–2013 that were used for testing and validation of our method.

The obvious reduction of the NIS forecasting error attracted the attention of our external collaborator, PJM, with whom we are now working on extending this work and integrating it into their system. The success of this project is demonstrated both in the continuation of funding from PJM and the series of prestigious conference and journal articles that summarize this work. A manuscript has been submitted to an international journal to document the developed uncertainty quantification methods.

PN11004/2326

Advanced Visual Analytic for the Power Grid

George Chin

To enable an effective time-critical power grid analysis, interactive visual analytic interfaces that combine complex models and data can provide necessary insights and understanding of the state and course of an existing or proposed power system.

Developing useful, effective visual analytics tools for the future power grid is a significant challenge. The power grid has three challenges of scale: system complexity, big data, and computational requirements. Presenting users with interactive visual analytic interfaces that can manage these three scales while providing an intuitive representation and navigable interactions will be the primary focus of our work. We will build on existing efforts in data management, modeling, and architecture. Developing systems for interactive analysis requires specific design criteria that optimize for interactive analysis. This is often at odds with designing for storage or computation. All three have to be accounted for in a way that balances the needs of each, where the goal is to provide the planners and operators with visual analytics tools that meets mission needs as efficiently as possible.

This project is developing a prototype visual analytics system for planners to assess future grid configurations for supporting power transmission and distribution in the dramatically changing grid over the next two decades. This work will emphasize adaptable visual interfaces supporting deep analysis while leveraging existing modeling and simulation projects and architecture projects. Additionally, we will demonstrate the value of a holistic view of the power grid of the future through powerful visual interfaces. As a result, this project will provide an integration demonstration capability for selected modeling and simulation tasks as well as for the GridOPTICS framework.

Though this project was a late FY 2013 start, our project team visited the Bonneville Power Administration, City of Richland, and Portland General Electric's control centers to observe and study the work that different types of power grid operators and planners perform and the physical and analytical environments in which they operate. The visits provided the team with important contextual information and understanding, and identified situational requirements for the development of visual analytics tools.

In-depth design sessions were conducted with GridOPTICS Software System (GOSS) and model developers to derive power grid analysis use cases to serve as target domain problems for a visual modeling and analysis environment (VMAE). The primary use case will be demand-response modeling using the GridLab-D distribution system to simulate retail markets at multiple feeders and the MatPower system to simulate wholesale markets. Communication between the two systems will be supported and synchronized using the next generation network simulator FNCS streamed into GOSS, where VMAE will access the communication stream, model parameters, and results. Visual analytics tools in VMAE will be instantiated to allow users to steer models, interactively explore data and results, and receive decision support assistance.

A second target use-case was designed for communication link modeling that employs GridLab-D, the power systems simulator Power Systems Toolbox, and the discrete-event network simulator NS3, where communication and data will again be streamed through FNCS and GOSS to VMAE. The Power Systems Toolbox will be applied to generate reduced-order models to evaluate high-speed resolution control. NS3 may then be used to introduce network link failures and delays to enable the study of how network communication problems affect power grid performance and reliability.

For analytics methods and tools, the project compiled an inventory of traditional power grid visualization techniques and advanced visual analytics techniques that have strong potential to support visual analysis of dynamic power grid data, relationships, and behaviors. Traditional techniques include one-line diagrams, contour maps, and line and bar charts. Advanced techniques include geospatial layering, bubble plots, circular and spiral visualizations, network layouts, parallel coordinates, heat and treemaps, and supply-demand as well as community detection visualizations.

In FY 2014, VMAE development and visual analytics tools will preliminarily focus of supporting the first demand-response use case. This will involve developing VMAE as a GOSS application that may publish and subscribe to the GOSS framework. The GridLab-D and MatPower applications will be modified such that VMAE may steer those applications through FNCS and GOSS parameters. Specific transfer mechanisms, protocols, and data formats passed in the communication stream will be fully elaborated and developed. Inventoried visual analytics techniques will be evaluated with power grid engineers and planners at PNNL. Based on the evaluation, a confined set of possible techniques will be iteratively employed starting from mock-ups to limited prototypes to fully functional java-based visual analytics tools. These will be linked into VMAE to be instantiated dynamically with the power grid models and applications integrated into FNCS and GOSS.

PN13081/2562

Battery Manufacturing Cost Model Tool - Flow Battery

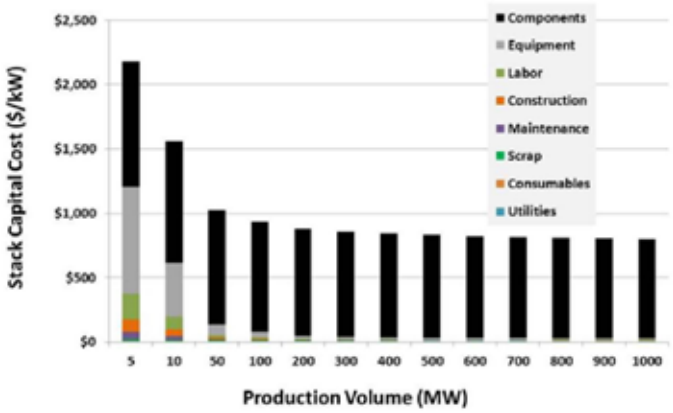
Scott A. Whalen

Our research will lead to lower cost flow batteries, making grid-scale energy storage and renewable integration more affordable.

Flow batteries for grid-scale energy storage and renewable integration will be adopted only if the market determines that they are cost effective. Currently, manufacturing costs alone exceed cost targets for the entirety of an installed system and must decrease if grid deployment is to be achieved. Though manufacturing cost models are not new, we are unaware of published efforts to model flow battery manufacturing costs for large-scale automated production. Any mention that we have seen related to manufacturing cost of flow batteries are unpublished and closely held by private industry. Open-source manufacturing models exist for lithium battery, PEM, LED, and other technologies, but they have limited manufacturing capability. These models typically calculate costs for a specific set of manufacturing processes, and they are inflexible for adding, deleting, and/or changing processes.

Our objective is to develop a manufacturing cost model for redox flow battery stacks. Specifically, the model will be used to perform three primary investigations: evaluate manufacturing costs for large-scale automated production; identify manufacturing processes where innovation would have the greatest impact on cost reduction; and determine costs to meet grid-scale energy storage cost targets. As PNNL has elected to invest heavily in flow battery development, this project was initiated to add another dimension to PNNL's flow battery expertise and research portfolio. The desirable outcomes for this research are to create a public domain user friendly tool, identify manufacturing cost pinch points to drive PNNL innovation investments, and increase PNNL's stature within DOE as the go-to laboratory for all things flow battery.

During the last fiscal year, we developed a redox flow battery manufacturing cost modeling tool for large-scale automated production. The model has a modular architecture that allows for creating modules that are populated with various manufacturing processes. For example, a cell assembly module may be described by multiple processes such as pick-and-place, heat treating, pressing, and dispensing. For each process, the model calculates cost for equipment, labor, construction, maintenance, and the like (see graph). Global variables such as tax and labor rates, depreciation, production volume, stack size, utility costs, and other items are used in the calculations. Initially, the model was developed in spreadsheet form and is now integrated as a tab within the Viswanathan/Crawford cost performance model GUI. The model allows for simple addition and deletion of modules/processes and contains library functionality for creating and recalling input. By design, the architecture is general enough to be used for any domain, energy storage or otherwise.



Manufacturing model with module architecture.

Our model generates a data file from which curves for stack cost ($/kW and $/kWhr) as a function of production volume are constructed. Costs are divided into module, process, and each cost category within a process. These plots show a classic "knee" in the curve, where increasing production volume no longer results in a significant cost savings. Total capital cost is also determined.

Our three most important findings are as follows:

- Subject to numerous assumptions whose validity could be debated, our model shows that stack cost plateaus at roughly $880/kW at a production volume of 200MW/yr. As such, claims by consultants and flow battery manufacturers that massive production volumes (in the GW range) will lead to commodity competitive electricity cost may be unfounded.
- The cost to assemble the stack is insignificant compared to the cost of the individual components. As such, we suggest investigation into the manufacturing costs of the bi-polar plate, membrane, and felt electrodes.
- Capital equipment and labor costs account for > 40% of the manufacturing costs.

Staff at PNNL has limited experience defining processes and equipment for large-scale manufacturing plants. As such, we collaborated with JR Automation to formulate a manufacturing approach and define equipment for a 100 MW and 1 GW flow battery stack manufacturing plant. This adds credibility to our analysis over a broad range of manufacturing volumes.

PN13040/2521

Decision Support for Future Power Grid Organizations

Gariann M. Gelston

This project examines the influence of inter-organizational communication on decision making by power grid operators.

Rapid and effective response to system events is a core challenge for transmission organizations. To help address this challenge, this project analyzes the relationship between inter-organizational communication and grid operator decision making in contingency response. We configured a prototype communication capture tool using the Common Operating and Response Environment (CORE™) toolkit to collect real and simulated communication data among multiple grid entities and investigated how communication affects decision making and grid reliability. While previous research focused on improving individual operators' situation awareness and decision performance, we chose to examine communication and decision making on the organizational level.

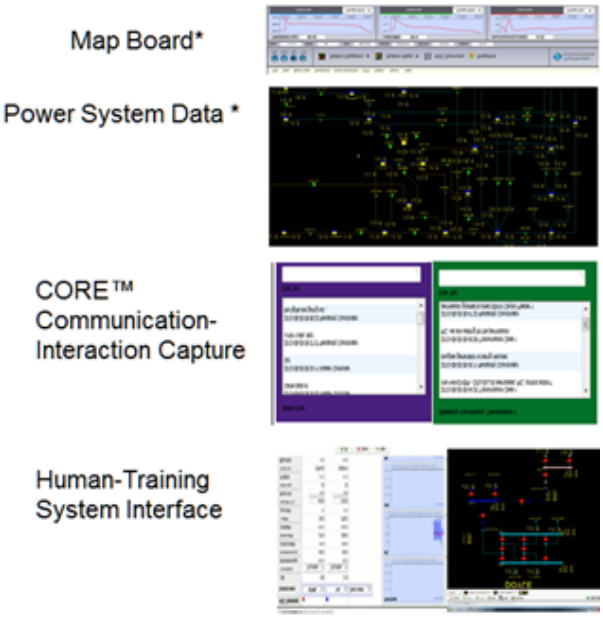
After exploring limitations in the current mode of grid communication, we configured a prototype communication capture tool CORE data collection toolkit and coupled it with an existing power simulator to provide decision support capabilities for improving decision performance of grid organizations in their contingency response and problem solving. Our aim was to develop a communication capture tool and conceptual framework to conduct meaningful grid communication analysis. The project team proactively engaged power industry experts to advise tool development and deployment strategies. The project fostered capabilities to integrate CORE with products from other projects and contributed by sharing industry contacts and lessons learned.

Important findings from previous program years suggest that inter-organizational communication is crucial in successfully addressing unknown, undefined, and unplanned contingencies. Further, the mode, extent, and frequency of communication for distributed decision making is highly variable and complex. Therefore, a systematic communication analysis needs to account for such variability and complexity. Guided by these findings, the project team developed the prototype communication capture tool (CORE) in FY 2012 (invention report filed) and demonstrated it at NERC meetings and industry site visits with positive feedback.

During FY 2013, the project advanced communication research through two activities: secondary analysis of communication data provided by an industry partner, and proof-of-concept coupling between CORE and a web-based commercial power grid simulator. For the secondary analysis, 60 audio files representing a number of grid system events were transcribed and analyzed. The analysis highlights the importance of analyzing communication in the context of the grid system condition that precipitated the communication. To that end, the project successfully identified an industry partner and integrated the CORE tool with the partner's simulator.

Three years of our research effort culminated in a power grid operator training session on August 22, 2013 in PNNL's Electricity Infrastructure Operations Center. During the training, operators utilized CORE effectively to share information and discuss solutions to a black start restoration scenario. This novel coupling aligned the communication data seamlessly with simulated system data to investigate the effect of communication on operators' situation awareness and decision making. Introducing new web-based technologies like the communication capture tool in decision making has the potential to diversify the mode of inter-organizational communication in grid operations. We are proud to be part of this pioneering effort.

Our project bridges across expertise in human factors, organization behavior, and emergency response to develop an organization-driven view of distributed decision support. This work can lead to an in-depth understanding of the key decisions and challenges in organizational communication and coordination, help facilitate broader information sharing and coordination within the power industry, drive innovation in decision support technology in the grid domain, and develop novel solutions to communication challenges for the future grid.



Integrated Communication-Simulation System.

PN11020/2342

Development of the Capability for Assessing Benefits of the Smart Grid to Building Owners

James A. Dirks

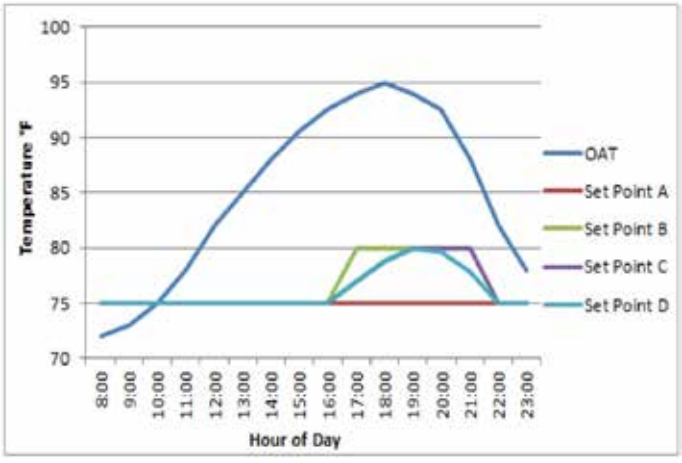
As challenges of affordable energy, energy security, and emissions reduction grow, significant interest in the analysis capabilities of the Facility Energy Decision System (FEDS) has increased. This project enhanced FEDS to allow users to determine the cost and value of various building responses to grid needs.

For the past two decades, the interest in and application of FEDS as the tool-of-choice for analysis of building energy efficiency investment opportunities has grown substantially and is expected to continue. Recently, interaction between buildings and the grid has gained attention, and it has been recognized that buildings may provide highly desirable services to the grid. The modifications made to FEDS as part of this effort allow a user to determine the cost and value of various building responses to grid needs. This tool allows exploration of numerous demand response actions, determination of the benefits to the grid, and the ability to evaluate what the cost is from the building owners/operators perspective in terms of delayed, missed, or less than desired building services. Additionally, we attempted to estimate the likely financial motivation required to induce building owners/operators to supply the desired grid services.

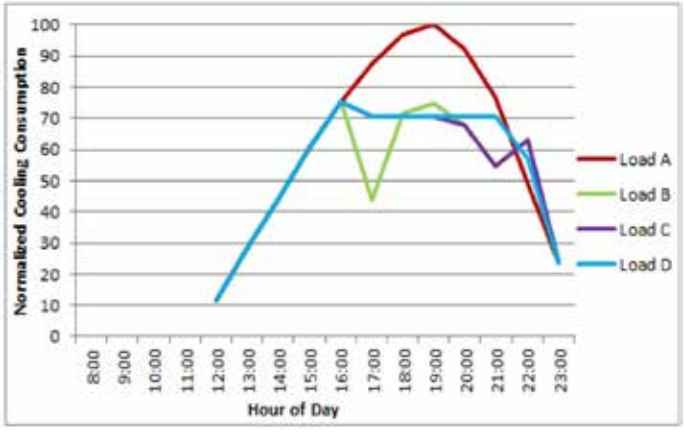
Resulting insights from employing this tool will help support efforts to expound the expected benefits of the smart grid to the public—compensation commensurate with their contribution and lower overall electric rates due to lower required generation, transmission, and distribution capacity in the system that is necessary only to meet the most extreme peaks. Commensurate compensation does not mean cost of contribution (which in fact is generally reasonably low in a monetary sense) but rather sharing the benefits through financial incentives required to induce the behavior desired. This knowledge will allow utilities to tune a demand response program to get the response they want while ensuring that building owners/operators receive the compensation they need. In addition, utilities can determine the optimal mix of demand response, capacity expansion, and load reduction (i.e., efficiency improvements) to meet a forecasted shortfall of generation, transmission, or distribution capacity. This capability directly supports ongoing efforts at PNNL and the nation to understand and quantify the value of Smart+Efficient grid-building integration from the perspective of the grid and building owners/operators.

The result of this effort is a tool that can perform complex evaluations of benefits to building owners from a Smart+Efficient grid-building integration. Through exercising the model, one can examine the value to building owners for varying levels of demand response and investments in efficiency under real-time pricing; determine capacity avoided as a function of building response to a utility signal; estimate the financial inducement required to achieve the desired response and tune a demand-response program at the lowest price; and determine the optimal mix of demand response, capacity expansion, and load reduction (i.e., efficiency improvements) to meet a forecasted shortfall of generation, transmission, or distribution capacity.

As an example, the outside air temperature for the day will generally follow a profile of rising temperatures through the day and falling in the evening. With FEDS, we can now



Example outside air temperature and cooling set point profiles.



Normalized cooling consumption under various set point profiles.

PN13045/2526

explore how varying the cooling set point impacts hourly building consumption. Consider four possible set point profiles: A—constant 75°F; B—constant 75°F until the demand response period (1700–2100) when the setpoint increases to 80°F; C—constant 75°F until the demand response period when the setpoint is increased gradually to 80°F; D—constant 75°F until the demand response period when the setpoint is increased gradually to 80° and then decreased to 75°F. The resulting cooling energy consumption profiles are described below:

- Cooling Load A peaks at 100, inside temperature is always 75°F, and consumption averages 90.7 during the demand response period.
- Cooling Load B peaks at 75 (same as peak before the demand response period), inside temperature averages 80°F during the demand response period, and consumption averages 62.6 (the minimum) during the demand response period.
- Cooling Load C peaks at 70 (tied for minimum during the demand response period), inside temperature averages 79.1°F during the demand response period, and consumption averages 66.8 during the demand response period.
- Cooling Load D peaks at 70 (tied for minimum during the demand response period), inside temperature averages 78.6°F (the minimum and providing the least thermal discomfort) during the demand response period, and consumption averages 70.5 during the demand response period.

Clearly, the choice of set point profile significantly impacts the cooling consumption and the best choice for a building owner depend on the owner's objective function (minimize peak demand, minimize consumption, minimize energy consumption, minimize discomfort) and the financial inducements that the utility is providing. FEDS can now evaluate real-time price, time-of-day rates, demand charges, and demand ratchets.

All projects goals and objectives were accomplished and demonstrated. FEDS can now model real-time electricity pricing and varying levels of demand response actions for each of a number of loads, which would include the following in order of importance:

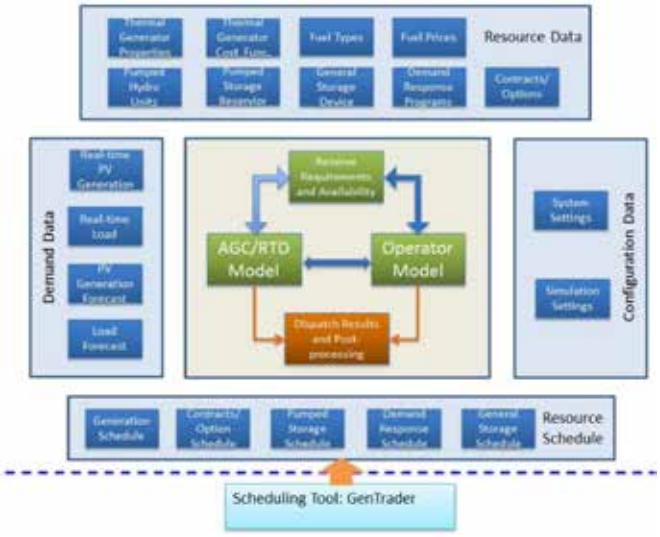
- HVAC set point modification, including recovery
- Demand controlled ventilation (outside ventilation air reduced during response period reducing HVAC loads), including pre-ventilating before the response period
- Lighting response
- Major appliance delay (e.g., clothes washer and dryer).

Electric System Intra-hour Operation Simulator

Shuai Lu

This project is developing a software program Electric System Intra-hour Operation Simulator (ESIOS) for high time-definition dispatch analysis of power systems. ESIOS enables accurate quantification of the impact of variable resources and an assessment of new operation technologies.

With increasing penetration of variable renewable generation (VG) such as wind and photovoltaics (PV), balancing power system generation and load has become more challenging. Operators often find that the scheduled generators have insufficient flexibility to follow real-time changes caused by load and VG. A more detailed look at intra-hour balancing processes is needed for these systems. New types of resources such as energy storage and demand response (DR) into automatic generation control (AGC) and real-time dispatch (RTD) services have been proposed to increase system flexibility. AGC and RTD control approaches are being improved to accommodate these new resources. Market rules to compensate for better regulation performance are under development from FERC Order 755 on Frequency Regulation Compensation in the Organized Wholesale Power Markets. Models and software tools are not yet available for control approach test and evaluation of market rules. Also, existing operation tools have focused on hourly scheduling of generating resources, an effort that is not adequate to solve the challenges from increasing wind and PV energy.



Structure of the ESIOS Program.

This project is developing a software program, Electric System Intra-hour Operation Simulator (ESIOS), for high time-definition dispatch simulations. Based on hourly resources, schedules are produced by unit commitment tools, ESIOS models system AGC, and RTD for frequency regulation and load/variable generation following, respectively. ESIOS adopts an open software architecture design, which allows it to be integrated easily with hourly generation production models, and its function modules conveniently expanded or replaced for various research purposes. ESIOS can be used as a planning tool for the assessment of system flexibility and as a testbed for new storage and DR control methods and market rules. The open software architecture is designed for ESIOS to be used by researchers as a convenient test platform for new scheduling and AGC control algorithms, which will help the simulator to evolve with new developments in relevant aspects of power system technologies and policies.

In FY 2013, we redesigned the structure of ESIOS to accommodate models of different systems. A graphical user interface (GUI) was developed to enable users not familiar with the program to build simulation cases conveniently. Data interface format with generation scheduling software was defined and tested with commercial software named GenTrader, which is widely used by utility companies in the United States. New models such as pumped energy storage and demand response were developed. Additionally, a pumped storage model was integrated with the ESIOS program.

As a result of the above effort, ESIOS software has been applied to utility resource planning studies as a tool to analyze VG integration impacts and mitigation approaches. It also has the potential to provide detailed resource capability forecast and dispatch decision support for real-time operations. The team is working toward this direction in a project with University of California, San Diego.

PN13062/2543

Future Power Grid Control Paradigm

Karan Kalsi

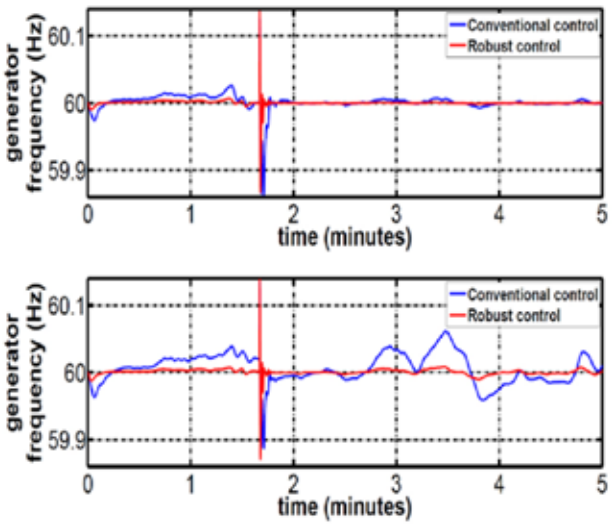
This project will introduce a novel control paradigm that enhances reliability and efficiency of the future power grid by exploiting the value of new monitoring and information technologies coupled with new business needs and policy changes. This work will lead to creating a more flexible and reliable U.S. energy infrastructure.

With the premise that transmission system loading will continue to increase and the trend will be toward maximum utilization of transmission system close to thermal capacity, security and reliability will are crucial issues for the future power grid. The problem of power system security and reliability is evident considering past events such as the 2003 Northeast Blackout, which swept out a large geographical area in 9 seconds, affected 50 million people, and caused ~ $10 billion dollar losses. With the advent of the smart grid, the current power grid requires a new look at monitoring and operation capabilities and a new from-the-ground-up design of the control architecture.

This project constructs a new control paradigm based on robust, coordinated, and distributed control architectures. The proposed control paradigm for the future power grid can enable system wide control, allow more integration of renewable energy sources, exploit the potential of demand response and significantly reduce the risk of cascading failures ultimately leading to blackout events. The performance of control scheme is benchmarked against current practice using several test systems. The suitability for high performance computing (HPC) will be evaluated to ensure broad applicability of the proposed architecture on large-scale systems.

In FY 2012, we offered a novel modeling framework that represented multi-machine power systems as large-scale interconnected systems. Automatic generation control (AGC) was modified and coordinated with decentralized robust controllers to restore tie line power to their scheduled values. We also studied the interactions of typical remedial action schemes (RAS) with the distributed hierarchical controller to identify a path for optimal coordination of two control actions. It was demonstrated that the proposed robust control strategy reduced the burden on the RAS, which in turn reduced system stress from tripping of generators.

Distributed generation, demand response, distributed storage, smart appliances, electric vehicles, and other emerging distributed smart grid assets are expected to play a key part in the transformation of the American power system. Due to the variability and uncertainty associated with these resources, there is much trepidation from system planners and operators about the controllability of such resources and how they affect the stability of the grid infrastructure. With the continuing addition of renewable generation, more resources are needed to compensate for the uncertainty and variability caused by large-scale integration of renewable resources. Our work in FY 2013 tested the robustness of the distributed robust control architecture with different levels of high penetration of wind in the system.



Generator 2 frequency (Hz), top : 10% wind generation penetration, bottom: 20% wind generation penetration).

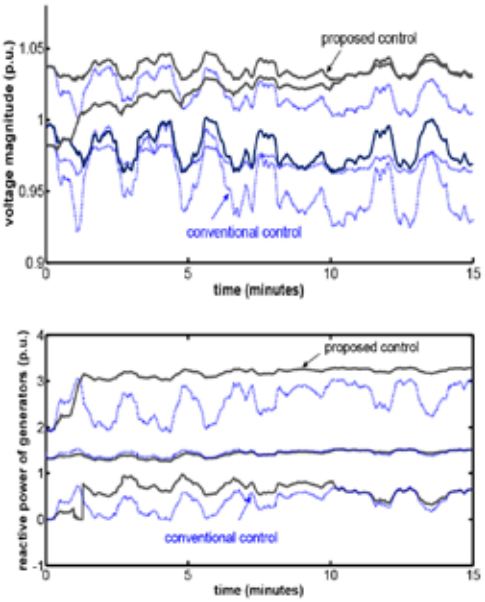
The above scenarios involve a three-phase short-circuit fault followed by tripping a generator, while the wind speed experiences a sudden increase. The higher the level of generation variability caused by wind power, the larger the frequency deviation from its nominal value under conventional control. On the contrary, the proposed robust controller manages to recover the frequency faster and keep it within tighter bounds.

Conventional generators are an important reactive power source and contribute significantly in maintaining voltage stability. The increasing penetration of wind farms will displace some of the conventional generators, and their capability to supply reactive power may need to be replaced. By increasing the amount of wind generation

PN12035/2436

in the power system, the system reactive demand is being changed with increasing demand/supply volatility. A three-level hierarchical voltage control architecture was successfully applied in Europe and China that consisted of primary, secondary, and tertiary voltage regulation. Implementation of this hierarchical scheme may vary for different requirements and definitions of voltage regulation in different countries.

Several weaknesses have been determined, however: the architecture does not consider discrete actions (such as on/off status of capacitor banks and switchable reactors), does not consider the future system status, and does not address uncertainties arising from the system topology in load/wind forecast.



Voltage and reactive power with 30% penetration of wind

Additionally during FY 2013, a predictive supervisory controller was planned that coordinates different voltage resources in a more efficient manner. The objectives are to optimize allocation of reactive power reserves, improve voltage profile, and minimize switching action of devices over prediction horizon. Uncertainties are addressed by solving the optimization framework over different scenarios. The optimal control strategy was tested on IEEE 10 machine 39 bus system with 30% penetration of wind generation. It was shown that the optimal control strategy led to an improved voltage profile and more efficient allocation of reactive reserve as compared to the existing controls.

In FY 2014, we will peruse engagement with utilities such as PJM, AEP, and ERCOT to take the voltage controller approach closer to the application and meet better the industry needs. At the same time, including discrete on/off actions in the voltage control approach will be investigated, which has the advantages of discrete reactive power resources like capacitor banks and tap changers. We will also investigate the question of predicting the effects of the proposed robust decentralized controller on system-wide parameters. The use of look-ahead dynamic simulations will be considered to evaluate possible issues like transmission congestion. Mitigation and coordination strategies will be developed by adjusting controller design based on wide-area information/prediction. Finally, the coordination of the distributed robust controller with tertiary controls (such as economic dispatch, unit commitment, etc.) will be investigated. In particular, impacts on existing market rules will be studied (e.g., analyze payback to generators).

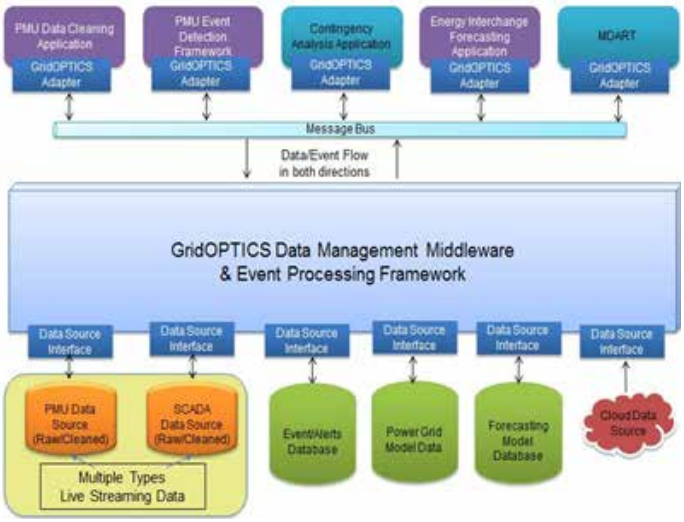
GridOPTICS

Bora Akyol

GridOPTICS Software System (GOSS) facilitates the creation of new, modular, flexible operational and planning platforms that can meet the challenges of the next generation power grid.

The U.S. power system is evolving at a rapid pace. New intermittent distributed power generation and storage technologies, new control paradigms, and new higher fidelity sources of monitoring system state such as phasor measurement units (PMUs) are being deployed today. All of these advances are causing the power system to become more dynamic and have more uncertainty. Computational analysis of the power grid for operational and planning purposes needs to evolve to match this emerging dynamic behavior. In turn, software tools available for power grid analysis were developed under a traditional, sequential, workstation-based software paradigm and are locked into existing vendor software systems for energy and distribution management. This makes it difficult to infuse novel and emerging analysis and data management techniques into power system operations to meet growing computational and big data needs.

GOSS project is developing a flexible, scalable software architecture that engenders innovation in power grid analytics and data access. Through a set of open application programming interfaces and a robust, high performance and extensible implementation, we are creating a framework in which new and novel data management, analysis and user-facing technologies can be easily integrated. Primary achievements for GOSS will include integrating a range of data collection, analysis, simulation and visualization technologies to support the operations and planning of the future power grid; providing a framework for integrating novel operations and planning technologies with external power grid systems, including energy and distribution management systems developed by software vendors; providing a client software interface to the GridPACK numerical library; and creating an interface to launch and collect data from high-performance computing applications.



GridOPTICS Software System Architecture.

In FY 2013, we developed a GOSS prototype that supports the requirements of high frequency data access and upload from operational, planning and visual analytics tools running on platforms from Windows workstations to high performance computers. The current prototype was integrated and tested with multiple tools and applications, including PMU data cleaning and event detection, graphical contingencies analysis, NIS forecasting and shared perspectives, and Distributed Systems Architecture.

To provide data to these applications and store generated results, current GridOPTICS framework supports data source interfaces for retrieving and uploading PMU measurements (raw/cleaned), load schedule, power grid network topology and contingencies for power grid. Currently, these data sources include MySQL and GridMW. Developers building software tools using GOSS reported that they benefited from data reusability and automated translation, timestamp alignment for different sampling resolutions, multiple data source support in one query, easy to implement client interface, multiple language/platform support at client side, and support for one-time and recurring (event-based) queries.

Based on feedback from stakeholders and experience gained by integrating different tools, our project has determined multiple target tasks for FY 2014. First, we will integrate the simulation-run capability in the GOSS framework for desktop as well as HPC machines and stream results back to application (time synchronized). We will implement associated rule mining for pre-fetching data based on past trend, which would increase the efficiency and particularly speed up run time of operational tools. Next, we will develop adapter code generation tool for data model conversions that would be implemented for a domain specific language. We will integrate with Distributed Systems Architecture to have service registry and semantic capabilities and work to implement security to include identity management, access control, application validation, and resource management. Additionally, we will develop a standardized and comprehensive data model to exercise all codes. This model will be based on CIM standards for power grid. We will also conduct a performance and scaling study to identify performance bottlenecks and possible improvements. Finally, we will integrate with other projects to align with the GridOPTICS vision.

PN13016/2497

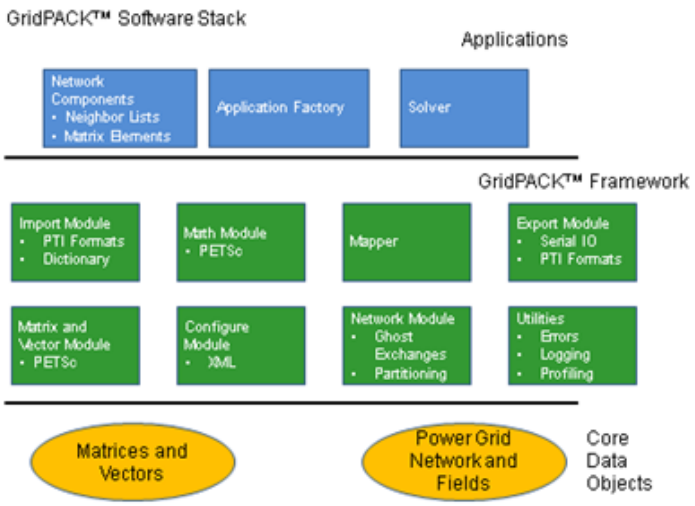
GridPACK: Grid Parallel Advanced Computational Kernels

Bruce J. Palmer

The software framework GridPACK™ will allow engineers to run more sophisticated models on larger systems in less time, providing more detailed predictions about the power grid and leading to greater reliability and more efficient utilization of grid resources.

Although the electric power grid represents the largest machine in the world, much of the modeling done on it is still performed using programs running on single processor workstations. Recently, there has been some movement to extend modeling to multithreaded codes running on a single compute node. However, this activity still represents a severe constraint on the size of the system modeled and the amount of computing resources that can be applied to a single problem. Parallel programs can enormously increase the amount of computing power to be applied to problems, but the difficulty of writing parallel programs has been a major barrier to adoption by the power grid community.

The goal of the GridPACK™ project is to create software libraries designed to support power grid modeling that will substantially lower the effort required to develop power grid applications that can run on modern high performance computer (HPC) architectures. This project will identify common software patterns required in parallel power grid codes and encapsulate these in libraries that can be used to write new applications quickly. In turn, these libraries will reduce the overall complexity of applications by providing simpler programming structures for writing code by reducing the need to re-implement common programming structures and eliminating the need for the complex indexing and communication programming that makes parallel programs so much more complicated than conventional serial programming.



Schematic of GridPACK software stack with three levels: basic distributed data objects (gold), framework components (green), and user-developed components specific to the application.

Our main accomplishment from FY 2013 was the development of a design for the GridPACK software stack. The original GridPACK design identified several major modules that would support software development of power grid codes on HPC platforms, which included the following components:

- A network module that would keep track of network topology, enable exchange of information between network components on different processors, and house a partitioner that would divide the network in optimal ways between different processors.
- A math module that would house distributed matrices and vectors and would support basic algebraic operations such as matrix-vector multiply, inner products, norms, and the like. This module contains the linear and non-linear solvers needed to solve the algebraic equations generated by power grid models and analyses.
- A mapper module that provides a mechanism for mapping between the network topology, where the original models are based and the matrices and vectors generated by the models. This is frequently a complicated, error prone portion of setting up parallel codes. Simplification of this task can significantly reduce coding effort and complexity.

Other modules identified as part of this project included those to support input and output, a configuration module to support user supplied parameters, and utility modules to support error handling, profiling and logging. These findings were summarized in a design document that has been used as the basis of continuing development of the GridPACK framework.

The specific intent for this scalable software framework is to provide high level modules accessible to the power domain expert that can be composed in a straightforward manners to create simulation applications. The framework is intended to be able to exchange data with existing (and possibly new) power system applications and become an I/O module of prime importance.

PN13017/2498

Hardware in the Loop Testing and Power System Simulation of High Penetration Levels of PV

Kevin P. Schneider

The growing integration of renewable sources of electricity into the nation's electrical infrastructure will provide many social and environmental benefits. Before high levels of penetration can be achieved, technical and operational problems must be solved and optimal control schemes must be developed to overcome current barriers to greater adoption.

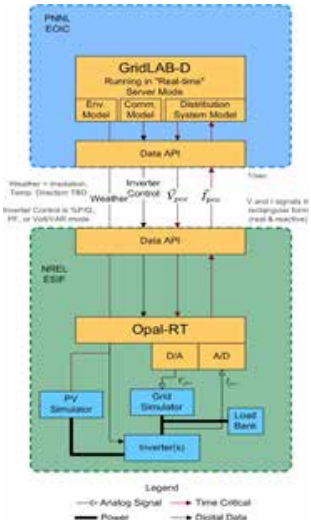
Operational impacts of high penetration levels of solar photovoltaics (PV), including voltage rise, voltage flicker, phase balancing, and power factor correction, are of significant concern to distribution utilities and may pose a barrier to integration. To study these impacts and develop effective mitigation strategies, it is necessary to combine detailed models of hardware as well as a time-series simulation environment that can capture system-level effects.

To build a platform that can successfully address these issue, PNNL and the National Renewable Energy Laboratory (NREL) are collaborating on the integration of real-world devices and detailed system-level simulations in a hardware-in-the-loop (HIL) environment. Expertise in system-level modeling within the GridLAB-D simulation environment at PNNL has been combined with knowledge of hardware and laboratory facilities at NREL. Leveraging the capabilities of PNNL and NREL, tools and capabilities essential to accelerate the integration of renewables can be developed.

The HIL system under development embeds PV inverters at the NREL Energy Systems Integration Facility (ESIF) into a simulation of a test feeder running within GridLAB-D in the PNNL Energy Infrastructure Operations Center (EIOC). The inverters communicate with the GridLAB-D model in the EIOC using Java Script Object Notation (JSON) at 1-sec intervals over a User Datagram Protocol (UDP) link. In FY 2013, all scheduled pieces of the HIL system were successfully tested. The IEEE 123 node distribution feeder was selected as the simulation model and populated with detailed physics-based models of both residential loads and solar PV. A secure tunnel was established between the EIOC and ESIF in accordance with the cyber-security requirements of both laboratories. Communication tests were successfully performed between a JSON link demonstration script acting as a proxy for GridLAB-D and for the hardware set-up on the Opal-RT real-time machine in ESIF. The test script used the same protocols, ports, exchange rate, and configurations that will be used for the final HIL setup. The lost packet rate was determined to be under 1%, and the long-term stability was established by operating continuously for 48 hours. Additional tests established the capability of a GridLAB-D simulation to operate in real-time mode, communicated via JSON link to a proxy script at ESIF, and sent continuously updated values to a local HTML user interface and display.

Hardware set-up at ESIF is complete, fully tested, and calibrated. Due to technical and safety related issues, hardware installation at the ESIF was delayed, requiring compression of the testing and validation schedule. However, communication and control pathways have been successfully established between the Opal-RT HIL real-time machine, the inverter, and the solar simulator. The necessary work to enable the Opal-RT machine to communicate via JSON link was completed and tested. The final step is the integration of each piece of the HIL system, which will be used to perform initial experiments, focusing first on validation of GridLAB-D inverter and PV models and moving into testing and developing inverter control schemes within simulation and on the actual inverters.

Efforts during FY 2014 will add capabilities to the GridLAB-D HIL platform on two fronts: expanded communication and new sub-second HIL experiments. The GridLAB-D hardware-to-simulation communication options will be expanded to allow the integration of PV inverters, building Energy Management System (EMS) and Community Energy Storage (CES) systems using a common communication protocol. This will enable the study of how the balancing potential of building EMS and CES can be leveraged to increase the level of PV penetration without negatively affecting power quality or reliability. Sub-second simulation capabilities of GridLAB-D will be made compatible with the real-time analysis mode and HIL experiments. Many of the operational impacts discussed above are expressed at a faster timescale than the 1s timescale of this past year's work. HIL experiments conducted with ms temporal resolution will allow more complete validation of GridLAB-D models of inverters, building EMS components, and more accurate modeling of control response.



The integration of PNNL EIOC (GridLAB-D) and NREL ESIF (Opal-RT) capabilities

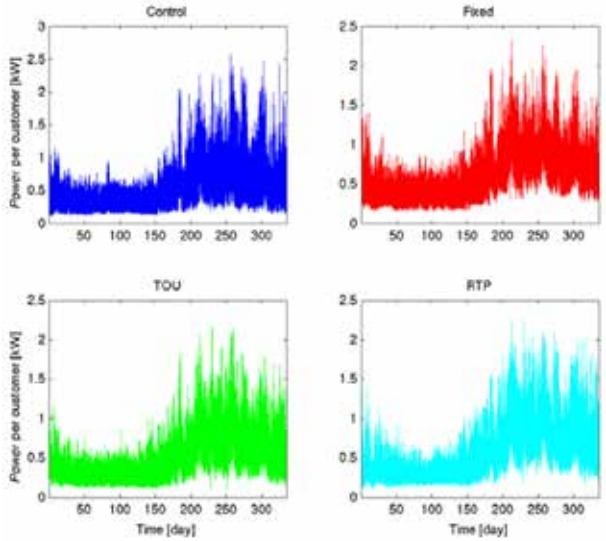
Hybrid Load Control Strategy for Demand Response System

Yu Zhang

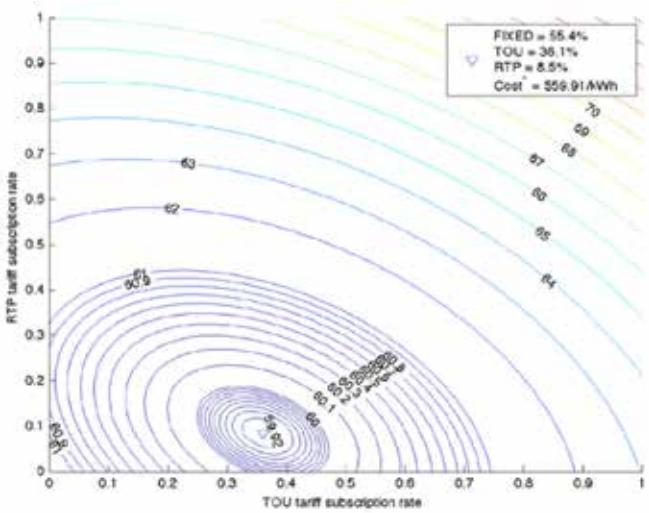
Our project is developing a portfolio analysis based control system for recruiting, committing, and dispatching electricity consumer demand response.

Load modeling is an essential component of utility planning and operations. Specifically, it can allow short-term optimization of utilities' tariff portfolio without requiring regulator's approval. Additionally, the modeling has a great potential to enhance the value of demand response significantly. Because a significant amount of work has been done on load models, aggregation methods and those of measuring and estimating aggregated loads, and different device level controls, these approaches can be used for creating this hybrid load control strategy. The portfolio analysis based methods have never been applied to the tariff subscription problem. The methods proposed are familiar to economists and readily applicable and has the potential to enhance the value of demand response significantly.

Recent developments in smart grid technologies, specifically thermostatically controlled appliances (TCAs), have traditionally drawn more attention than other end-use loads because of their ability to shift energy consumption within a certain time period by storing electricity as thermal energy. Traditionally, the TCA loads are scheduled using unit commitment either to reduce peak loads or provide spinning reserves so that system operation cost is minimized or integrated with interruptible load management. This method provides instantaneous reserves for ancillary services in deregulated power system. The optimized direct load control on TCA loads is to turn the unit on or off directly instead of shifting the unit's thermostat based on its states in the group. The output of this optimized direct controlled load can be controlled precisely with mitigating changing the diversity of the load and impacting the customers comfort. However, the optimized direct load control only fit for a relative small proportion of customers. By contrast, the indirect load control and autonomous load control will fit for a considerably large scale. Indirect controlled or autonomous load customers are given incentives to change their electricity demand according to system requirements. They use local voltage, frequency, or global price information as control signals and contribute to overall system stability without two-way peer-to-peer communication networks. However, most indirect and autonomous control strategy cannot increase or decrease the load precisely, compared to the direct control.



Raw telemetry for the Olympic Peninsula project.



Optimal subscription rates for minimizing the cost of the Olympic Peninsula project

Through this project, we aimed to understand the strengths and weaknesses of existing aggregate load models from an aggregate load control viewpoint and existing device-level load control from an aggregate load control viewpoint. We will lay the technical foundation for the development of a portfolio analysis and relevant optimization problems of customer subscription rate.

An easy-to-use solution has been found to an emerging smart-grid cost reduction opportunity for utilities wishing to maximize smart meter technology and new tariffs made possible by recent innovation in smart-grid technology such as those demonstrated in the Olympic Peninsula project. Utilities can use this method to determine the optimal mix of customer tariff subscriptions for a variety of business objectives using high-resolution metering data collected from smart meters.

In the case of the cost-minimization objective, we have shown that the problem is convex when considering both energy cost and the cost of serving an uncertain load peak. Thus, the problem is solved efficiently in closed form and becomes amenable to large utility operations without significant computational burden.

PN13103/2584

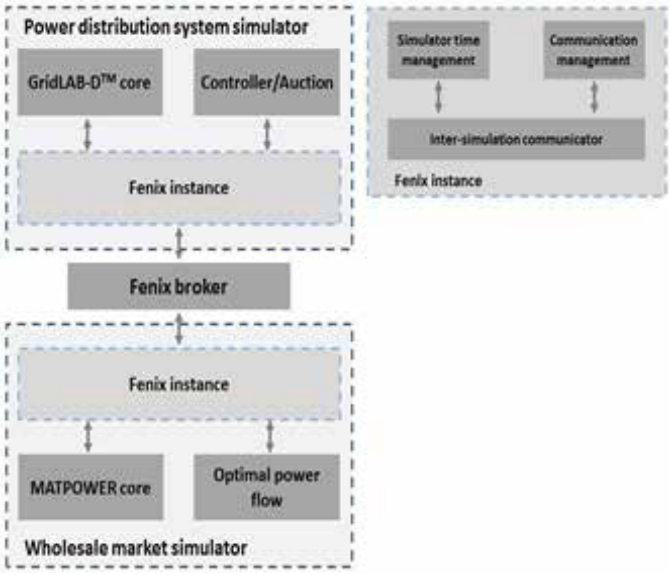
Market Design Analysis Tool

Abhishek Somani

This project will create tools that will help in the analysis and design of market mechanisms that enable participation by demand response resources (DERs; retail customers) into electricity markets.

The retail sector in the U.S. power system has not undergone a change similar to the wholesale sector, where market-like constructs are used to transact electricity on various timescales. Hence, retail customers have no mechanisms or incentives to participate actively in transactions. The area of wholesale-retail integration is also not extensively studied; the little information that does exist is limited by simplistic assumptions used to model distribution systems. To improve power grid efficiency, however, it is imperative to understand the relationship between wholesale and retail energy markets, and how one influences the other. The most effective way to evaluate this interaction is through detailed simulations and analysis of different operation scenarios. Simulating these two independent power markets in tandem proves is challenging because they do not normally coexist under the same simulator framework.

The project will design mechanisms that will enable large-scale participation by retail customers into wholesale markets. The issues addressed constitute the core of transactive control mechanisms that PNNL researchers have been working on for years. Our project will advance this mission by designing additional transactive mechanisms that expose customers to a wide array of value streams, originating both at the wholesale and retail layers of the electric power system. As the system is rapidly changing from centralized generation and distribution to one where power is generated using distributed energy and DERs, it is important to establish the right mechanisms that provide sufficient incentives to grid services.



Two intricate, robust simulators are MATPOWER and GridLAB-DTM. MATPOWER is an open-source power system simulation toolbox written in MATLAB. It is used to solve AC/DC and optimal power flow (OPF) solutions at the transmission level. Additionally, it incorporates functions to execute OPF-based auction markets and co-optimize reserves and energy. These resource allocations and prices are part of the auction clearing mechanism and are determined by the OPF based on cost (generators offers to sell, and loads bids to buy). GridLAB-DTM is a software platform developed to study and analyze the power distribution systems for end-use customers (GridLAB-DTM). Among many features, it allows a detailed modeling of how retail markets interact with latest end-use technologies, and distributed energy resources. Integrating these two simulators to study the "smart market" is done using the Fenix framework. The main purpose of this project in integrating these simulators under the same framework is to create a test-bed that would allow a thorough analysis of how these power markets operate over accurately modeled transmission and distribution grids. Transactive control mechanism used in PNNL's gridSMART smart grid demonstration project is our test case to demonstrate the capabilities of this test-bed.

In FY 2014, the integrated retail-wholesale test-bed will be used to design transactive mechanisms that enable retail customers to participate and extract value, by providing energy and ancillary services to the grid, in the wholesale markets. These mechanisms may be market-based, such as double-auction markets, ¬bilateral negotiations, or tariff-based, such as time-of-use and critical-peak-pricing programs. This will enable the study of the impact of DERs on wholesale energy and ancillary services prices. Additionally, the participation of these resources into newer markets, such as markets for ramping or load-following products, will help evaluate their ability to integrate more renewable power resources into the power system.

Further, we will investigate alternative business models that may be used by utilities and third-party service providers. The models employed by these various entities range from providers of electrons to that of value-added services to customers. In the latter case, the project will also utilize concepts from finance theory such optimal portfolio to construct the optimal package of services to be provided to the customers.

PN13029/2510

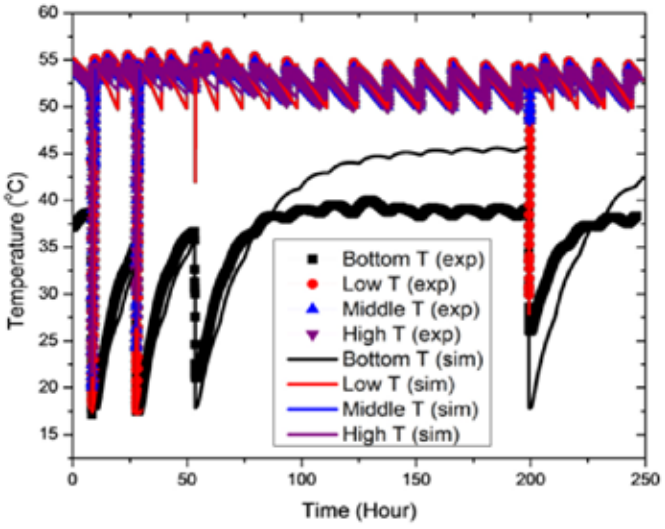
Modeling of Distributed Energy Resources in the Smart Grid

Shuai Lu

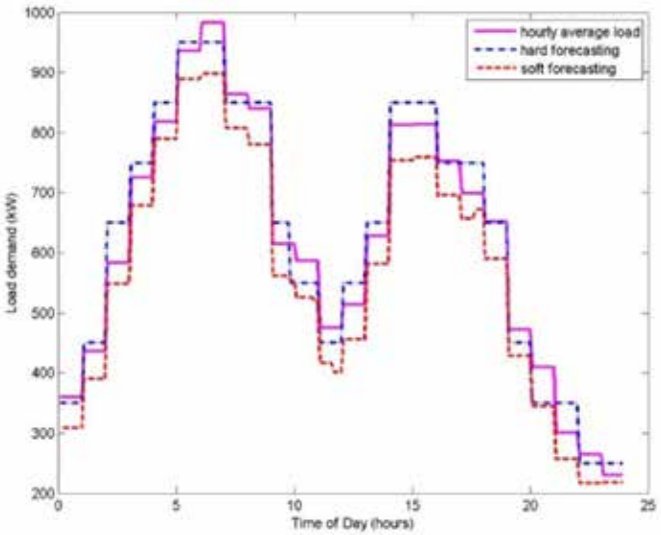
This project developed the aggregated models for demand response (DR) resources for the design, operation, and planning studies of the power system, removing the technical barrier of including DR in these studies.

The electric grid is seeing increasingly higher penetration of distributed energy resources as a result of smart grid developments. The smart grid will feature several distributed energy resources such as PV systems, EV batteries, and demand response from loads (e.g., A/C devices and electric water heaters) that can respond to system control signals and events to improve operation reliability and efficiency. Modeling these resources is important to the design of relevant technologies and to the planning and operational studies of electric power systems. Significant effort has been made in the development of models for distributed PV systems and storage devices, but lacking are models for DR resources, even though it is impractical to model them in detail in the simulation of interconnected transmission systems.

Through this project, we are focused on developing aggregated models that represent the performance of the entire population of distributed devices. Day-ahead prediction of DR availability, dynamic behavior of DR during deployment on the operation day, and the approach to determine the model parameters for a population of DR resources are three major topics investigated in this research. Through these developments, we will provide the essential elements for DR planning and operation tools, removing the barrier for large-scale implementation of DR programs and therefore improving power grid reliability and efficiency.



Temperature curves from detailed electric water heater model compared with field measurements.



A/C load prediction from DBN model compared with "actual" load used in validation.

Dynamic Bayesian Network (DBN) and state transition (ST) models are the two approaches initiated at the start of the project to derive the aggregated models for DR resources. DBN uses training data obtained from simulation of detailed models or from historical DR data. The trained DBN can help forecast the availability of resources and estimate the distribution of device parameters. The ST model aggregates the DR devices into a few states of temperature, and significantly reduces the time needed to simulate the dynamic behavior of DR.

In FY 2012, we made significant progress with the following activities: aggregated A/C and electric water heater modeling, electric water heater field experiment for detailed model validation, sensitivity study and uncertainty quantification on the parameters of the A/C model, characterization of A/C demand response performance, dynamic Bayesian network model for A/C load forecast, and multi-scale modeling of transmission systems with demand response. In addition, the calibrated electric water heater model is to be integrated with PNNL's widely used distribution system analysis software GridLAB-D.

We continued development of aggregated electric water heater model in FY 2013 based on the previous year's results, developed calibration techniques that help determine the aggregated model parameters of A/C resource, and refined the DBN model for the prediction of DR availability day ahead for A/C devices and electric water heaters. These techniques establish the foundation for an operation and planning tool that can help build models for both DR resources and predict the DR load shape from minutes to days.

PN11053/2375

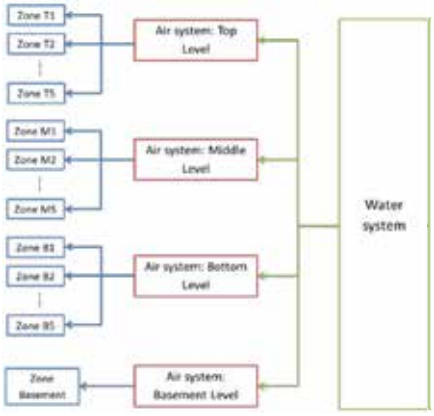
Multi-Zone Dynamic Commercial Load Modeling

Yu Zhang

Power system simulation and analysis of so-called "smart grid" systems must consider the short-term power consumption of loads, often the least understood part of an electric power system. Load modeling is an essential component of utility planning and operations, especially for commercial buildings that use more than one-third of the nation's electricity.

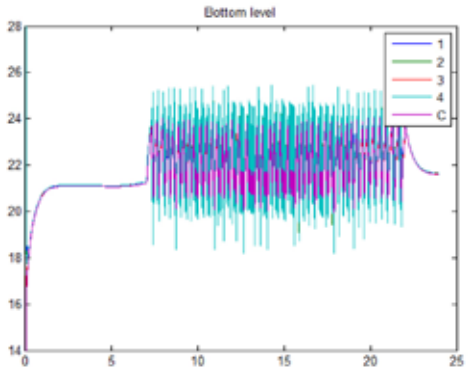
Significant progress has been made for single-family residential loads using the equivalent thermal parameter (ETP) model developed at PNNL. However, for the commercial buildings, there is no existing load model that is suitable for use in dynamic power system simulations. Some tools like EnergyPlus™, HVACSIM+, and TRNSYS require large amount of detailed characteristics data as inputs to estimate the heating/cooling/ventilation end-use profiles. Others such as DOE-2 Building Loads Analysis and System Thermodynamics (BLAST) do not support sub-hourly time steps. Moreover, the computational resources required are a barrier to interconnection-scale and even community-scale simulations, which can require thousands of concurrent building models with the associated distributions of key parameters.

Commercial buildings come in a wide variety of building sizes, vintages, usage, construction and locations. Depending on these factors, they are equipped with different air-conditioning systems that may range from 3-ton packaged units to thousands of tons built-up air handling systems served by a central plant. The varieties of mechanical and control technologies behind each system have different impacts on the grid. To reduce these complexities and capture the building load impacts on grid simulation, we are developing a set of modules for the commercial building thermal and AC systems specifically designed for power simulations. The structure will allow modification and reconfiguration in the future to support others common system types that serve different types of commercial buildings.



Whole building and HVAC system configuration.

We developed the most significant modules for a multi-zone large office constant air volume (CAV) system. It contains a multi- thermal zone large office served by a centralized CAV fan on the air side and constant speed water-cooled chiller on the water side. For the thermal zone of the large office, there are top, middle, bottom, and basement levels, with five zones on the last three levels and one zone for the basement. Zone air nodes and the mass nodes of the walls, floors, and ceilings are considered. A 54-order ETP multi-zone thermal model for a large office envelop is derived.



Zone air temperatures at bottom level.

The HVAC system includes both air and water sides. For the air system, we modeled an outdoor air box (including the fresh air intake, exhaust air release, mixed air delivery and economizer performance), cooling coil component, fan, and zone reheat. Within each zone, the supply air is mixed with room air, and thermal interactions with solar radiation ambient air temperature, wall, floor and ceiling masses, internal loads, and adjacent zones are considered. The water system contains a chiller, a constant speed chilled water pump, cooling tower, and a constant speed condensing water pump.

The modeling and control for the multi-zone dynamic large office and its HVAC system is implemented and simulated on MATLAB platform. The framework of the modules is structured in GridLAB-D. Numerical solver and controls will be deployed in the future. The dynamic model requires less input data, uses less computation resources, supports secondly-hourly dynamic simulations, and provides active and reactive power estimates for aggregate populations of buildings. We can simulate the zone air temperatures and other temperatures, mimic heat transfer process, capture power and energy changes with outdoor temperature change. These will enhance the accuracy of important power system studies, such as power flow analysis and demand response. They will also enhance the composite load data for FIDVR analysis required by WECC.

PN13047/2528

Numerical Simulator for the Utilization and Storage of CO_2 in Natural Gas and Petroleum Reservoirs

Mark D. White

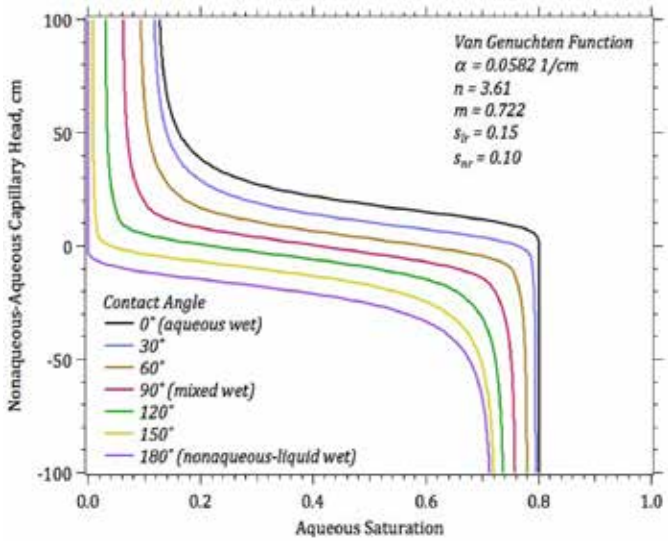
We are developing a new version of the STOMP simulator for investigating oil production and CO_2 sequestration using a combination of CO_2 and brine injection for conventional and unconventional enhanced-oil-recovery (EOR) techniques.

Proprietary and commercial numerical simulators help guide decisions made daily in the oil and gas industry. Scientific investigations concerned with the storage and utilization of CO_2 in geologic reservoirs benefit from numerical simulation when the mathematical descriptions of the subsurface processes are accessible. This project supports the development of a new open source numerical simulator, STOMP-EOR, for the carbon capture, utilization, and storage (CCUS) scientific community. Scientists and engineers on the Southwest Regional Partnership on Carbon Sequestration (SWP) will be applying STOMP-EOR over the next 5 years as one of the analytical tools for understanding the subsurface processes during the commercial-scale injections CO_2 for enhanced oil recovery and storage at the Farnsworth Unit of the Anadarko Basin, TX.

The new STOMP-EOR simulator will have capabilities for modeling non-isothermal subsurface systems involving water, petroleum components, CH_4, CO_2, and salt where three fluid phases are possible: aqueous, gas, and nonaqueous-liquid. The appearance of fluid phases will depend on the component concentrations, temperature, and pressure. The number of total components will be flexible; however, most EOR simulation applications will limit the number of petroleum components to reduce the required computational effort. As with other recent versions of STOMP, geochemical simulation capabilities will be incorporated into STOMP-EOR with the ECKEChem module, and geomechanical simulation capabilities will be incorporated via the EPMech module. Inclusion of these modules will provide STOMP-EOR with the fully complement of thermal-hydrological-geomechanical-geochemical (THMC) capabilities.

The multicomponent flexibility of the simulator will be accomplished using cubic equations of state, which allow for the prediction of phase occurrences and compositions (i.e., equilibria). The critical components of this project will be the development of an efficient set of routines for computing phase equilibria, and numerical solution algorithms for phase transitions. The equilibria and other constitutive equations make up the collection of routines and algorithms of the equation of state (EOS). Similar cubic EOS routines have been previously developed within the STOMP suite of simulators, which will greatly reduce the development effort. Numerical efficiency of multifluid simulators is strongly dependent convergence during phase transitions. Considerable effort is dedicated to the development of phase transition algorithms in the STOMP suite; these approaches will be adapted for the primary variable switching scheme and phase transitions in STOMP-EOR.

DOE's shift in emphasis from deep saline formations to partially depleted oil and gas reservoirs as target geologic sites for sequestering CO_2 and other greenhouse gases (GHGs) was the impetus for this project, which was a mid-FY 2013 start. Without the existing STOMP simulator framework and associated software, this code development project would not have been attempted. To realize a preliminary release of the STOMP-EOR by the close of the fiscal year, five critical code elements were required: 1) flexible EOS for petroleum reservoir fluids, GHGs, and water; 2) numerically efficient relative permeability-capillary pressure-saturation functions for three-phase mixed wettability; 3) numerical solution scheme with primary variable switching for any combination of aqueous, gas, and nonaqueous liquid phases and liquid wettability; 4) flash calculations routines for initial conditions, boundary conditions, and sources; and 5) integrated geomechanical capabilities.



Mixed-wetting capillary head vs. saturation function with wettability defined via the contact angle using the van Genuchten formulation as the base functional form.

PN13075/2556

At the core of all petroleum reservoir simulators are EOS routines, which are generally the most exercised calculations yielding phase conditions, phase compositions, and properties. The STOMP-EOR EOS was created using the standard cubic equation formulations but emphasized a balance between numerical efficiency and accuracy. The EOS allows for user specified components and cubic equation formulations. A novel capillary pressure vs. saturation (s-P) for mixed wettability systems was developed for STOMP-EOR, after reviewing the literature for approaches for modeling mixed wet systems. The new scheme uses conventional s-P functional forms and allows for fluid entrapment and extensions below residual saturations.

STOMP-EOR solves a system of governing conservation equations for energy and every user-defined component for each computational grid cell that are strongly nonlinear. Contributing to the nonlinear nature of the problem is the possibility for phases to appear and disappear. The primary variable switching scheme developed for STOMP-EOR provides for any combination of phase occurrences with a direct calculation path from the primary to all secondary variables. A unified suite of flash calculation routines were developed applicable for specifying and computing boundary conditions, initial conditions, and sources. This new approach for STOMP will provide the user with full input flexibility and eliminate differences between input specifications.

A specific effort was made to validate the newly developed geomechanical capabilities implemented in the STOMP suite of simulators. This capability provides STOMP-EOR with fully integrated and open source geomechanical capabilities, eliminating the need to execute commercial software to complete the geomechanical calculations.

STOMP-EOR will not become a mature, recognized open-source numerical simulator for petroleum reservoirs for several years, but it is ready for debugging and the validation stage of its development cycle. The U.S. DOE NETL under the SWP Phase III project led by New Mexico Institute of Mining and Technology will support these next stages of the development.

Operations and Planning Fusion

Yousu Chen

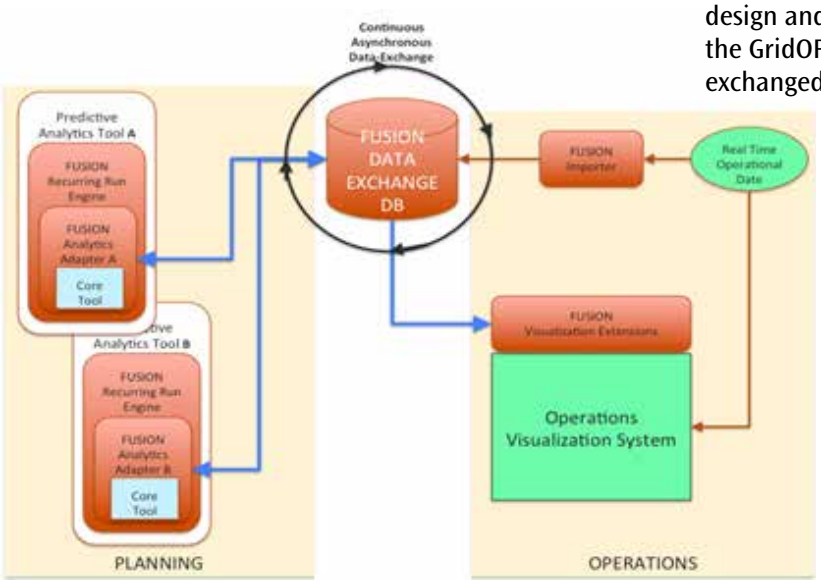
This project aims to advance decision support for the next generation of power system operation by better integrating the analysis results from planning into real-time operation and improving planning functions with fresher and more accurate information from operation.

The U.S. power system is accommodating new intermittent power generation technologies, controlling paradigms, and supporting the deployment of automation technology at all levels of the system. With this dynamic behavior brings the pressures of incorporating the predictive capability from planning in order to enhance today's power grid operation. These pressures are manifesting themselves in a smoother, better integrated interplay between the functional roles of planning and operations.

The objective of this project is to provide an initial set of tools that enable a bi-directional exchange of data between operations and planning environments based upon existing software integration and messaging framework developed under GridOPTICS™. In particular, the tools should allow the results of traditionally operational planning studies to be reflected in relevant real-time tools available to the system operator and support initialization of the operations planning analysis tools with current data from the operational system. By understanding a set of use case scenarios that drive the interplay between operations and planning, the project will articulate a set of common tools to help bridge these two worlds. While this objective can service a large number of operational scenarios, the project is structured to identify a core set of capability that can be used to demonstrate the bidirectional flow of information in a small subset of scenarios. The outcome of this project is the capability of alleviating the pain of integrating operations and planning while respecting the natural differences between their domains and design centers.

In FY 2013, we focused on selecting the candidate planning-oriented tools to support one or more use cases that demonstrate the benefit of operation and planning fusion. The project team worked closely to identify existing tools, understand the advantages and disadvantages of these tools, and how they are connected to the fusion project. These cases emphasized the use of predictive analytics for providing the statistical based information in a meaningful way to operators. A set of detailed documents were completed, including documentation describing each candidate tool, a set of illustrative use cases involving each of the candidate tools, and a summary of the evaluation process that was used to select a use case and tools for actual demonstrative implementation. These documents will provide a solid ground for moving the project forward.

Our plans for FY 2014 will continue our efforts of system design and identify the requirements for the extensions to the GridOPTICS framework and to support the information exchanged between planning and operation. This work will provide guidance to develop the design specification for the prediction analytics tools interface and GridOPTICs extensions, which are expected to adapt the information models in the analytic tools with those used in the operations environment. The extensions to existing analytics tools will be designed and developed based on GridOPTICS, but the operator interface area is expected to take more time to develop. We will engage actual operators or those familiar with the operating environment to help us identify reasonable approaches for designing operator interfaces. The selected use cases will be developed to discover requirements and to help prepare mock-ups of ideas from cognitive/decision support researchers.



General architecture of how operation and planning components interact in the fusion project.

PN13091/2572

Optimal Sizing Tool for Energy Storage in Grid Applications

Chunlian Jin

We are developing an optimal sizing tool for energy storage in grid applications to maximize the life cycle energy storage profit at the selected location.

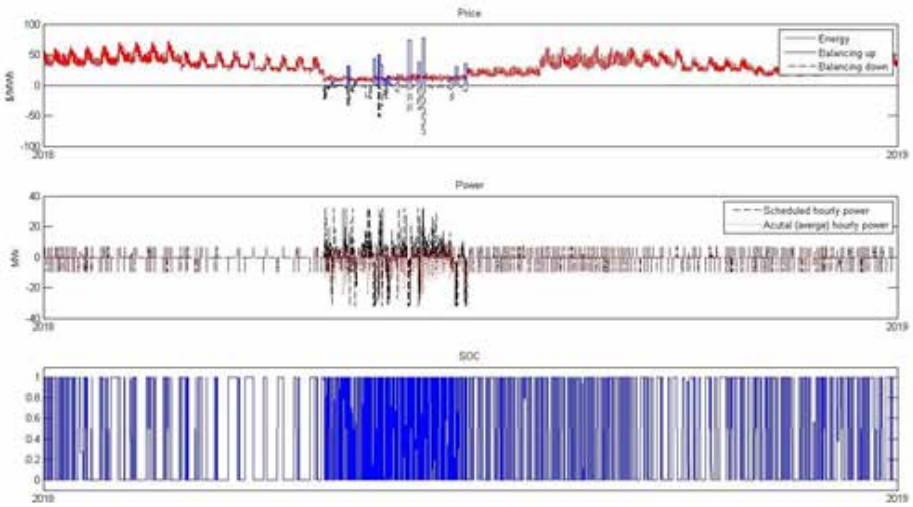
Variable generation (VG) plants are required to limit their power ramp rate at the point of interconnection in some systems. The plants are charged with high penalties if the actual generation deviates significantly from its forecast. PNNL's interactions with utility representatives and storage developers indicated the lack of engineering and economic tools to determine the optimal size of energy storage for a specific application. Based on our research, none of the tools or models reviewed provides features that seek optimal sizing options within a transmission or distribution system context. For a wide-scale deployment of energy storage or even for an arrangement beyond the pilot stage, both utilities and VG plant owners require analytic tools to assess the optimal size of energy storage to meet their system need with the consideration of cost recovery.

For the sizing of energy storage, the main challenges are storage cycle life, bundled grid services, energy constraints or desired state of charge range of energy storage units, and operation strategy. In this project, a novel optimization problem is being formulated to minimize the total cost, including capital and operating financial needs of the energy storage units to meet system requirements for the study horizon (30 to 40 years). The final product is a graphical user interface (GUI) software for which potential users are utility companies and storage vendors. The ultimate tool created will allow trade-offs among different grid services and determine the optimal size of energy storage needed.

In FY 2013, we developed a mathematical formulation that captures the cost and performance characteristics of energy storage, implemented the optimization in Matlab, developed a use case for Puget Sound Energy system, and tested and solved the use case by using an open source solver GNU Linear Programming Kit (GLPK) package. To understand the exact needs of potential clients of the sizing tool, the project team performed a survey. After designing the questionnaire, the team spoke with over 10 potential energy clients. Most survey participants showed a great interest in the sizing tool and are willing to provide data such as performance characteristics of different types of energy storage technologies, utility system data, and price information. Additionally, most of these companies would like to test the tool when it is available.

For example, potential value streams for the Bainbridge substation in the Puget Sound Energy system are energy arbitrage, balancing service, capacity value, outage mitigation, and substation upgrade deferral. To model these value streams, we collected the multiple data streams, including minute-by-minute system load and wind generation, 15-minute SCADA load at the storage location, historical power outages and forecast uncertainties, simulated energy price, balancing service price, and cost of alternative solutions to mitigate outage and transformer overloading. The historical outage information was processed and cost out based on outage cost information in a separate report.

The optimization problem has been formulated in a standard format so that commercial and open source mixed integer programming solvers such as IBM CPLEX and the open source GLPK package can be used. The nonlinear constraints were converted into linear constraints with integer variables using big M method. Because the outage constraints introduce many binary variables, the optimization problem takes a very long time to solve. The sizing



Top: Hourly price for energy, balancing up and down services for the study year 2018 in PSE; Middle: Energy arbitrage optimal schedule vs. the final optimal schedule of the energy storage system with the consideration of five potential value streams; Bottom: The state of charge of the energy storage system.

PN13096/2577

results have been obtained for the Bainbridge substation, which is 32 MW and 8 MWh with a charge/discharge rate of 4C. The tool does not differentiate the cost with charge/discharge rate at this time, but that will be considered in FY 2014.

A main finding in FY 2013 is that the size of energy storage is mainly determined by the requirements of local services such as distribution deferral and outage mitigation at a distribution location. Our work during this past year also revealed the complexity of the problem. With consideration of one technology and simplified representation of outage mitigation, the simulation time is already over 48 hours.

In FY 2014, different types of energy storage such as flywheel, flow battery, Li-ion battery, NaS battery, pumped storage hydro, and compressed air energy storage will be modeled. Therefore, the sizing tool will allow competition among different energy storage technologies. Further, different market mechanisms such as pay-by-performance will be modeled to make the tool applicable for power systems in different power markets. A GUI will also be designed and developed to facilitate the use of the tool.

Recyclable Methane Hydrate Inhibitors

David J. Heldebrant

We are creating a recyclable methane hydrate inhibitor with benefits that include improved economics, increased safety, and reduced environmental impact, all while providing comparable efficiency of separation compared with conventional hydrate inhibitors.

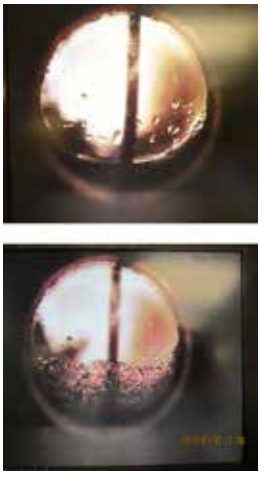
Hydrates are a significant problem in natural gas production and transportation that cause plugging of processing equipment and transmission pipelines. Thermodynamic inhibitors are effective but need to be added in significant quantities to prevent hydrate formation that potentially add substantial cost and have large storage requirements. Methanol is the most commonly used inhibitor because it is widely available and cheap, which is advantageous because the large volume requirements needed. However, methanol is challenging to work with, especially on offshore rigs because of large storage volume requirements, logistics of transportation, and high flammability and toxicity that requires responsible environmental disposal or recovery.

With these options, it is desirable for the industry to adopt a process that could regenerate inhibitors on site and find a method that could be recycled and recovered. Conventional thermodynamic inhibitors are too volatile; thus, we chose switchable electrolytes and their comparable performance to conventional inhibitors with enhanced safety and low environmental impact. A recyclable methane hydrate inhibitor based on CO_2-activated switchable electrolytes is being developed to deliver hydrate "inhibitor on demand." Such a process would be ideal from a logistic, safety, environmental, and potentially cost standpoint.

We strived to use data gathered from past studies and develop a working method for gas processing. To measure the hydrate curve depression, we designed an experimental method to observe the inhibition of the hydrate by detecting the water phase transitions under a series of pressures and temperatures common for natural gas processing. After design and preparation, a series of experiments were performed to validate our custom-built equilibrium measurement cell and validate the testing method. Specifically, methane and water mixtures were subjected to cooling/heating cycles. Cell temperature and pressure were recorded and compared with known data of methane-water system. Cooling and heating was completed at controlled ramp rate at the external circulator bath. Temperature ramp rates were 0.05–0.17°C/min for heating and cooling between –10°C and 20°C.

During these experiments, we observed the hydrate formation zone via changes in temperature and pressure in the cell in addition to visual confirmation. Video monitoring of the cell content revealed formation of solids, which was concurrent with observed increase in the pressure reduction rate (i.e., the formation of methane hydrate). Several freezing/thawing cycles were performed under a series of pressures and temperatures common for gas processing. From these curves, we constructed hydrate P&T equilibria curves at different loadings of inhibitor with and without CO_2.

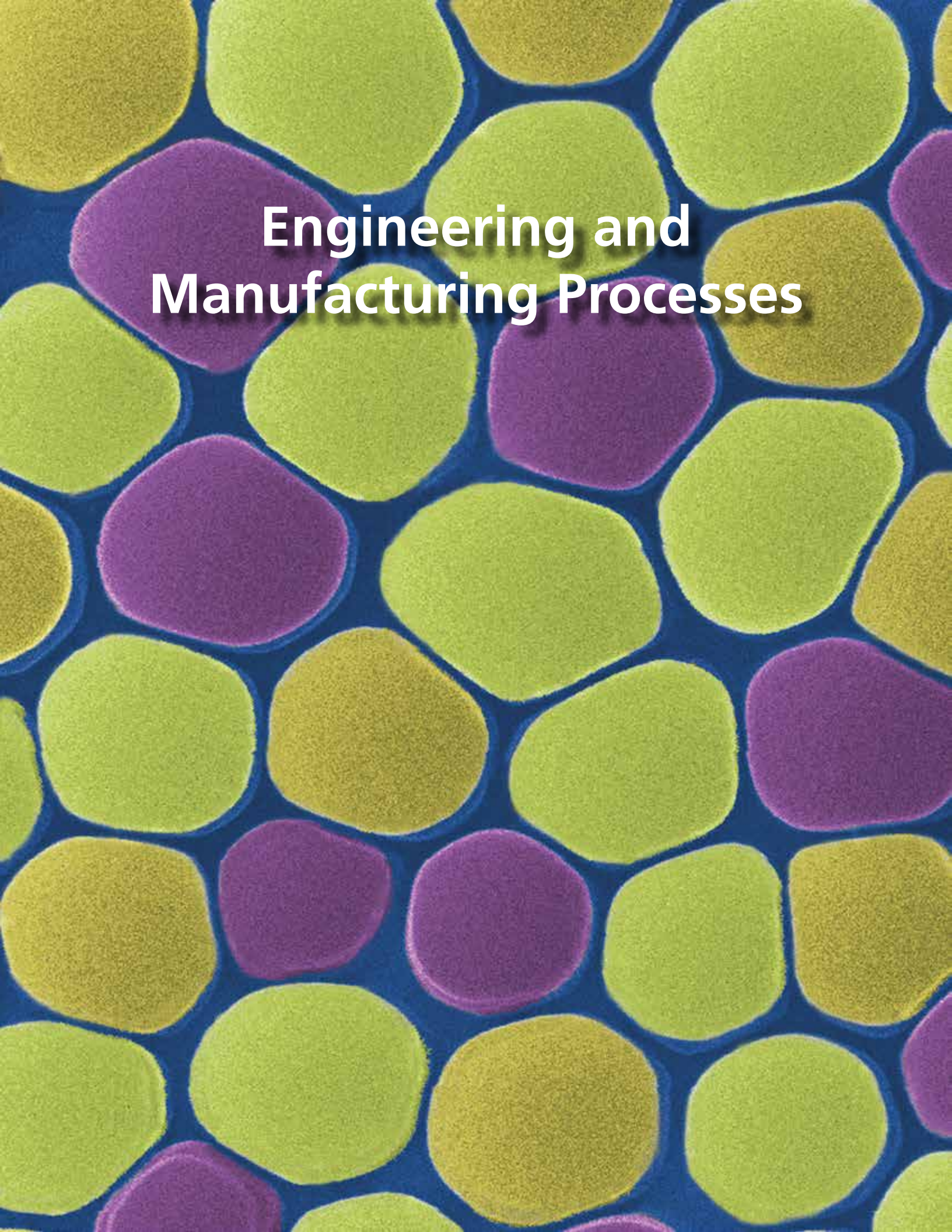
Looking at the produced data curves, it is clear that the inhibitors perform as conventional inhibitors when they have been CO_2-activated. On average, at 5-wt%, the inhibitors showed a 0.5–1.5°C destabilization in the hydrate curves. In the unactivated state, the inhibitors show some destabilization of the hydrate curves near 0.5°C, but CO_2 activation decreases hydrate onset (on average) by an additional 0.5–1°C. At 10 wt.%, the inhibitors show up to 3.5°C depression, confirming the effectiveness scales with loading. For comparison, inhibitors in the CO_2-activated state perform at the same rate on a mol % loading as methanol and sodium chloride, indicating that they work as conventional hydrate inhibitors, albeit activated when in contact with CO_2 in the gas stream. This is key, as the strength of the inhibitors is needed to validate process efficiency.



Visualization of the hydrate formation using a sapphire-windowed autoclave.

Our results indicate that CO_2-activated hydrate inhibitors can function as conventional inhibitors with comparable loadings and effectiveness of inhibition. The next step is to construct a viable process configuration and project sizing and efficiency to validate safety and logistic improvements over conventional systems. ASPEN plus analysis of the process are still being performed and will be updated when complete. As the result of our work, an invited manuscript is being drafted for a special CO_2 activated materials themed issue of *Green Materials*, and a patent will be filed upon completion of the process modeling.

PN13049/2530



Engineering and Manufacturing Processes

Manufacturable Gel-based Membranes for Gas Separations

David W. Gotthold

We used a thin film polymer deposition capability previously developed at PNNL in combination with ionic liquid expertise to establish and demonstrate a thin membrane suitable for dehumidification applications.

Separation membranes are a $3.8 billion per year industry dominated by polymer membranes for water filtration with other applications in gas and chemical processing, food preparation, and pharmaceutical and medical uses. Liquid membranes have shown promise for advanced separations because they can have much higher transport rates than solid membranes coupled with solubility selectivity. In particular, room temperature ionic liquids (RTILs) are of interest because they have both a very low vapor pressure and the potential for enhanced catalytic behavior.

The goal of this project was to develop a scalable deposition process to integrate a thin (between 0.1 and 10 µm) semi-solid RTIL with a flexible porous structural support suitable mass production. Current work in the field of RTIL-based membranes has focused on relatively thick films (typically above 20 µm), but the ability to manufacture even more thin films is critical to achieve industrially useful permeation rates. Previous research has demonstrated that gelled RTIL membranes could be both performance and cost competitive for CO_2 separation if the films could be made 1 µm thick rather than between 10 and 20 µm. This research leverages ongoing work at PNNL on membrane development, manufacturing, and RTIL development.

For membrane application, ionic liquids require low viscosities to allow rapid diffusion through the liquid and hydrophobicity to prevent water accumulation inside the membrane. There are numerous ionic liquids that meet both metrics, as several anions have been studied and were considered for the thin-film gelled ionic liquid membranes. Our initial work used published ionic liquids to refine the fabrication techniques and move to other advanced ionic liquids to refine membrane fabrication by adjusting key properties such as ionic liquid pore volume, contact angle, and surface tension. The second phase was transferring the selected RTIL materials to the monomer deposition system available at the Research Technology Laboratory at PNNL to demonstrate deposition on a manufacturing scale.

Specifically, our initial approach was to use vacuum spray deposition to deposit a thin layer of RTIL (1-Butyl-3-methylimidazolium bis(trifluoromethylsulfonyl)imide) on a polymer substrate. In addition to the vacuum spray deposition of liquids on a membrane support, we discovered that techniques based on gelation of the ionic liquids after spray deposition were possible. The addition of small quantities of 12-hydroxystearic acid as a gelating agent enables a facile transition of this ionic liquid to gel at temperatures below approximately 70°C. This environment makes it possible to use a spray coating of the ionic liquid with its subsequent gellification to enable improved control over both the resulting film thickness and pore penetration depth.

To make sure that the liquid RTIL formed a continuous layer, it was crucial that the substance wet the surface of the substrate, which requires that the polymer be hydrophilic. However, we found that with a standard hydrophilic membrane, a liquid (or its gelated form) quickly penetrated into the pores, forming a thick structure where the effective membrane thickness was equivalent to the thickness of the support polymer (typically 90 µm). To prevent this penetration, we developed an O_2/Ar plasma treatment process that converts just the surface layer of a hydrophobic porous polypropylene substrate to be hydrophilic. Because the plasma process is line-of-sight, only the very top surface is converted, while the bulk of the support remains hydrophobic. This multi-functional structure then enables wetting of the surface without the liquid penetrating into the bulk of the support, enabling the formation of thin (approximately 5 µm) liquid layers. This process worked so well that we were able to avoid the need for vacuum spray coating of the RTIL and use only simple solution coating processes.

To validate the performance, we completed mixed-gas permeation testing using a combination of CO_2, N_2, O_2 and H_2O to simulate wet flue gas. We observed the CO_2 flux increasing from 1.3×10^{-3} mL•cm-2•s^{-1} to 1.6×10^{-2} mL•cm-2•s^{-1}, a factor of 12.3 times. This correlated almost exactly with the reduction of thickness from 89 µm (the polypropylene support) to 7 µm (estimated by weight) for a total thickness reduction of 12.7 times. This was achieved while maintaining the selectivity at approximately 25:1 for CO_2/N_2.

In short, we successfully demonstrated a manufacturable process that enables the formation of extremely thin liquid membranes on mechanical support structures.

PN13046/2527

Materials Science
and Technology

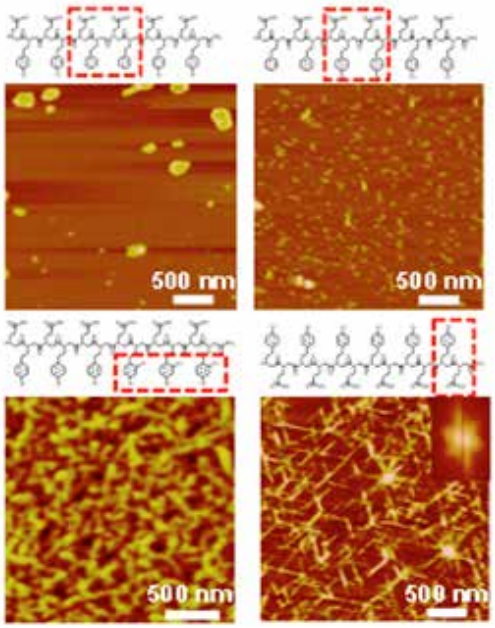
Design and Synthesis of Peptoid-based Functional Materials

J.J. De Yoreo

We are developing a new class of polymers that mimic the basic architecture of proteins and naturally assemble into functioning materials, with an initial target of membrane-based separation for biofuel production and CO_2 capture.

The complex functions seen in biological systems are in large part enabled by the action of proteins, which are natural polymers that provide the molecular machinery of life. In contrast to manmade polymers, proteins perform sophisticated functions because their construction from 20 distinct amino acid monomers endows them with high information content, which in turn codes folding, assembly, and function. However, because protein structure arises through weak intra-protein H-bonds inherent to amino acid-based polymers and the abundance of natural proteases that cleave peptide backbones, proteins comprise poor building blocks for the construction of useful materials.

In this project, we will construct functional materials by designing self-assembling high information content polymers called peptoids, in which the peptide backbone is immune to protease digestion and weak H-bonds are replaced with stronger ionic and covalent bonds. In parallel, we will use in situ imaging and spectroscopy to follow the peptoid assembly into macroscopic materials. This work goes beyond previous research in the field by examining the synthesis of short peptoids with repeating units that either do not assemble or simply make homogeneous non-functioning solids. It represents the first attempt to observe peptoid assembly and derive the physical parameters that control the process. With this combination of synthesis and characterization, we seek to develop both a set of functioning membranes for water and CO_2 separation as well as a quantitative model that relates peptoid design to assembly. Developing the ability to create synthetic materials that exhibit the high level of function achieved by living organisms would be transformative in the fields of energy, information technology, medicine, and environmental stewardship.



AFM images of the final morphology of four different peptoids deposited on mica showing importance of sequence in controlling organization. The three-fold symmetry seen in the bottom images are due to the interaction between the peptoids and the underlying mica substrate.

To achieve our goal, we synthesized a suite of peptoids that we have purposely designed to link together into a mesh-like network. The peptoids have chemical groups that extend into the pores of the mesh to control what can pass through them. In addition, we developed a physical picture of the way in which the peptoids assemble into the mesh by using molecular imaging tools that allow us to visualize the process and measure the rates at which molecules add onto the growing mesh. Our specific project approaches and processes are described below.

Design and synthesis of network forming helical peptoids. Our first peptoid design is based on helical units that assemble through coil-coil interactions and create functional pores through the choice of side chain chemistry. To create the structural motifs that enforce helicity in peptoids, we designed and synthesized the amphiphilic peptoids using inexpensive, commercially available (S)-(-)-α-methylbenzylamine as the basic unit to generate a helical backbone.

In these peptoids, monomers in one-third of the peptoid backbone positions are anionic monomers N-(2-carboxyethyl)glycine (Nce) or cationic monomers N-(4-aminobutyl)glycine (Nab), and these two monomers are known for forming strong charge-assisted hydrogen bonds for stabilizing peptoid helical bundles. Monomer N-[2-(3,4-dimethoxyphenylethyl)] glycine (aminobutyl)glycine with metal-chelating capability was also introducted in the C-terminals of these two peptoids to stabilize helical bundles by mimicking metal-directed assembly of protein helices. The 6-aminohexanoic group was introduced in the N-terminals for adjusting the overall peptoid amphiphilicity and reaction activity for further peptoid assembly.

PN13086/2567

Design and synthesis of network forming non-helical peptoids. Our second peptide design utilizes a three-fold non-peptoid hub to organize non-helical peptoids into a pore-formine network. Complementary side-chains containing primary amine and carboxylic acid groups were introduced for formation of strong charge-assisted hydrogen bonds, and pepotids containing only carboxylic acid side chains or N-[2-(3,4-dimethoxyphenylethyl)]glycine (aminobutyl)glycine were introduced for metal-chelating capability. These functional side-chain groups are introduced to mimic the β-sheet like structures found in peptide- and protein-based materials. We will conjugate these non-helical pepotids with organic building blocks with C3-symmetry such as trimesic acid and use resultant C3-symmetric peptoid-conjugates for assembling porous materials.

Peptoid synthesis and purification. All peptoids were synthesized on a commercial Aapptec Apex 396 robotic synthesizer using a solid-phase submonomer method which consists of a two-step monomer addition cycle: acylation and displacement. First, bromoacetic acid activated in situ with N,N'-diisopropylcarbodiimide acylates a resin-bound secondary amine. Second, nucleophilic displacement of the bromide by a primary amine introduces the side chain. The two-step cycle is iterated until the desired chain length is reached.

All reactions were performed at room temperature. Peptoid chains were cleaved from the resin using 95% (v/v) trifluoroacetic acid (TFA) in water. Following cleavage, peptoids were dissolved in a mixture (v/v = 1:1) of water and acetonitrile for HPLC purification. All peptoids were purified by reverse-phase HPLC on a Vydac C18 column using a gradient of 5-95% acetonitrile in H_2O with 0.5% TFA. Finally, all final products were analyzed by analytical reverse-phase HPLC and electrospray mass spectrometry trap system or matrix-assisted laser desorption/ionization mass spectrometry.

Analysis of peptoid assembly. To investigate the impact of sequence variability on the assembly of peptoids into two-dimensional networks, we used atomic force microscopy (AFM) to examine the assembly of a set of previously synthesized peptoids deposited from solution onto mica substrates. The results are two-fold. First, peptoids that adopt ordered structures do so through a two-step process that begins with condensation into amorphous deposits before molecular re-organization brings about the emergence of order. Second, small variations in side chain chemistry lead to large-scale changes in the assembled architectures, with many peptoids remaining disordered. Based on the chemical structure of the peptoids that assemble into an ordered state, we hypothesize that a balance of peptoid-peptoid and peptoid substrate interactions is required to achieve order.

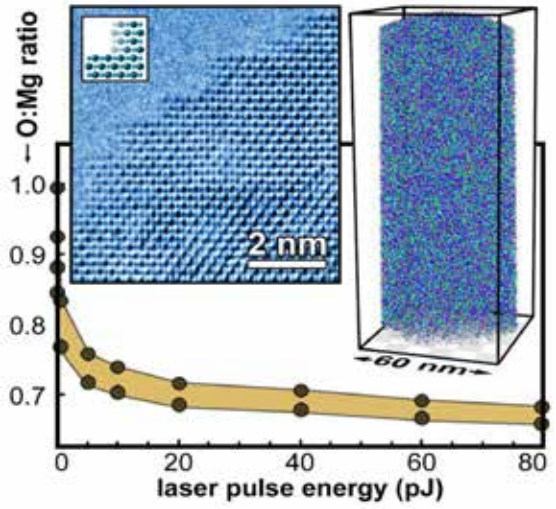
The results obtained here represent the first steps toward a capability to create peptoid-based pore networks for molecular separation. They illustrate a design concept that exploits the natural ability of three-fold elements to assemble into pore-forming networks as well as the ability of in situ AFM to follow the assembly process.

Developing Next-Generation Multimodal Chemical Imaging Capability by Combining STEM/APT/STXM/HIM

Suntharampillai Thevuthasan

This project will develop a common analysis platform for integrating multiple microscopies to characterize commercially viable catalyst materials with metal or metal-oxide nanoparticles supported on porous metal oxides and Li-related energy storage materials.

Understanding complex energy conversion and storage systems hinges on analyzing the structure and chemistry with sub-nanometer spatial and chemical resolution and ppm-level mass sensitivity. These requirements exceed the current capabilities of any individual technique, necessitating an integration of multiple tools and methodologies. To this end, we are combining chemical imaging capabilities by atom probe tomography (APT) and high resolution structural characterization with scanning transmission X-ray microscopy (STXM). In one integrated platform, these tools will preserve a sample registry across different techniques. Through this multi-modal approach, we aim to understand the nanoscale structure, composition, and chemical state of elements in Na and Li-ion battery cathode materials as a function of various synthesis conditions at different stages of electrochemical cycling. We aim not only to improve the fundamental scientific understanding of starting and degradation of microstructure with electrochemical cycling but also to aid synthesis conditions and develop protective coatings and other microstructure modification methods. Collectively, these efforts will prevent battery capacity fading during long-term use through charge-discharge cycles.



Understanding laser-oxide interaction during APT analysis by direct coupling of aberration corrected TEM and APT analysis.

Initially, aberration corrected transmission electron microscopy (TEM) imaging and atom probe tomography (APT) analysis were correlated to understand the laser-oxide interaction during APT analysis of MgO. The result was published in *Journal of Physical Chemistry Letters*. Subsequently, the coupling of aberration corrected scanning TEM (STEM) imaging with APT analysis and computational modeling provided a deeper understanding of possible aberrations during analysis of metal-oxide composites, the results of which are currently in manuscript preparation.

With the Advanced Light Source beam time (lines 11.0.2 and 5.3.2), we conducted STXM imaging and XANES of pristine and cycled Li-ion battery cathode materials explicitly layered $Li_{1.2}Ni_{0.2}Mn_{0.6}O_2$ and $Li_{1.5}Mn_{0.5}Ni_{0.5}O_2$ spinel materials for Mn and Ni L as well as O K edges. Subsequently, we collected STXM images and XANES spectra from a number of pristine and cycled Na-ion battery cathode materials synthesized differently, which led to different starting capacities as well as a differing extent of capacity decay during battery testing. Specifically, $NaMn_{0.33}Ni_{0.33}Co_{0.33}O_2$ and $Na_{0.44}MnO_2$ materials were analyzed in STXM and XANES for Mn, Ni, Co L edges, and Na, O K edge.

Parallel efforts were made to develop sample preparation methodologies for fabricating atom probe needle samples from nanoparticles of Li-ion battery cathode materials on both stand-alone APT sample holders and a STXM-TEM-APT compatible multimodal sample holder. The APT analysis results from $Li_{1.2}Ni_{0.2}Mn_{0.6}O_2$ and $Li_{1.5}Mn_{0.5}Ni_{0.5}O_2$ spinel materials provided the evidence of phase separation of Mn and Ni for the layered material, but not for the spinel, within which a uniform distribution of the elements were observed. In addition, quantifiable evidence for Li partitioning toward Mn rich regions in layered $Li_{1.2}Ni_{0.2}Mn_{0.6}O_2$ was obtained by APT analysis that was absent in $Li_{1.5}Mn_{0.5}Ni_{0.5}O_2$ spinel materials.

A parallel effort was focused on developing a level set computational modeling capabilities for understanding the field evaporation of heterogeneous materials during APT analysis. The level set framework developed in FY 2013 can simulate the evaporation of layered heterogeneous materials with alternating layers of high and low field material in addition to composite materials with a precipitate with a certain evaporation field embedded in a matrix with different evaporation field. These models are also validated by our experimental results of direct combination of TEM and APT analysis.

We will continue the expertise developed in FY 2013 for conducting STXM, XANES, TEM imaging, and diffraction along with APT analysis of the same needle sample using a multimodal holder. Specifically, we intend in FY 2014 to seek to understand the fundamental microstructural factors that are influencing the initial capacity and capacity decay during cycling of Li and Na ion battery cathode materials.

PN13033/2514

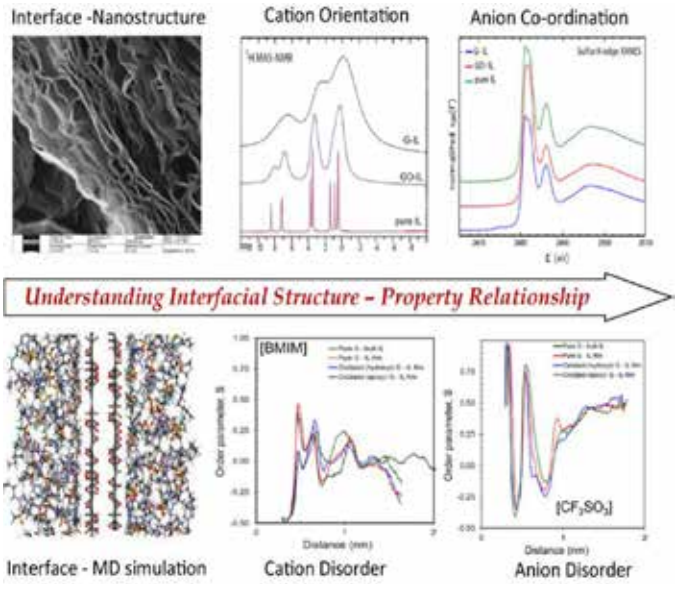
Development of Graphene/Ionic Liquid Hybrid Material for Ultracapacitors

Vijay Murugesan

This project is developing graphene-ionic liquid based designer interfaces to achieve high energy density ultracapacitors that can power modern technologies ranging from portable electronic devices to large-scale electric grids.

Molecular interaction at electrode-electrolyte interfaces is the basic science that drives all the energy storage/conversion devices. However, our predictive understanding about the impact of electrode surface chemistry and nature of electrolyte ions on their molecular level interfacial interactions is seriously limited. This impedes our ability to predict, control and design optimal interfaces. For example, formation of electric double layer capacitance (EDLC) which arises from interfacial interactions is the driving force of supercapacitor performance. Hence, by rationally designing the interfacial region we can achieve optimal EDLC and the much desired high energy density supercapacitors. This work focuses on designing an optimal interface between graphene (G) and ionic liquids (IL) by understanding their interfacial molecular structure.

Graphene is widely reported as suitable electrode material for high energy density lithium-ion battery and supercapacitor technology. However, fickle performances are reported for graphene based electrode materials. For example, the engineered surface modification of graphene such as nano-structural designs, doping and/or surface functionalization yields relatively higher performance. Evidently many research efforts are ongoing to design optimal graphene surface for supercapacitor applications. These fickle performances can be related to the purity and structural homogeneity of the graphene material which heavily depend on their surface properties. Therefore, to design the graphene based electrode materials rationally, we need clear understanding about the role of graphene surface defect groups and its effect on final device performances.



Consilience through combined multi-modal experimental and theoretical approaches provided the critical structure-property relationship of complex G-IL interfaces.

The major challenge in developing a G-IL nano-hybrid structure in FY 2012 was to prevent restacking of graphene sheets to allow intercalating the ILs in desired thickness between individual graphene layers. Such an intercalation required controlled interlayer spacing (i.e., IL thickness) and more favorable affinity between the ILs and graphene. To develop such an optimal interfacial region, it was essential to have a comprehensive understanding of the molecular structure and dynamics of these complex interfacial regions. To understand the effect of ion size and symmetry in the formation of EDLC, we chose different anions such as CF_3-SO_3^-, $(CF_3$-$SO_2)_2N$-, HSO_4^-, $AlCl_4^-$ and Br- with common imidazolium based cation known as 1-butyl, 3-methyl imidazolium (BMIM). We developed an industrial friendly synthesis method for preparing G-IL composite material using these ionic liquids. This simple solution-based method allowed us to successfully intercalate IL molecules between successive graphene layers with significantly reduced restacking phenomena. Additionally, this allowed us to utilize the enormous surface area of the graphene electrode materials.

In FY 2013, to gain comprehensive understanding of the molecular structure and dynamics, we carried out combined experimental and theoretical analysis of complex G-IL interfacial region. In particular, nuclear magnetic resonance (NMR), sulfur K-edge X-ray absorption study (XANES), X-ray photoelectron (XPS), and infrared (FTIR) spectroscopy analyses on this G-IL material were performed. These analytical results were then correlated with statistical and electronic structure-based computational analyses. As a first step, classical molecular dynamics (MD)-based modeling method was used to develop molecular structure at the interfacial analysis. However, this classical MD simulation offers a reliable result only when the chosen interatomic potentials closely represent real atomic interactions. To overcome these inherent limitations, feedback from experimental spectroscopic parameters such as ^1H MAS NMR spectra, sulfur K-edge XANES spectra, FT-IR spectra were used to benchmark the simulation. For this, the time-averaged spectroscopic parameters such as NMR chemical

PN12023/2424

shift, vibrational frequencies, and core-level spectra were calculated at the DFT level for the molecular models taken from snapshots of MD simulation trajectories. This interplay between experiment and computational approach led to reliable interfacial molecular structure.

Our uniquely combined theoretical and experimental approach led to fundamental molecular insight about role of surface functional defect groups on G-IL interfacial region, with the results summarized as follows. First, BMIM cations orients parallel to the graphene layer due to π–π stacking interaction and form a primary interfacial layer, which is subsequently capped by a layer of TfO anions. BMIM cations can interact with functional groups on the graphene surface just as they interact with regions of defect-free graphene; however, the disorder in BMIM orientation near electrode surfaces are greatly amplified by the presence of defect groups. Next, TfO anions are repelled by pure graphene surfaces due to π - orbital interactions, however hydroxyl defect groups can attract the TfO anions by offering hydrogen bonding environment. Finally, the distance between electrode and electrolyte's ion is controlled by the defect groups through protrusion forces. Because the EDLC is directly related to ion-electrode separation distance, the fickle performance is from the non-quantified surface defect groups in the graphene electrodes. It is therefore essential to appraise the surface defects to predict the device performance.

Molecular-level understanding gained in this project through combined theoretical and experimental analysis explained the fickle performance of graphene based electrodes widely reported in literature. In addition, G-IL hybrid nanostructure with optimal interlayer spacing is achieved owing to better understanding about interaction between IL molecules and oxygen containing functional defect groups on graphene surfaces. Results from this project thus far are disseminated in two publications, one additional manuscript, and two conference presentations.

As a next step, we will focus on testing these rationally designed G-IL nano-hybrid materials under various electrochemical charge/discharge cycling properties. Macroscopic observables such as specific capacitance and electrochemical cycling properties will be derived from optimized MD simulation and verified against the experimental observation of specific interfaces. This combined approach will elucidate the relationship between electrode-electrolyte interfacial structure and their macroscopic critical device properties. Specifically, we will acquire preliminary results to show the feasibility of XAS analysis to evaluate the extended interfacial structure. In addition, we will explore the market analysis and performance requirement for a successful ultracapacitor.

Development of Hard X-Ray Emission Spectroscopy Nanoprobe

Nancy J. Hess

We are developing an* in situ *probe of the electronic structure of battery materials during operation to gain fundamental knowledge of irreversible transformations that adversely impact battery performance.

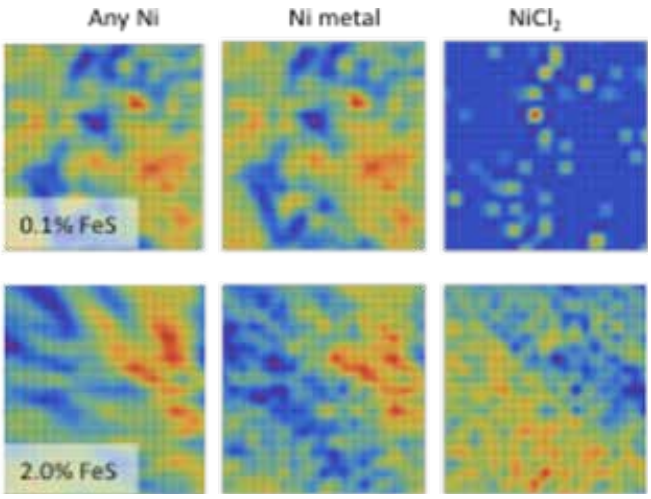
Battery lifetime and performance can be improved by elucidating mechanisms that lead to irreversible transformations and changing electronic structure. Much of the previous work on battery materials relies on *ex situ* measurements and less direct measurements of electronic structure with lower spatial resolution. Understanding the detailed electronic structure of multivalent ions is fundamental to many fields of chemistry and materials engineering, particularly in energy storage and materials. X-ray emission (XES) is complementary to X-ray absorption spectroscopy (XAS) as it provides element-specific measurement of oxidation state, bond length, and neighboring ligand identity. However, unlike XAS, XES is a "more local probe" of the electron-electron interactions, so it gives more detailed electronic structure information such as orbital splitting, spin state, magnetic properties, and bond character and is less influenced by structural disorder and thermal effects. Through the use of low-cost, element-specific XES spectrometers, we will construct a modular system capable of investigating the electronic structure of the transition metal, lanthanide, and light actinide elements within a variable temperature sample environment and with spatial resolution ultimately at the nanometer scale.

In collaboration with the University of Washington and Advanced Photon Source, we used the Ni K-beta mini-XES spectrometer developed by this project in FY 2012 to measure XES maps of regions of Na/$NiCl_2$ (ZEBRA) battery cathodes. The maps were obtained on discharged cathode samples from multi-cycled batteries fabricated with varying FeS loading from 0.1 to 2 wt %. Chemical maps of the Ni speciation were obtained with a 5 μ × 5 μ resolution across a 100 μ × 100 μ area by linear combination fitting of the XES spectrum at each point in the map to Ni and $NiCl_2$ standards. The high quality of the fitting, the standards, and our recently published XANES measurements in the *Journal of Power Sources* validate our ability to determine of the Ni to $NiCl_2$ ratios within a single XES map acquisition without the need for inclusion of other Ni species such as nickel oxides.

Preliminary analysis of these XES maps along with complementary SEM/EDX imaging indicate that the concentration of the FeS additive influences the spatial distribution of the Ni and $NiCl_2$ phases within the cathode. For example, at low FeS concentrations $NiCl_2$ is found in isolated regions within the cathode compared with Ni metal. At 2% FeS loading, $NiCl_2$ is much more homogenously mixed with the Ni metal regions. This experimental result may underlie the empirical observation of increased battery performance with the addition of 1 wt % FeS. Namely, enhanced mixing of Ni and $NiCl_2$ phases allows the electrochemical reaction $NiCl_2 + 2Na = Ni + 2NaCl$ to completion more readily because of the close proximity of $NiCl_2$ to Ni metal, the electron charge carrier network. Further analysis is currently in progress.

Our collaboration with the University of Washington has resulted in the design and fabrication of a new emission spectrometer as described in *Review of Scientific Instruments* that maximizes the focusing ability of X-ray polycapillaries. The polycapillary optic collects the emission from the sample over a wide solid angle and transmits an intense parallel beam to an analyzer crystal for energy discrimination. The high count rate lends the design to application in *in situ* studies.

The high operating temperature (250–300°C) and extreme sensitivity to moisture present major technical barriers to *in situ* studies of Na/$NiCl_2$ ZEBRA batteries. We therefore plan to extend our research project to other battery systems, and in particular the Na-Mn-O and Na-Mn-Ni-Co-O systems. The XES techniques we developed will provide complementary information to STXM investigations underway in other projects, both from a chemical and spatial perspective. We will also work to modify an existing *in situ* sample design from the Stanford Synchrotron Research Laboratory (SSRL) research team. The SSRL thin film battery is designed for transmission XAS experiments and is well suited for incorporation into the miniXS spectrometer.



XES maps showing the spatial distribution of Ni phases in 0.1 and 2.0 wt % FeS ZEBRA battery cathodes.

PN12024/2425

Directed Mesoscale Synthesis of Tunnel Structured Materials for Energy Applications

Eugene S. Ilton

The rational design of materials with properties optimized for energy applications is vital to a sustainable future. At its most fundamental level, theory to design materials are being used in this project with useful properties relevant to energy storage and conversion.

Explicitly called out in the area of mesoscale science as examples of designed and self-organized systems, the synthesis of epitaxial thin-films can be guided by theoretical predictions. Specific materials of interest are tunnel structured Mn oxides that are intrinsically mesoscale in spatial and compositional dimensions and are crucial in technological applications such as Li batteries, catalysis, fuel cells, electrochemical capacitors, sensors, and groundwater remediation. However, the complexity and poor quality of natural, hydrothermal, and electrochemically synthesized Mn oxides has hindered efforts to understand their fundamental structure-property relationships, thus preventing full technological potential. To address this issue, we are using molecular beam epitaxy (MBE) and pulsed laser deposition (PLD) to make high-quality MnO_2 single-crystal films for characterization and experimentation in conjunction with computational modeling. Most previous thin film work has been more art than science. By contrast, we are melding experiment with theory to provide a predictive capability that forms the basis for a rational design of films with unique properties.

In FY 2013, we successfully synthesized 1 × 1 tunnel structures with a novel compositional twist and closely integrated theory and experiment. Specifically, the successful growth of Mn oxide 1 × 1 tunnel structures was achieved by an insight that led to the novel approach of mixing some Ti into the lattice to stabilize MnO_2 in the rutile structure and IV oxidation state. Synthesizing pure β-MnO_2 thin films on TiO_2 substrates was challenging, initially resulting in Mn_2O_3 dominant films. Several potential reasons for synthetic difficulty were considered and evaluated by theory to help overcome them: 1) the marginal lattice match between the TiO_2 substrates and β-MnO_2 in combination with an accidental lattice match with Mn_2O_3; 2) *ab initio* thermodynamics calculations indicate that the experimental conditions are on the edge of stability regions for competing Mn(III) and Mn(II) phases; and 3) recent calorimetric work with the University of California-Davis indicates that the Mn(IV)O_2 stability field is strongly restricted to lower T as particles move into the nano-size regime.

Theoretical calculations addressing stability issues suggest that "protecting" the Mn in a TiO_2 matrix by co-deposition would be beneficial (i.e., stabilizing the desired rutile structure at the expanded lattice dimensions and destabilizing the competing Mn_2O_3 and Mn_3O_4 phases). These considerations also provide insight to issues encountered during early MBE experimental work by other groups using PLD to synthesis β-MnO_2. This approach met with initial success even though the resultant $Mn_xTi_{1-x}O_{2-\delta}$ films suggest some degree of oxygen vacancies. Single-crystal epitaxial smooth films of $Mn_{0.125}Ti0_{0.875}O_{2-\delta}$, $Mn_{0.25}Ti0_{0.75}O_{2-\delta}$, and $Mn_{0.50}Ti0_{0.50}O_{2-\delta}$ were deposited by PLD on TiO_2(110) and TiO_2(001) substrates. X-ray photoemission spectroscopy (XPS) records some Mn reduction which increases near the surface. XPS indicates that $Mn_xTi_{1-x}O_{2-\delta}/TiO_2$(110) films are Mn-rich near the surface while $Mn_xTi_{1-x}O_{2-\delta}/TiO_2$(001) films are Mn-rich near the interface. XRD indicates that the films are coherently strained to the substrate which may also influence the oxygen non-stoichiometry. Aberration corrected transmission electron microscopy corroborated the XPS and XRD results and indicates a potential Mn-dependent defect.

First principles DFT calculations were linked with experimental thermodynamics using the method of *ab initio* thermodynamics, which enables calculation of the stability of solid state phases, surfaces, and defects, as a function of temperature, pressure, and concentration of components. Again, it should be emphasized that the demonstration of how such calculations can explain and guide the synthesis of epitaxial films of strategic importance should help position PNNL in the area of mesoscale material design.

Calculations indicate that the experimental conditions are in a narrow region of stability for the desired pyrolusite structure (i.e., rutile) but close to the stability region for the undesired bixbyite (Mn_2O_3) and hausemannite (Mn_3O_4) phases. This is consistent with the historical difficulty in synthesizing pure MnO_2 pyrolusite. Expanding the MnO_2 pyrolusite lattice to fit the TiO_2 rutile lattice destabilized the structure by 0.19 eV per Mn atom. The computations show that adding Ti to make a mixed oxide, however, stabilizes Mn in the expanded rutile lattice as well as destabilizing the competing bixbyite and hausemannite structures. Further, calculations of the bulk phases indicate that for a

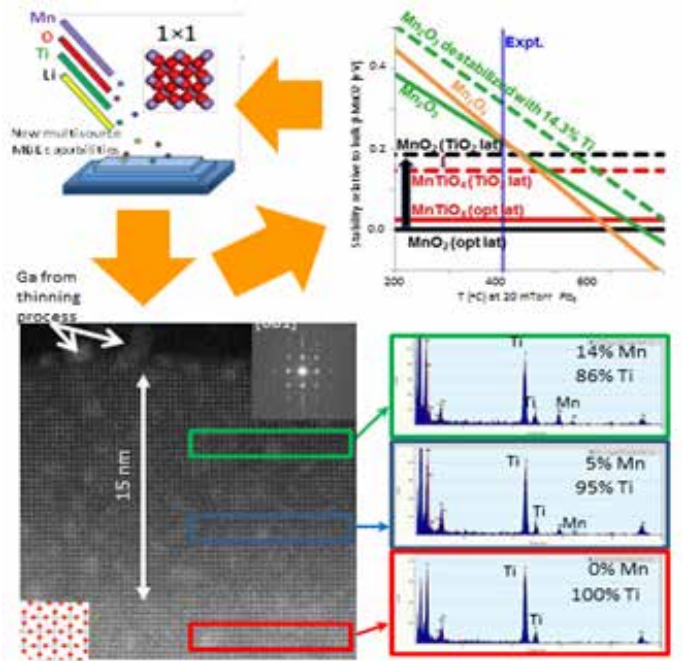
PN13004/2485

1:1 mixture of Ti:Mn the synthesis temperature could be increased by 50–75°C at PO_2 = 20 mTorr to help overcome kinetic barriers without entering the stability fields of undesired phases.

To facilitate the development of a structure-property framework, electronic properties associated with the mixed Mn:Ti oxide structure were calculated. The 1:1 doping of Mn into TiO_2 introduces states into the band gap, thus reducing the calculated gap from 1.84 to 0.48 eV. These calculations indicate that the mixed Mn-Ti oxides may exhibit excellent electronic properties for energy applications.

Interatomic potential parameters for simulating pure and lithiated Mn oxides have been derived and will allow for modeling mesoscale diffusion of Li in mixed Mn/Ti thin films. The new model represents an improvement over existing models in terms of both the accuracy with which the structures of Mn oxides are reproduced and the range of structures that can be modeled. The new model yields good agreement with experimental data on the structure of mixed Mn/Ti lithiated spinels, which gives us confidence that planned calculations for mixed Mn/Ti lithiated rutile structures will also yield accurate results.

For FY 2014, we plan to test the performance of these films as battery materials using a novel *in situ* TEM cell developed at PNNL, which allows for the real time observation of gradients in structure, Li content, and Mn valence states as a function of applied voltage. The results of these experiments will be compared to our predictions using a combination of MD and *ab initio* level theory, as described. A strong renewed interest in rutile structures has been spurred on by the relatively recent discovery that very high capacitances per gram can be reached by using nano-size materials. The mechanism responsible, however, is controversial and explanations range from intrinsic size effects on the energetics to the effect of size on access to fast diffusion pathways (i.e., the relatively unstable [001] surfaces). Such issues can be directly studied by our methods. We also plan to extend our predictive synthesis capability; thus far, the predictions are for bulk phases, but the phase stability regions will likely shift for surfaces and thin films. Consequently, calculations for the energies of surfaces and thin films will be performed. Stability boundaries established will be tested experimentally. One goal would



Top left: PLD or MBE manufacture of thin films; Top right: Ab initio thermodynamics, showing stability of relevant phases and the effect of lattice mismatch for MnO_2 vs. $MnTiO_4$ grown on TiO_2 substrate; Bottom: TEM of fibbed section highlighting 15 nm thick film of $(Mn,Ti)O_2$ with accompanying EDS analyses. The view is directly down the tunnels (i.e., perpendicular to [001] surface). The white dots are individual Ti and Mn atoms, where O atoms are not visible.

be to maximize the concentration of Mn in the rutile structure thin films, which should increase conductivity of the material. This will require careful consideration of the effect of the substrate on film stability as well as the interplay of Mn, Ti, and oxygen defect densities. Finally, where effort in the first year of this work focused on the 1 × 1 tunnel structure, new structures, including 1 × 2 and 2 × 2 tunnels, will be attempted. Theoretical input into synthesis conditions will be crucial, as these structures are likely less stable than the 1 × 1 structures. As before, the parameters of interest are T, PO_2, Mn/Ti ratios, and substrate structure. The concept of using a mixed oxide framework for the tunnel structured materials increases the design flexibility and chemical space to achieve tunable properties. Future work should further highlight the power of combining theory and experiment rationally to design materials of high interest in energy applications.

PN13004/2485

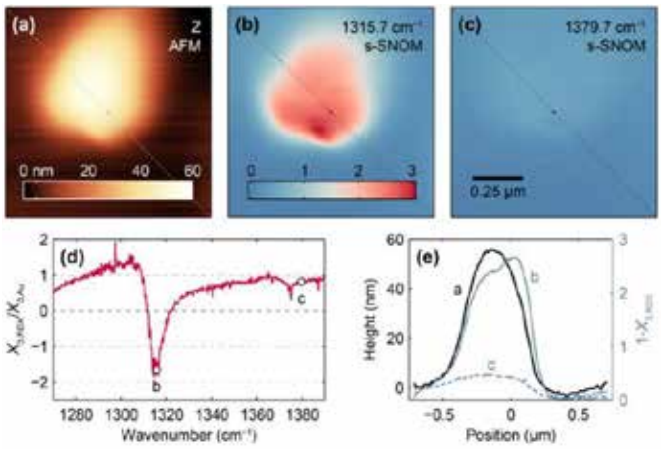
Facet Specific Chemistry of Noble Metal Nanoparticles Using an Enhanced Scattering Infrared Scattering Near-Field Optical Microscope

A. Scott Lea

We are adding a new capability to existing equipment that allows investigation of reactive sites on various faces of crystalline nanosized noble metals or model oxide surfaces and examining the physical processes that occur on these nanostructured materials.

Molecular and inorganic nanostructures, polymer and supramolecular assemblies, proteins, correlated systems, and many other natural and synthetic materials gain their unique functionalities from intra- and intermolecular interaction and electron correlations on mesoscopic length scales of 10s of nm. Gaining a molecular level understanding of the materials' structure and function has remained a major experimental challenge from the lack of techniques that routinely provide chemically specific spectroscopic identification with simultaneous spatial resolution on the relevant length scale associated with the size and interactions of the molecular building blocks within the 10–100 nm range.

This project integrated a quantum cascade laser with an atomic force microscope to develop a novel, state-of-the art capability in scattering infrared near-field optical microscopy to investigate the distribution of bonding configurations and heat of adsorption of probe molecules on different crystal facets of noble metal nanoparticles. External cavity quantum cascade lasers (ECQCL) were coupled to this instrument as a complementary infrared light source to the femtosecond OPO chain. These devices work extremely well in the molecular fingerprint region (8–12 μm), the range of many fundamental vibrational bands. ECQCL's rapid scanning capabilities in the 100s of Hz allowed infrared spectra to be taken point-by-point across the sample, forming rapid spectral data coverage. The second advantage of quantum cascade laser incorporation is its ease of use (permitting wide-spread usage), low cost, and ruggedness. Further, we established an inelastic electron tunneling spectroscopy (IETS) capability on EMSL's low-temperature scanning tunneling microscope (LT STM) that provided spectroscopic information about individual species involved in surface-mediated reactions at the atomic level.



Near-field chemical imaging of RDX. (a) AFM topography map of a particle of RDX. (b,c) s-SNOM contrast images taken at two different frequencies as indicated in the spectrum in (d). The color bar in (b) applies to both optical images, and the scale bar in (c) applies to all three images. (d) Normalized spectrum taken at center of the RDX. (e) Line scans across the diagonal of each image.

In FY 2011, we custom built an ECQCL with a continuous mode operating frequency range of 1260 cm^{-1} to 1400 cm^{-1} and identified an experimental system with nm- to μ-size spatial variability in its infrared absorbance from 1280 cm^{-1} to 1380 cm^{-1}. For FY 2012, the *s*-IR SNOM system was disassembled, transported to PNNL, and re-assembled and qualified for operation with the femtosecond laser system and ECQCL. For C-O adsorption experiments, we purchased a 5.02–5.52 μm ECQCL, successfully integrated it into the *s*-IR SNOM instrumentation, and demonstrated field enhancement of the second harmonic (2Ω) scattering SNOM signal imaging slot antennas.

During FY 2013, we used ECQCL spectroscopically to image the spatial distribution of a number of nitro group containing molecules on the surface of gold using IR wavelength tuned to the oscillator and demonstrated sensitivity, low noise, and rapid acquisition that greatly exceeds what can be obtained on femtosecond laser systems. Additionally, we used a finite dipole model to reproduce the qualitative aspects of the self-homodyne SNOM signal.

We incorporated the inelastic electron tunneling spectroscopy into the LT STM and developed a detection technique on the basis of sample bias voltage modulation and lock-in amplifier to improve a signal-to-noise ratio of IETS spectra. We carried out initial experiments using combination of LT STM with hyperthermal molecular beam to access the structural and dynamic aspects of various processes during adsorbate-substrate interactions for water molecules at the model TiO_2(110) surface. Using collimated H_2O beams with well-defined kinetic energies allowed probing reactions at a low temperature (80 K). Varying the energy of H_2O molecules, we directly identified active sites and activation energy barriers for various processes, including H_2O molecular adsorption, dissociation, and abstraction. Further insights were achieved via comparison of experimental results with time-resolved DFT calculations.

PN11035/2357

High Energy Density Non-aqueous Metal-Organic Redox Flow Battery

Wei Wang

We are establishing the framework for frontier research and development of a high energy density redox flow battery. Successful development of such system will enable the battery for high energy/power applications such as electrification of transportation sector, integration of green energy source, and deployment of the smart grid.

The demand for stationary energy storage has rapidly changed the worldwide landscape of energy system research, which recently brought the redox flow battery (RFB) into the spotlight. RFB is a leading technology in providing a well-balanced approach for the imminent need of a large-scale energy storage system. Several advantages originating from its unique architecture afford RFB the potential to advance and transform the energy storage technology. The unique RFB configuration provides a simple design that can limit the negative impact of inactive materials in a battery on the system energy density. Fundamentally different from the architecture of current secondary batteries, an RFB system has the potential to afford a transformational improvement in cost, manufacturing scalability, safety, and energy density as future energy storage system.

In this work, using Li/Li^+ as the anode and organic redox active agent as the active catholyte, we developed a high energy density lithium-organic RFB (LORFB). This flow system has not been demonstrated or reported before. The concept and system we demonstrated offer great promise to provide improvements such as system energy density, scalability, and cost. Our research consists of two areas. On the cathode side, we modify and engineering the molecular structure of the active organic agent to achieve high solubility resulting in high energy density catholyte. On the anode side, a new advanced hybrid lithium (Li) anode is developed to enable the long-term stable cycling. The combination of high energy catholyte and advanced hybrid anode contribute to the successful development of high energy LORFB. Because traditional RFB is a low energy density system, it limits its application to stationary energy storage. This effort will demonstrate the new chemistry that enables the RFB for high energy/power density application. Through this project, we will demonstrate for the first time that high energy catholyte can be achieved through molecular structure engineering to improve the solubility of the redox active organic agent.

We successfully developed an organic ferrocene (Fc)-based high energy density catholyte. Although the Fc possesses a high potential of ~3.5V, its solubility is extremely low in the organic electrolyte solvent. The ferrocene structure was, therefore, chemically modified to graft a polar IL pendant of tetraalkylammonium-TFSI to increase its solubility in the polar solvent. The ferrocene-based IL compound (Fc1N112-TFSI) was prepared via a nucleophilic substitution reaction of a commercially available ferrocene derivative, (dimethylaminomethyl) ferrocene (Fc1N11), with bromoethane in acetonitrile, to produce the dimethyl ethyl ferrocenylmethyl ammonium bromide (Fc1N112-Br), followed by exchange of the Br^- with $TFSI^-$ in deionized water to afford the final product at an overall yield of 91%. The resulting Fc1N112-TFSI shows a dramatically enhanced solubility of 1.7 M in the EC/PC/EMC solvent system, and thus dissolved in the 1.2 M LiTFSI in EC/PC/EMC supporting electrolyte to form 0.85 M high energy density catholyte. The theoretical energy density of the catholyte is nearly 90Wh/L, more than three times higher than the current aqueous electrolyte.

A new hybrid anode has been developed to accommodate the high energy catholyte. Traditional Li metal anode suffers from dendrite growth and breakdown during cycling and cannot support the high current charge/discharge necessary for excellent flow cell cycling performance. We designed a completely different anode structure where the interfacial redox reaction is shifted away from the Li surface. The graphite felt/Li connected in parallel forms a shorted cell where the graphite felt is always lithiated at equilibrium state and maintains a pseudo-equal potential with the Li metal. As such, it functions as a dynamic "pump" that supplies Li^+ ions on demand. It also serves as an artificial SEI layer of Li metal. As the result, the Li organic flow cells were able to cycle at much higher concentration and current density.

Combining the high energy catholyte and hybrid Li anode, we successfully demonstrated stable cycling of a nonaqeuous RFB with actual energy density of about 50Wh/L, two times higher than the current aqueous flow battery systems. We also filed four immediate closure reports. One U.S. patent application has been filed, and another three are in process.

The most important finding of the work is to demonstrate that the solubility of the redox active organic agent can be improved through the molecular structure engineering and the surface chemistry of the Li anode can be modified to preserve the stability. These findings are critical toward our overall goal of a high energy density RFB. The molecular structure tailoring increase the solubility of the Fc more than ten times as the energy bearing active materials, while the hybrid anode structure enable the flow cell to be cycled at the maximum energy density.

PN13023/2504

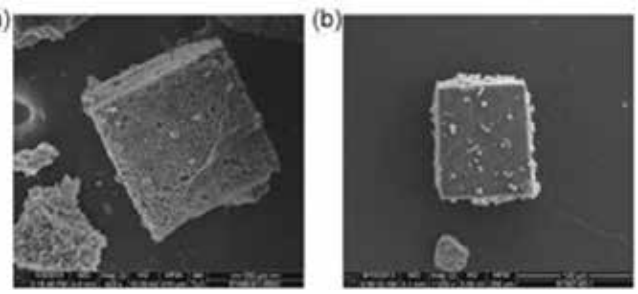
Hybrid Electrodes for Next Generation High Energy Ultracapacitors

Satish K. Nune

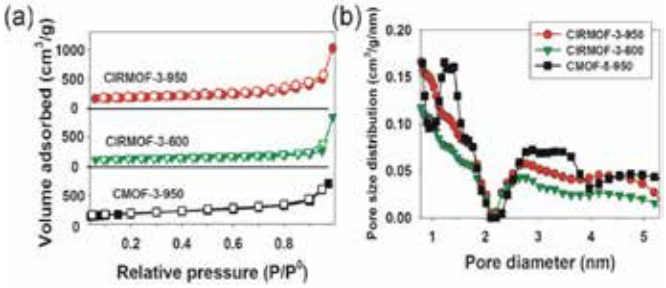
We are developing solution-based strategies for carbon architecture synthesis with large surface areas and controlled pore distribution for improved mass/charge transfer properties from an enhanced electrode/electrolyte interface.

Supercapacitors with high power and energy densities in combination with batteries serve as potential alternatives to increase energy densities while providing high power performance. Carbon nanotube (CNT) and graphene nanosheets (GN) with outstanding properties including high surface area, thermal conductivity, and stability have been used as candidate materials. However, the energy density of existing state-of-the-art capacitors is an order of magnitude lower than metal hydride and Li-ion batteries. GN tends to aggregate and stack in multilayers from vanderwall interactions affecting performance, high cost of CNTs, and limited functionalization strategies limiting their use. Inexpensive mesoporous carbon fiber-based electrode materials with efficient ion diffusion pathways serves as potential alternative. Understanding fundamental mechanisms and properties required for improved performance enables modification of synthesis strategies that offer electrode materials with enhanced electrode/electrolyte interfaces. Characterization and electrochemical measurements of synthesized materials will advance our understanding of fundamental mechanisms for improved performance reducing American dependence on foreign energy (oil) imports.

In FY 2013, we performed a synthesis and complete characterization of nitrogen-doped porous carbons with controlled composition and high surface area using IRMOF-3 as a sacrificial self-template as a precursor without any additional carbon and nitrogen sources. To investigate the nitrogen doping effect on electrochemical performances, MOF-5 was used to produce porous carbons. The as-synthesized MOFs were carbonized directly in an inert atmosphere (Ar or He) at 600, 700, 800, and 950°C. The carbons derived from these MOFs had high surface areas and significant differences in nitrogen doping depending on the MOFs ligand. When the carbonization temperature is maintained at 950°C, *in situ* reduction of ZnO to Zn occurs followed by Zn sublimation at > 908°C to produce high surface area carbons (CMOF5 and C-IRMOF3). As CMOFs carbonized at 950°C, additional hydrochloric acid treatment was not required. Our method to produce nitrogen-doped porous carbons from nitrogen-containing MOFs did not require external hard templates or any additional nitrogen and carbon sources, a significant advantage compared to other methods. To the best of our knowledge, this is a first report to produce nitrogen-doped porous carbons using MOFs through direct carbonization.



SEM images of (a) CMOF-5 and (b) CIRMOF-3



(a) Nitrogen adsorption-desorption isotherms; and(b) pore diameter calculated by DFT.

In the SEM images, the porous structure of CIRMOF-3-950, and CMOF-5 was clearly observed. To characterize the porous nature and determine BET surface area, nitrogen adsorption/desorption measurements were performed on MOF precursors (MOF-5 and CIRMOF-3) and on porous carbons obtained after carbonization. The synthesized CMOFs showed large uptakes at low relative pressure, indicative of the presence of abundant micropores. Following a large uptake of nitrogen at low pressure, plateau region and steep increase of adsorbed nitrogen at high relative pressure were observed, which likely originated from large meso and macro pores from an interstitial void between particles. The pore size distribution was obtained from nitrogen isotherms using density functional theory (DFT) method, and analysis reveals the presence of micropores and good degree of mesoporosity in C-MOF5 and C-IRMOF3, vital for achieving high specific capacitances.

Using XPS mapping, we also verified that the nitrogen and oxygen in CIRMOF-3-950 and oxygen in CMOF-5-950 were

PN12039/2440

evenly distributed rather than located in specific region. To further characterize the nature of nitrogen in porous carbons, the nitrogen peak of IRMOF-3-950 was deconvoluted into four different peaks. Pyrrolic nitrogen (N-5) was reported to be thermally less stable than other forms of nitrogen, and quaternary nitrogen (N-Q) could be generated from other forms of nitrogen such as pyridinic (N-6). Considering the account stability of each nitrogen form, the dominance of N-Q is quite reasonable and expected.

To investigate electrochemical performances as a supercapacitor electrode, the electrodes were prepared using MOF-derived porous carbons with carbon black and polyvinylidene fluoride (PVDF). Symmetric coin cells were assembled with separator and 1M sulphuric acid electrolyte. Electrochemical performances were assessed using cyclic voltammetry (CV) and galvanostatic charge/discharge measurements. All porous carbons derived from IRMOF-3 at various carbonization temperatures exhibited almost rectangular shape of CV curves, which is typical behavior of ideal capacitors. Nitrogen doped carbon architectures (C-IRMOF3) exhibited much higher specific capacitance than simple carbon (C-MOF5) at 2 mVs^{-1} of scan rate, respectively. The CIRMOF-3-950 has the largest CV curve compared with other porous carbons, even though it has the lowest nitrogen content among CIRMOF-3. The excellent electrochemical performances of CIRMOF-3 can be ascribed to its less defective structure as well as high specific surface area. Nitrogen in high specific surface area porous carbons induce additional pseudocapacitance and helps wettability, resulting in increased electrolyte-electrode interaction for high-performance supercapacitors. On the other hand, CMOF-5-950 showed distorted CV curves, which might be from the large internal resistance because of inferior wettability with electrolyte. This result is well consistent with its poor wettability with water caused by hydrophobic nature of CMOF-5-950.

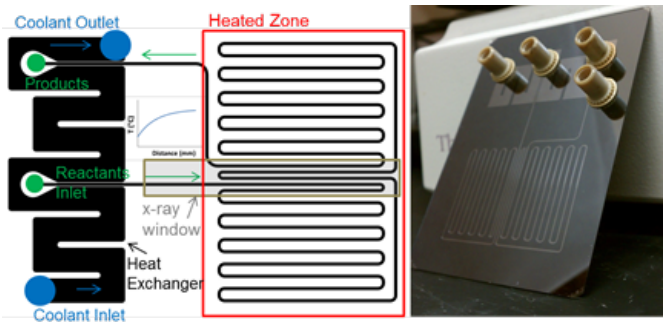
In summary, we developed a strategy to develop highly efficient electrode materials through control of surface area and composition. Nitrogen-doped porous carbons synthesized through self-templating approach without additional carbon and nitrogen sources exhibited significantly improved super-capacitive behavior compared to pure carbons due to enhanced wettability and additional pseudocapacitance induced by nitrogen.

Imaging the Nucleation and Growth of Nanoparticles in Solution

Ayman M. Karim

This project is developing* in situ *characterization tools for imaging the synthesis of nanoparticles in solution. Direct real time visualization would allow a fundamental understanding of synthesis mechanisms and therefore accelerate synthesis development to make materials with specifically engineered nanostructures for a wide range of energy storage and conversion applications.

The synthesis of nanoparticles in solution is the most promising tunable method but is not well understood due to the difficulty of experimentally observing the nucleation and growth processes. Consequently, most synthesis efforts rely on trial and error from the lack of a fundamental understanding and paucity of *in situ* characterization tools. Our goal is developing new *in situ* tools for imaging the nucleation and growth of nanoparticles in solution by combining microfluidics, X-ray absorption spectroscopy, and atomic-resolution aberration-corrected transmission electron microscopy (TEM). The project will result in an *in situ* microfluidic cell and development of data collection and analysis techniques for the study of nanoparticle nucleation and growth with ms time resolution using synchrotron techniques and ex situ atomic-resolution aberration-corrected TEM.



Microfluidic reactor for millisecond time-resolved X-ray absorption spectroscopy studies of nanoparticles synthesis. The heat exchanger at the inlet and the heaters by the serpentine channel allow for two temperature zones and a well-defined t_0. The X-ray window covers the inlet, two middle, and outlet channels, which allows for ms time resolution during the nucleation and up to minutes time resolution during the growth.

We designed and fabricated a novel microfluidic silicon/pyrex reactor for Pd nanoparticles synthesis that allows for ms time resolution during the entire synthesis duration to observe nucleation and nanoparticle growth. The reactor allowed us to study Pd nanoparticles synthesis with XAFS and SAXS in the same reactor. Combining *in situ* XAFS and SAXS, we determined the synthesis mechanism, which was found to be nucleation limited–a slow, continuous process while the growth was surface catalyzed by the nanoparticles. According to the literature, slow continuous nucleation typically leads to large nanoparticles with broad size distribution. In contrast to the classic nucleation and growth theory, we discovered that instead of the particles continuing to achieve large sizes, growth was arrested at a certain size. XAFS results showed that the arrested growth was from increased bonding with the capping agent, causing existing particles to stop growing after reaching a certain size. New nuclei continued to form new particles through the same mechanism (although at slower rate) until the remaining precursor in solution was depleted. Therefore, we showed that it is possible to synthesize Pd nanoparticles with narrow size distribution despite slow nucleation rate compared to growth rate. This is the first report of Pd nanoparticles in the 1 nm size range with narrow size distribution based on understanding synthesis mechanisms and not trial and error.

The reactor used for these studies was designed for Pd (K absorption edge = 24.35 keV) and other metals with similar absorption energies (> 20 keV). We showed that the reactor design can be modified easily to accommodate thinner windows for the characterization of elements with lower X-ray absorption energies (as low as 9 keV) and high quality XAFS was collected for Pt (11.56 keV) precursor solution in the modified reactor.

Our work in FY 2014 will focus on the design and fabrication of a droplet flow microreactor for high temperature nanoparticles synthesis and developing data acquisition methodologies for XAFS and SAXS. The droplet flow reactor has advantages over the single phase flow reactor that we used in the first two years of the project. This new reactor eliminates metal deposition on the walls and fouling, a major issue for high temperature synthesis. In addition, the droplets act as well stirred nano-batch reactors (on the order of nanoliter each), where the concentration gradients are minimized, leading to much narrower particle size distribution. The droplet flow microfluidic reactor and XAFS/SAXS data acquisition methodology will be developed using Pt nanoparticles synthesis; however, it can be generalized to many other nanoparticles synthesis systems.

PN12041/2442

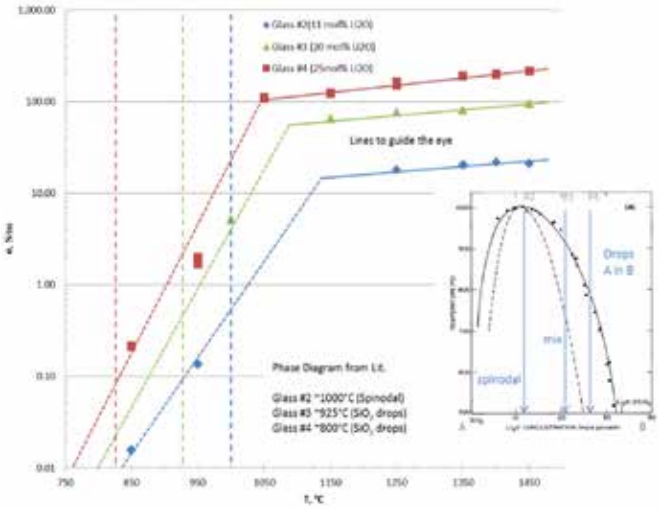
Impedance Spectroscopy: Next Generation Tool for Waste Form Characterization

Jarrod V. Crum

This project demonstrates the viability of impedance spectroscopy (IS) for studying the evolution of nuclear waste forms and applies the techniques and instrumentation to the waste form problem.

High-level nuclear waste contains several components (Mo, F, S, and P, to name a few) that tend to undergo phase separation in borosilicate glass upon cooling when above their solubility limit. In some cases such as in glass ceramics, the phase separation also nucleates crystalline phases. An incomplete understanding of phase separation leads to difficulty in modeling the process by which a single-phase melt transforms into a multi-phase glass or glass ceramic. Phase separation is a dynamic process that occurs in silicate melts upon cooling as a function of chemistry and thermal history. Therefore, it is difficult to quantify *ex situ* because it cannot be stopped with rapid quenching. For this reason, an *in situ* technique such as IS is required to characterize the phase separation process in a molten state.

IS is a powerful technique that is used to examine the electrical properties of multiple components within a system. In this case, the components are assumed to be electrodes, electrode-to-melt interface, melt(s), and/or crystal(s). For this research, IS was used to investigate the phase separation process of a glass in situ upon cooling from a single phase melt. The simple and well known binary Li-silicate glass system was chosen for this work. If successful, this technique could be applied broadly to study phase separation (liquid or crystalline) that occurs in a melt state (waste forms, commercial products, and even igneous petrology). This would be a valuable tool that could be used to understand and/or optimize the phase separation process and subsequent crystallization. Accurate modeling of these processes will result in better control of the phase assemblage of the final form.



Conductivity plotted in log scale vs. temperature (immiscibility dome inset from Haller, Blackburn, and Simmons [1974], *Journal of the American Ceramic Society*).

IS was performed on three lithium silicate glass compositions (11, 20, and 25 mole % Li_2O, balance SiO_2) as a function of temperature in the melt state. The intent was to deconvolute the impedance data into their respective portions (melt A, melt B, and/or crystal C). However, there are significant interference effects (such as induction) that make full deconvolution difficult for melts at this temperature. For this reason, only the bulk conductivity of the melt could be determined at this time.

In addition to temperature, melt conductivity was determined by fitting a portion of the data to a circuit model (ZARC) and plotting versus temperature. At high temperatures, the glass conductivities all follow an Arrhenius relationship with respect to temperature. At lower temperatures, however, the conductivity of the glass melts all decrease noticeably. This shift indicates a significant change in the morphology of the overall melt, most likely because of phase separation and/or crystallization. In addition, there was a temperature region (1150–1050°C) between the two conductivity regimes where the impedance signal was extremely weak and noisy.

Interestingly, the change in slope of the conductivity vs. temperature for each of the glasses tested occurs between roughly 100–200°C above the immiscibility dome (figure inset) that was determined for the Li_2O–SiO_2 system in previous research. However, the dome was determined visually by opalescence after quenching. These differences suggest that we are observing the effects of a smaller scale phenomenon, such as clustering, that precedes the development of optically detectable phase separation.

If verified, this research represents a breakthrough in the detection and evaluation of the mesoscale processes leading to phase separation. Confirmation requires other *in situ* techniques to verify the presence of clusters and to determine how these clusters impact the conductivity of the melt.

PN13041/2522

Improving Magnetoelectric Coupling in Novel Single-phase Multiferroic Thin Films of the $MTiO_3$ (M = Fe, Mn, Ni, . . .) Family

Tamas Varga

The project will enable us to design multiferroics via the control of chemistry and physical properties, providing key knowledge to the rational synthesis of novel materials with potential technological impacts on information storage, energy, and the semiconductor industries.

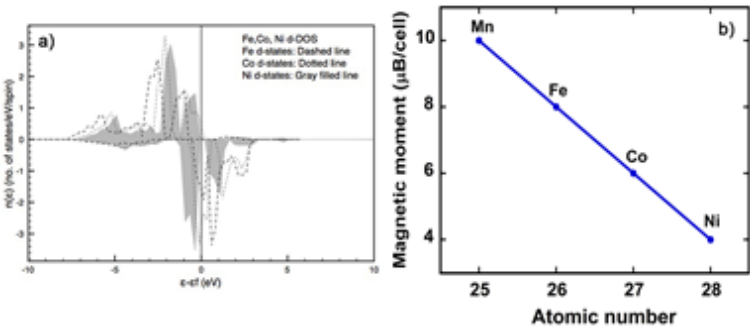
Multiferroics (materials that are both ferromagnetic [FM] and ferroelectric [FE]) can be employed as new four-state memories for future data storage applications and in photovoltaic devices for energy applications. However, the coupling of FM and FE in single-phase multiferroics is inherently weak. The objective of the project is to understand the physical origin of the coupling of polarization and magnetization through theory and use that knowledge to design novel multiferroic materials that perform better. Building on our success in stabilizing $NiTiO_3$ in the theory-predicted structure in epitaxial thin film form and showing the coexistence of polarization and ferromagnetism, our hypothesis is that alloying/doping $MTiO_3$ will affect FM (and possibly FE) properties and, in turn, their coupling. Ultimately, we wish to demonstrate that the resulting multiferroic properties are superior to earlier results in related single-phase systems.

Designing new multiferroic materials as the prediction of the desired compositions that meet our design criteria, we began by performing a series of theory calculations to study the origins of FM/FE formation and stability in the dopant metals Cr, Co, Mn, V, with an eye on making doped $MTiO_3$ phases such as $Cr_xM_{1-x}TiO_3$ next. We examined how the electronic structure of the transition metal in the magnetic site (M = Fe, Co, Cr, Mn, Ni) affects the energetics and magnetism in $MTiO_3$ phases. To make better multiferroic materials in the $MTiO_3$ family, we will need to increase the magnetic moment on the transition metal M and/or enhance the ferroelectric polarization in the compound.

We first looked at the magnetism by calculating the density of states (DOS) for metallic Fe, Co, and Ni, and the magnetic moment in $MTiO_3$ phases for M = Fe, Co, Mn, and Ni. There appears to be a rigid shift in the calculated DOS with the number of electrons on the magnetic ion. This results in a linearly decreasing magnetic moment going from Mn to Ni. These findings suggest that the smaller atomic number metals (such as Mn and Fe) may be better candidates as the M cation in $MTiO_3$ multiferroic materials.

Density functional theory calculations of the ground state energies for the antiferromagnetic (AFM) and FM states of Co, Cr, Fe, Mn, Ni, and V suggests that Cr may be a good candidate to create doped $MTiO_3$ (M = Fe, Ni, Mn, Co) such as $Cr_xM_{1-x}TiO_3$ with enhanced FM moment relative to the MTiO3 end member as the FM state of Cr has a favorably low ground state energy in the desired structure. Based on this theory work, efforts are now underway to synthesize both pure $CrTiO_3$ or Cr-doped $NiTiO_3$ phases.

On the synthesis side, building on previous results of the PI on related materials, efforts to date have been focused on two directions: attempts to stabilize new end members $FeTiO_3$ and $MnTiO_3$ with the desired structure as epitaxial thin films; and explore a new deposition and substrate options to improve film quality. For the new end members, $FeTiO_3$ and $MnTiO_3$ have been successfully synthesized in epitaxial form using pulsed laser deposition (PLD). Physical property characterization of the samples is underway, but structural results suggest that epitaxial films with the desired multiferroic structure were made. Molecular beam epitaxy was attempted to prepare $FeTiO_3$ and $NiTiO_3$ thin films on Al_2O_3 substrates following the successful PLD growth of these samples. Preliminary results indicate that under the first synthesis conditions, the samples contained a mixture of binary oxides Fe_2O_3 and Fe_2O_3 as well as metallic Ni and TiO_2 instead of $FeTiO_3$ and $NiTiO_3$, respectively.



a) Calculated density of states (DOS) for Fe, Co, and Ni metals exhibiting a rigid shift to higher eV in the DOS with the number of electrons on the magnetic ion; b) Linearly decreasing magnetic moment (calculated per unit cell in R-3 $MTiO_3$) with an increasing atomic number of the transition metal.

PN13072/2553

	V		Cr		Mn		Fe		Co		Ni	
	R3	*R3c*	*R3*	*R3c*	*R3*	*R3c*	*R3*	*R3c*	*R3*	*R3c*	*R3*	*R3c*
Magnetic moment/cell	2.468	2.887	4.606	4.000	10.000	10.000	8.000	8.000	6.000	6.000	4.000	4.000
Ground state E for AFM (eV)	-89.995	-90.048	-88.933	-88.859	-88.198	-87.809	-84.280	-84.405	-80.447	-84.406	-77.791	-77.918
Ground state E for FM (eV)	-89.516	-89.743	-88.836	-89.130	-88.044	-87.695	-84.717	-84.401	-80.645	-84.401	-77.757	-77.878
Ground state E for NM (eV)	-89.324	-89.312	-87.271	-87.774	-85.086	-86.002	-82.794	-83.918	-82.794	-83.918	-76.725	-77.073

Results from density functional theory calculations of the ground state energies for the AFM, FM, and nonmagnetic states of V, Cr, Mn, Fe, Co, and Ni as M in fictitious $MTiO_3$ with both the ilmenite (*R3*) and the acentric $LiNbO_3$ (*R3C*) structures, of which the latter is predicted to be multiferroic. The case where the FM ground state energy for the *R3C* structure is the lowest for a particular metal with significant favor of the *R3C* structure relative to the *R3* is found for Cr (highlighted); E = energy.

Previous results from earlier work on $NiTiO_3$ indicated that employing Fe_2O_3 instead of Al_2O_3 as substrate may result in smoother film growth, better crystalline quality, and perhaps improved physical properties (the lattice mismatch between $NiTiO_3$ and Fe_2O_3 is small, thus strain is reduced). To explore this graded approach to introducing lattice strain, films on both single crystalline Fe_2O_3 substrates, and on thin Fe_2O_3 films deposited on Al_2O_3 substrates were grown by PLD. Our structural characterization results suggest that the latter, $MTiO_3/Fe_2O_3/Al_2O_3$, films show promise, while the Fe_2O_3 substrate itself promotes the formation of the corresponding binary oxides (Fe_2O_3 + TiO_2, or NiO + TiO_2) rather than the desired titanates.

During the next year, we will continue theory calculations to determine alloying scenarios that will enhance the magnetic moment and/or the ferroelectric polarization (thus, their coupling). Certain alloys naturally create displacement effects, which may yield a strategy to alloy the M site with Fe+Cr, Fe+Mn, Fe+Ni, and similar combinations to create the material with the desired multiferroic properties through such effects. To this end, we will include rare earth metals in the calculations (large magnetic moment). In addition, we will perform experimental characterization of the physical properties of the end members to match theoretical values and establish a relative scale for judging property enhancement upon doping. The synthesis of the phases predicted by theory calculations will be determined and will iterate between experiment and theory as experimental results become available. Finally, we will perform synchrotron beamline experiments to probe magnetoelectric coupling at the Advanced Photon Source at Argonne National Laboratory.

Improving the Performance of Li-Air and Li-S Batteries Using Polymeric and Metallic Nanomaterials

Priyanka Bhattacharya

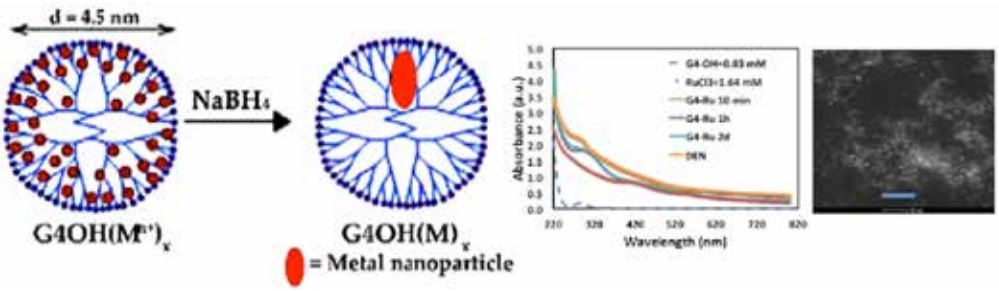
This project is developing novel nanomaterials to enhance the performance of lithium (Li) batteries and understand their fading mechanisms.

An increased demand for high energy density batteries based on environmentally friendly materials has led to the research and development of rechargeable Li batteries, which include Li-air (Li-O_2) and Li-sulfur (Li-S). A challenge with the Li-O_2 battery is recharging the insoluble, poorly conducting Li_2O_2 discharge products to Li^+ and O_2 during charging. Hence, Li-O_2 batteries require electrocatalysts for reducing the over-potentials of both reactions. For Li-S batteries, charge/discharge involves the creation of highly soluble polysulfide intermediates in electrolytes, causing an irreversible loss of active sulfur. Currently, efforts are directed toward polymers and nanostructures to trap polysulfides, thus improving the electrochemical cell cycling. The remarkable physico-chemical properties of dendrimers motivated research into the application of dendrimer-encapsulated Ru electrocatalysts in Li-O_2 batteries as well as their application in trapping polysulfides in Li-S batteries. Thus, our objective is to identify novel, efficient materials for high energy density applications and place PNNL at the frontier of dendrimer research in energy storage.

Dendrimer-encapsulated ruthenium nanoparticles (DENs) as catalysts in Li-O_2 batteries. We synthesized monodispersed and stable DEN of average size (~ 2 nm). Once complexation was complete, the Ru^{3+} were chemically reduced in situ to Ru metal using 10 molar excess of $NaBH_4$. Complexation and reduction were monitored using UV-vis absorption spectroscopy. After reduction, the DEN solutions were dialyzed using 10,000 MWCO dialysis membranes to remove impurities. The average size of the DENs with both the hydroxyl and amine terminated dendrimers was ~ 2 nm. However, while the hydroxyl-terminated DEN solutions remain stable for > 2 months, amine-terminated DENs aggregate within a few days. This is because some of the Ru^{3+} complex with the primary amines of the dendrimers that when reduced remain at the periphery of the dendrimers and unprotected, leading to aggregation over time. Subsequently, all experiments were conducted using hydroxyl-terminated DENs (G4-OH-DENs).

PAMAM dendrimers encapsulating nanoparticles demonstrated the potential for being catalysts in Li-O_2 cells. The DENs were stable for months and were electrochemically stable toward Li metal and electrolyte solvents. The materials improved the cycling efficiency of KB carbon by ~ 3.3 times and lowered the charging overpotential by ~ 0.4V, thus creating a new state-of-the-art catalyst for Li-O_2 batteries. Most importantly, the amount of Ru used as a catalyst is 6 times less than that in the state-of-the-art work published to date. Thus, dendrimer-stabilized Ru nanoparticles enable the use of a much lower amount of the noble metal catalyst.

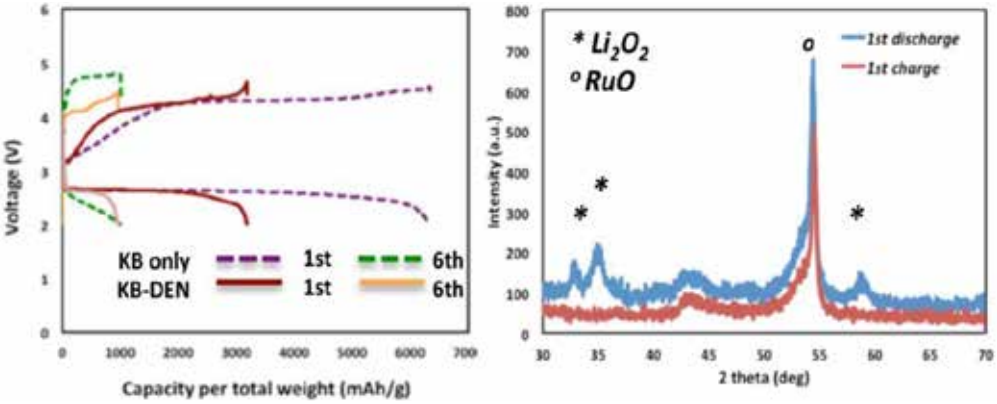
To analyze the nature of the discharge products, scanning electron microscopy (SEM) and X-ray diffraction (XRD) spectroscopy were conducted on discharged cells and washed in dimethoxyethane (DME) for 3 days to remove electrolyte salts. Both SEM and XRD showed the presence of Li_2O_2 as the major discharge product, with some electrolyte decomposition products also observed. Charged cells showed little or no Li_2O_2, suggesting good rechargeability. XRD results showed the presence of ruthenium oxide (RuO_2) rather than Ru metal, which suggests oxygen adsorption on Ru metal surface in ambient conditions. This indicates that O_2 diffuses into the dendrimer periphery and enables the use of RuO_2 as catalysts, which have been recently shown to be better OER catalysts than Ru metal. Preliminary results from this work have been presented at two conferences. A provisional patent has been filed, and a manuscript is in preparation.



The synthesis of DEN, UV-vis absorption spectra of dendrimer-ruthenium complexes and DEN, and a scanning tunneling electron microscopy image showing ~2 nm DENs as white dots (scale bar = 10 nm).

PN13013/2494

Dendrimers for capturing polysulfides in Li-S cells. G4-OH and G4-NH_2 dendrimers were mixed with KB-S composites in aqueous solutions for 1 day. The dendrimer wt% was 10 wt% and S was 80 wt% in the composite. Water was dried under low vacuum at 40°C overnight, and the dendrimer-KB-S composites were cast into a slurry using 20 wt% poly(vinylidene) fluoride (PVDF) in N-methyl-2-pyrrolidone (NMP) solvent. Cathodes were prepared by coating aluminum foils as current collectors with the slurry. The film thickness was ~ 150 µm. 1.4 cm diameter cathodes were punched out and dried in low vacuum at 40°C overnight. 2032 coin-cells were assembled in an argon-filled glove box with O_2 and H_2O content < 0.1 ppm. The cells were rested for maximum 6 h to prevent their self-discharge and tested. Initial results showed a poorer performance of these composite electrodes. Thus, PAMAM dendrimers did not improve the performance of Li-S batteries. This could be due to the insulating nature of PAMAM dendrimers. Further testing using DENs of different generations will enable us to evaluate the performance of dendrimers in Li-S batteries.



Capacity-voltage profile and XRD data of Li-O_2 cells with KB carbon and DEN catalysts. A 0.4 V reduction in charging voltage is obtained with DEN as catalysts.

In FY 2014, we will investigate the structure and oxidation state of metals in DENs, both parameters of which influence catalytic activities. Structures of dendrimer-Li peroxides and dendrimer-polysulfides will be studied to understand the nature of the complexes. These studies will enable us to optimize the dendritic structure for applications in high-energy storage, including the finding of alternate, more cost-effective hyperbranched polymers. We will also employ the technique of atomic probe tomography (APT) in investigating the distribution of Ru within the dendrimer interior. Finally, we will investigate the function of dendrimers as synthetic analogues of ferritin proteins, which will open a new area of application for APT for drug delivery.

PN13013/2494

Meso-scale Science and Technology: Manufacturing of Nanostructured Soft Magnetic Materials

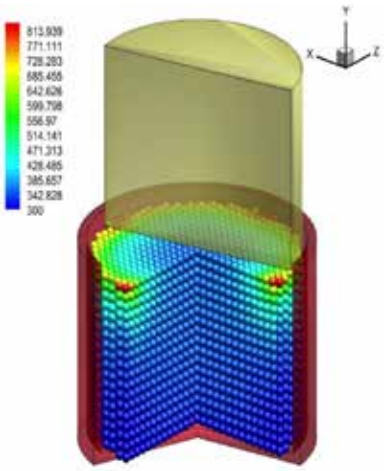
Jun Cui

We are developing novel consolidation methods that will enable robust and economical manufacturing of nano-structured bulk soft magnetic materials to the United States in energy security and technology leadership.

Over 80% of the global electricity used in 2012 was consumed or processed by motor and power electronics. As the world pursues greater energy efficiency and wider application of green energy, soft magnetic materials, which are critical components of motors and power electronics, are lagging behind permanent magnetic materials and wide band-gap materials. The currently dominant soft magnetic materials are FeSi and ferrites; compared to them, advanced soft magnetic materials (ASMMs) offer better combinations of magnetization, coercivity, permeability and electric resistivity, and may lead to higher efficiency, smaller size, and lower cost for power conversion devices. Limitations of current manufacturing methods, however, restrict ASMMs to thin ribbon form; thus, they cannot be widely used for electronics devices that require soft magnetic materials in bulk form.

We are attempting to understand the thermodynamic and kinetic laws governing the formation of nanostructures in bulk materials, pushing the thermodynamic and kinetic boundaries theoretically and experimentally with the ultimate goal of developing a method to produce nanostructured bulk materials cost effectively. Friction consolidation, cold spray, and shock compaction are nonequilibrium processing methods exploiting unique combinations of thermodynamic and kinetic parameters. They are identified as promising methods to produce nanostructured bulk materials. The main objectives are to 1) identify a viable modeling technique to simulate the three processing methods and complete preliminary modeling studies; 2) establish infrastructures for the three processing methods and perform early stage experiments; and 3) explore scalable nanosynthesis methods.

On the modeling front, the smoothed-particle hydrodynamics (SPH) approach was identified as the most suitable method for modeling friction consolidation, shock compaction, and cold spray processes. Because of its Lagrangian particle nature, SPH has several advantages for modeling the composite deformation and material flow of dissimilar materials: 1) complex interface and free-surface flows/deformations can be modeled without significant computational efforts; 2) accurate solution of momentum dominated flows is possible; 3) coupled physics such as melting or solidification, heat generation, strain and strain rate histories, stick-slip conditions, fracturing, void formation, and chemistry can be implemented. In FY 2013, a 3D SPH model was established to simulate the temperature profile for the friction consolidation process. The SPH model was able to predict the trend of the temperature profile as a function of load, plunge rotation speed, and process duration. While the model correctly predicted the trends, it failed to predict the temperatures accurately. The relatively large error is caused by the lack of accurate data on the stress-strain-temperature properties of the raw materials.



A clip created by the 3D SPH model showing the temperature profile during the friction consolidation process.

On the experimental front, friction consolidation was the major method under investigation in FY 2013. FeNiMo flakes were used as friction consolidation. Feedstock flake ($\sim 1 \times 50 \times 50\ \mu m^3$) was obtained by ball milling melt-spun ribbons. After consolidation, the flakes formed a fully dense sheet with submicron microstructure (~ 500 nm). Increasing processing time may reduce the grain sizes but also raises the temperature, which in turn causes the materials to decompose. The consolidated FeNiMo grains were decorated with $FeNi_3$ precipitates, indicating that processing temperature exceeded the maximum allowable. This is a good example showing the need to optimize processing parameters, stress, time, and temperature to obtain the desired grain size without decomposing the material. To expand the parameter range, a cooling fixture with chilling water circulating around the containment die was developed. A testing run showed that, for a consolidation run

PN13061/2542

with 10,000 lbf load and 125 RPM rotation speed, the peak temperature can be reduced by 135°C. The addition of active cooling allows the feedstock material to be processed with larger load and longer time without risking material decomposition. We received the second batch feedstock powders (~ 1 kg) and are ready to run it with the new cooled setup.

To create the shock wave, we originally selected, electromagnetic force generated by a large pulse current. Unfortunately, the large capital required by the capacitor bank and the electromagnetic coil forced us to take a more conventional approach. Currently, the team is working on setting up a drop tower (~ 20 mph) and an explosive nail gun (~ 40 mph). Due to the high energy involved in the experiment, the safety review is rigorous and lengthy.

Two nanosynthesis approaches were investigated in FY 2013, namely nitrate explosion and cryo-attrition milling. About 20 g of MnO particles with ~ 20 nm diameter was successfully prepared using the nitrate explosion method. Both hydrogen and CO were used to reduce the MnO at the elevated temperatures suggested by the Ellingham diagram. Unfortunately, only a small fraction of the MnO can be reduced to metal. Additionally, experiments were performed to develop a cryo-milling method that can bring the particle size down to submicron range without alloying or crystallization. Due to the mechanical energy delivered by the attrition balls, particles are often alloyed or decomposed. If feedstock materials are amorphous, the attrition mill can cause crystallization. The attrition mill can process a kg-size batch, so it is considered a scalable approach. The key feature of our approach is the use of liquid nitrogen during the milling process. The extremely low temperature not only keeps the feedstock brittle (and increases milling efficiency) but also keeps the feedstock particles below the crystallization temperature. In addition, the ongoing vaporization of liquid nitrogen keeps the feedstock free from oxidization. Previously, the smallest particle size we achieved was 40 µm; as of the end of FY 2013, we pushed the limit to 2 µm.

In FY 2014, we will continue the modeling, synthesis, and consolidation work started in FY 2013. We will refine the 3D SPH model to enable accurate prediction of temperature (±20°C) and grain size. In addition, we will establish SPH models for shock compaction and cold spray. We will complete development of the cryo attrition mill method with the goal of 500 nm particle size without crystallization or alloying. Further, we will develop a carbothermic reduction method to obtain magnetic particles of nano size. Finally, we will continue to use the friction consolidation method as the major approach to attain nanostructured bulk materials. After completing the drop tower and gas-gun setups and using them to consolidate powders, we will collaborate with United Technologies Research Center to investigate the cold spray method.

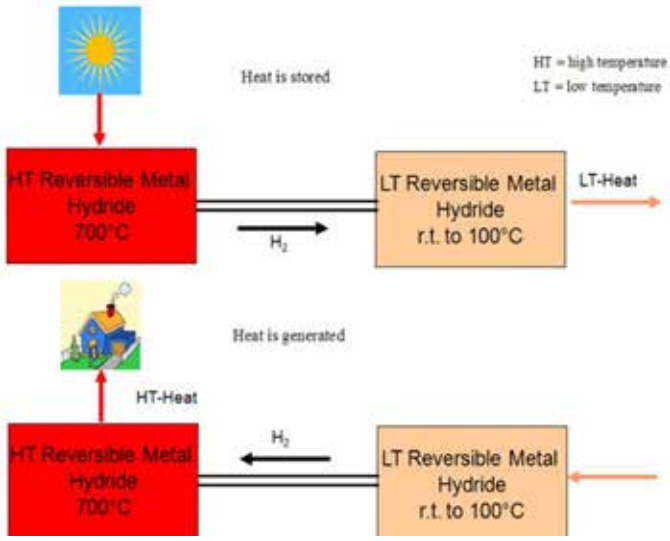
Metal Hydrides for Thermal Energy Storage

Ewa C.E. Rönnebro

To enable intermittent solar power technologies, we are developing a low-cost efficient thermal energy storage for large-scale utility heating when the sun is not shining or when excess energy is needed during peak times.

With a future of renewable energies for a sustainable environmental future, a key enabling technology is heat storage to generate electricity. Thermal energy storage (TES) technology holds high-temperature heat (solar energy or waste heat) that holds a concentrating solar power (CSP) and other applications. The concept is based on a reversible thermal cycle driven by gas-solid thermochemical reactions. To this end, we are developing a TES based on dual metal hydride (MH) beds for solar applications that have 10 times higher energy density than molten salts, the current state-of-the-art. The metal hydride materials must be selected and developed carefully to meet efficient storage requirements. We are using a high temperature (HT) metal hydride that can generate heat at 700°C when a chemical bond is formed between a specific metal and hydrogen.

When the sun is heating the HT-bed, hydrogen is released from the metal at atmosphere pressure and moves into a low temperature (LT) hydride bed, where the hydrogen will be stored. When the heat is needed, the hydrogen moves back to the HT-bed, and heat is generated to produce power. The beds are based on metals that reversibly absorb hydrogen to form metal hydrides. The compositions of the metal hydrides are crucial to obtain the desired performance; therefore, a materials development program is necessary. We previously identified the HT-hydride, but we also need to determine the LT-hydride composition to complete the dual bed TES system. Therefore, we embarked on synthesizing several new alloy compositions to identify a material that can operate near room temperature. To keep costs low, we investigated materials based on the metals Ti, Fe, Ni, Mn. The goal is to identify a low-cost alloy composition for the LT-bed that absorbs/desorbs hydrogen reversibly near room temperature at atmosphere pressure.



Metal hydride thermal energy storage

To identify a LT-hydride that can store hydrogen near room temperature and at atmosphere pressures, we prepared a dozen metal alloys of various compositions during FY 2013. The best performing LT-hydride materials will be down selected for future testing with the HT-hydride to investigate if they match up with respect to kinetics and hydrogen capacity to enable rapid hydrogen diffusion between the two beds for heat generation upon demand. Further fine tuning of the material's composition may be necessary, which will involve finding the optimized composition of a selected materials system.

To select LT-hydride candidates that potentially match with the HT-hydride, we performed a comprehensive literature study. Fine tuning of the LT-hydride composition involves synthesizing alloys to find the optimal relation between Ti, Fe, Mn, and/or Ni to match the performance with the HT-hydride. The most important characteristic is that the kinetics of hydrogen absorption and desorption must match between the two beds to guarantee adequate hydrogen diffusion in between beds. A low-cost metal hydride of interest is TiFe; however, there have been performance issues. Recently, nanosized TiFe or TiFe alloyed with a third metal has been determine to perform favorably. Other potential candidates include $TiMn_2$ and Ti_2Ni alloys that also require fine tuning. Although a little more expensive, an interesting option is quasicrystals in the Ti-Zr-Ni system. In the end, we chose three materials systems with ternary compositions of Ti-Fe-Mn, Ti-Ni-Fe, and Ti-Zr-Ni. For our applications, three compositions in particular are promising ($TiFe_{0.8}Mn_{0.2}$, $Ti_2Ni_{0.9}Fe_{0.1}$, and $Ti_{0.8}Zr_{0.2}Ni_{0.1}$), which are known to operate at atmosphere pressures and near room temperature with rapid hydrogen diffusion.

Alloys that absorb hydrogen are typically prepared by arc-melting followed by heat treatment and rapid quenching. The synthesis route must be optimized to lower cost of materials production. Phase composition is determined by X-ray diffraction analysis. Currently, we are in the process of performing hydrogenation tests using a quantitative volumetric technique to learn the material's performance with respect to operation pressure, temperature, reaction kinetics, and hydrogen storage capacity. Knowing the kinetics and hydrogen capacity will allow us to select the most promising candidates.

By the end of this project, we will have down-selected the most promising LT-hydride candidates to be tested in the future with the HT-hydride. The discovery of a suitable LT-metal hydride composition will enable a path forward to realize the metal hydride high-temperature thermal energy storage for solar applications.

PN13084/2565

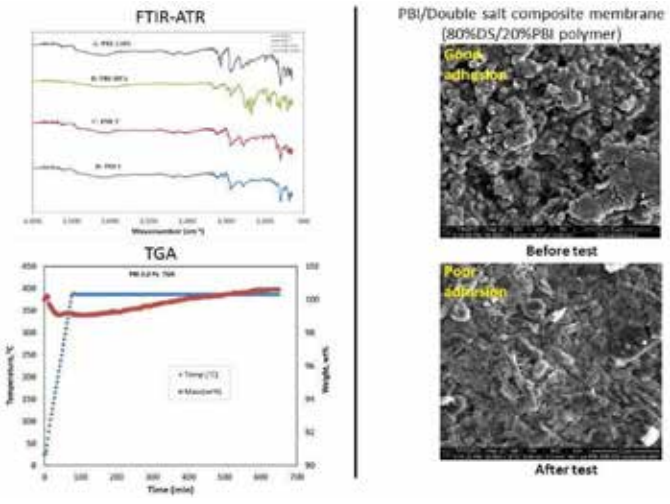
Novel CO_2-Selective Polymer/Double Salt Composite Membranes for Continuous CO_2 Removal from Warm Syngas

Rong Xing

We are developing a robust CO_2-selective double salt/molten phase composite membrane for CO_2 separation from warm syngas and incorporating the membrane with water gas shift (WGS) reaction to enhance CO conversion and H_2 recovery.

Current existing technologies for carbon capture from warm syngas are mostly based on solvent absorption that is energy intensive and requires inefficient heating and cooling of the syngas stream. The membrane separation process is considered as one of the least energy demanding processes if a good membrane that can fit the syngas process conditions is available. Unfortunately, there has been no membrane developed for practical CO_2 separation from warm syngas due to the limitation of existing materials. Although some inorganic membranes (e.g., metal, zeolites, and silica) show good separation performance, they are all H_2-selective and are limited by extremely high production costs.

We aim to develop a first-known CO_2-selective composite membrane that can separate CO_2 from warm syngas at an optimal temperature range of 300–400°C. Compared with a H_2-selective membrane, a CO_2-selective membrane has additional benefits: the generation of high purity H_2 on the high-pressure side, thus eliminating the recompression of H_2 for downstream uses; and a low membrane area that can be used. This is the first known attempt to develop a CO_2-selective double salt/composite membrane for CO_2 removal from warm syngas. Success of this work will provide the first step in PNNL advancing the state-of-the-art in composite membrane technology for pre-combustion fuel decarbonisation.



Characterization of synthesized polymers and polymer/double salt composite membranes

We conducted extensive studies to understand the relationship between membrane composition and its separation performance. A total of 16 compositions have been selected for membrane preparation. Our evaluation of these compositions in a membrane testing cell proved the concept that double salt can be developed into a functioning CO_2-selective membrane. Through our research, we determined that the amount of molten phase is a critical factor affecting membrane performance because the molten phase not only helps to facilitate CO_2 transport but also acts as a barrier to reduce or eliminate H_2 transport. Thus, we will continue the membrane screening tests for a wider structure spectrum to achieve the optimal composition.

Originally, we were to develop a robust polymer/double salt composite membrane to achieve our vision. The addition of polymer to double salt matrix was expected to seal the voids between double salt particles as a method for creating a pin-hole free membrane. However, our progress showed that the synthesized polymer/double salt composite membrane exhibited poor separation performance due to the oxidation of the polymer by molten $NaNO_3$, a promoter that secures high CO_2 sorption capacity and fast sorption kinetics of double salts. Our efforts were re-directed to develop polymer-free membranes. Significant progress was made: a self-supported CO_2-selective double salt/molten phase composite membrane was formed, and the membrane achieved our target performance temperature.

All of these findings are important to our project mission and have contributed to meeting our project goals. Specifically, the improved understanding of correlation between sorbent and membrane separation provided guidance for the membrane fabrication, selection of membrane composition, and testing conditions. The study of the membrane gas transport mechanism provided a fundamental understanding of the transport process through the membrane. Based on the mechanism, we can develop a separation model to simulate reaction and transport inside the membrane reactor. All new ideas are centered on how to achieve a CO_2-selective composite membrane with targeted separation performance. Based on these findings, we submitted two invention disclosures and will present the results in at a national conference.

Our plans for FY 2014 will cover multiple areas. We plan to: 1) investigate the role of the molten carbonate phase on membrane separation performance; 2) improve the membrane separation performance to meet our FY 2014 target; 3) standardize the membrane fabrication process that allows us reproducibly to make good membranes; 4) refine the gas transport model based on advanced material characterizations; and 5) incorporate a promising flat-sheet membrane into a WGS reactor to evaluate CO conversion and H_2 recovery.

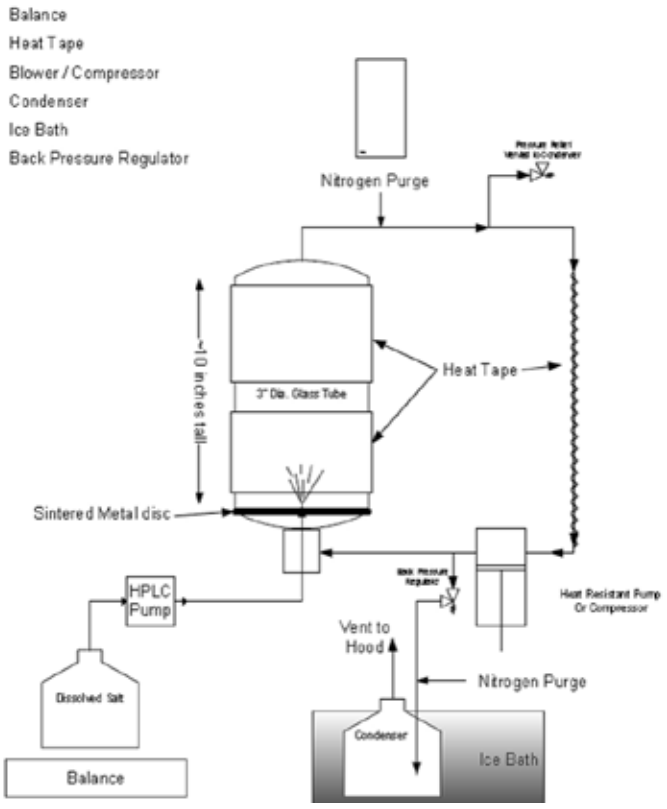
PN13006/2487

Novel Phase Selective (Drygel) Synthesis of Metal-Organic Frameworks

Radha K. Motkuri

This project increases possibilities for a continuous process of complex metal organic framework (MOF) synthesis using a fluidized bed approach that will enhance the synthesis efficiency from gram to kilogram scale.

During the last few years, MOF research at PNNL has been successful in developing new processes for gas sorption applications, including adsorption of CO_2, hydrocarbons, xenon/krypton, and recently, refrigerants (water and fluorocarbons) for efficient adsorption cooling applications. Many projects using MOFs require pure products with kilogram-scale synthesis for industrial-scale demonstrations. The general synthetic methodology for MOFs is similar to molecular sieve synthesis and usually involves hydrothermal or solvothermal crystallization of dissolved reactants in suitable solvents under conventional heating methods with autogeneous pressures. Recent reports of alternative methods such as microwave and sonication have issues with scalability, consistency, batch synthesis and long crystallization times. However, scaling up with the existing batch synthesis process not only takes time and effort but also produces challenges with reaction productivity and yield. Alternative synthetic approaches for bulk syntheses to enable MOF industrial applications are thus critical to the success of these applications.



Fluidized bed design for continuous MOF synthesis (construction in progress)

The objective of this project is to develop new methodologies for MOF synthesis using a solid-vapor method known as drygel. We will demonstrate a fluidized bed approach for the continual production of complex MOFs, which currently requires multiple synthesis and treatment steps. The drygel method was successfully adapted to the synthesis of three sets of MOF structures: microporous, the same with high open metal centers, and mesoporous. With success of this project, it is anticipated that we will create a new capability to perform scalable synthesis of MOFs, offer a high value product with excellent licensing potential, and create publications in high-impact journals.

During FY 2013, we successfully demonstrated the synthesis of nine MOFs: structures of HKUST-1, IRMOF-3, IRMOF-9, Zr-UIO66 (microporous), Ni-MOF74, Co-MOF74, Zn-MOF74 (high open metal centers), and mesoporous Cr-MIL-101 and Fe-MIL100. All structures were characterized with powder X-ray diffraction measurements and thermogravimetric analysis. In the case of MOF-74, two synthetic routes were successfully tested, with the THF (heating THF and water in a 1:1 ratio at 110°C for 3 days) resulted increased over all yield (70–80%). Surprisingly, the kinetics of the drygel resulted in decreased reaction time (from 72h to 24h) with comparable crystal morphology (PXRD) and chemical activity (gas sorption studies). Similarly, the yields of other MOFs observed increased (almost double that of conventional methods). In addition, we successfully demonstrated the synthesis of two MOFs (Ni-MOF74 and Co-MOF74) in a single pot reaction using a single set of solvents. This is a completely novel approach and will likely result in a new patent filing. The extension toward three or greater MOFs in single reactor is in progress.

The salient features of this project are aimed at goals that include Ni-MOF74 (~75% yield) in a 24 h reaction time (55–60% yield in 3 days = conventional); successful demonstration of various MOFs syntheses; an increased yield in Cr-MIL101, HKUST-1, and UIO-66; multiple MOF (Ni-MOF74 and Co-MOF74) synthesis in one batch reactor (the reaction was repeated twice); an attempt to synthesize MOF composites is in progress; and a new design and development of a fluidized bed approach for the continuous synthesis of MOFs (for bulk production).

In FY 2014, we will thoroughly characterize all MOFs and compare the activities with traditional synthesis, which will include peer-reviewed publications. We will continue the multiple MOF synthesis in a single batch reactor and construct and demonstrate the fluidized bed approach for continuous synthesis of MOFs (bulk synthesis).

PN13089/2570

Novel Window Coatings for Dynamic Thermal Control via Infrared Switching

Kyle J. Alvine

The goal of this project is to develop low cost dynamic infrared (IR) window coatings to control solar heating. If successful, this new technology is anticipated to save up to 1.2 quadrillion BTUs of energy or 20–30% of annual primary heating/cooling energy use.

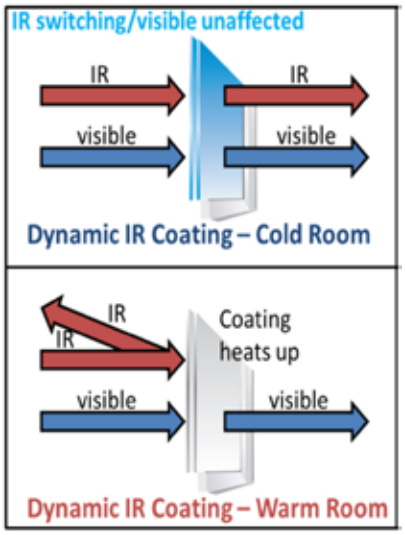
The ability to control the solar IR heating dynamically through windows has the potential for significant energy savings for buildings. Current state-of-the-art thermal management class windows are static low-e, electrochromic, and thermochromic. While low-e has a good market acceptance and low cost, potential energy savings are limited by the static nature of the coating. Electrochromics are highly effective at solar heat blocking but are handicapped by extremely high cost (approximately 16 times higher than standard double glazing), high installation cost, and simultaneous blocking of visible light. Thermochromics can also be highly effective and are lower cost compared to electrochromic (about 9 times higher vs. standard double glazing) but also significantly block visible daylight from tinting. The high cost and visible light blocking features significantly limit the marketability of this technology, which dismisses any overall energy savings. This information suggests that a novel dynamic IR coating that does not impact visible light and is low cost would be invaluable.

The objective of this project was to gather preliminary data to show the feasibility for a novel window coating technology that could dynamically transmit solar IR heat with only negligible impact on the visible light transmission. Within this project, we expected to gather feasibility data on individual components needed for the fabrication of the dynamic coatings. To demonstrate feasibility of the different components of the dynamic IR coating, we employed the combination of numerical modeling simulations and experimental fabrication work to prepare preliminary samples for measurements. In addition, microscopy was used to characterize samples, and preliminary optical data were taken to demonstrate feasibility.

The technology for this dynamic IR switching coating is based on the integration of novel optically responsive nanostructures with a thermally switchable order/disorder transition. These optically responsive nanostructures can be fabricated with a tailored wavelength response based on their size and other fabrication parameters.

Because the nanostructures are not classical optical structures, numerical modeling was needed to predict the wavelength response as a function of size and other fabrication parameters such as layer thickness. The numerical model solves Maxwell's equations at all points on a mesh defined over the nanostructure and yields both near-field wavelength response and far-field scattering predictions. Simulations were also carried out that showed a marked difference in the IR reflectivity with an order/disorder transition of the nanostructured coating.

In addition to the numerical modeling results, multiple laboratory scale (1 in by 1 in) nanostructured films were fabricated to generate preliminary data and determine the feasibility of these coatings for dynamic IR transmission. Coatings were fabricated with multiple different size nanostructures and reflection in the near infrared was confirmed for particular choice of fabrication parameters. In addition, preliminary feasibility studies on films with a temperature dependent order/disorder transition were performed. Multiple potential candidate materials were identified and reversibility was observed in some systems.



Schematic illustration of the dynamic IR coating technology. Only the IR transmission is affected as the coating heats up, allowing for blocking of IR during cooling periods and acceptance of IR during heating periods.

Additional work is needed to integrate the dynamic order/disorder and the optically responsive nanostructures to demonstrate proof-of-concept for the technology.

PN13034/2515

Optical Properties Modification in Complex Oxide Epitaxial Films via Alloy Formation

Ryan Comes

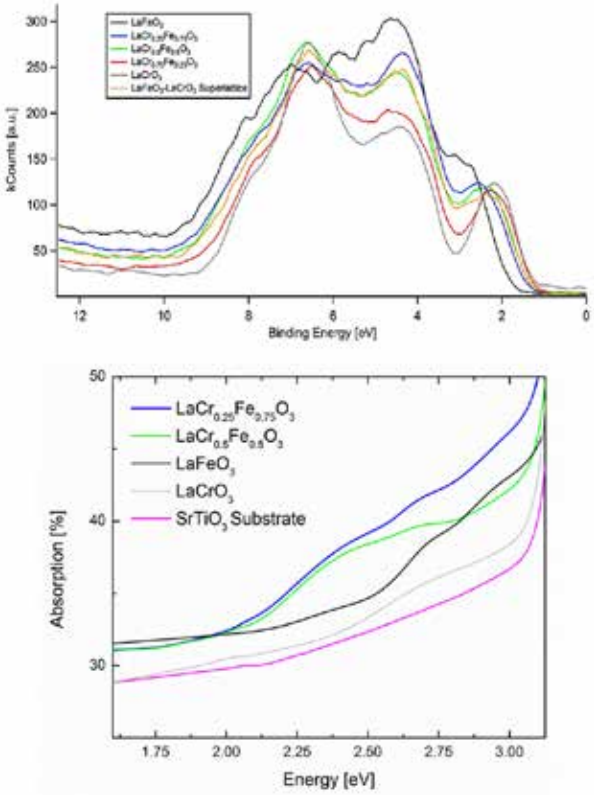
This research explores the optical and electronic properties of complex oxide thin films, examining their potential applications in both solar photovoltaic cells and hydrogen cells.

There is great interest in finding new materials that can harvest sunlight effectively for converting photovoltaics and powering photochemical reactions. Because the solar spectrum is more abundant in visible than in UV light, finding photoactive materials that absorb in the visible is of particular importance. Many recent popular photovoltaic materials use elements that are either toxic or rare; moreover, their surfaces form oxides in the atmosphere, which can alter their otherwise useful properties. An ideal class of materials for future photovoltaic applications is complex metal oxides, which are already stable in air and afford flexibility in designing and synthesizing tailored materials. Complex oxides also exhibit a wide range of optical and electronic properties. Commonly, they are used in solid oxide fuel cells but offer a wide variety of properties for potential use in electronic and alternative energy technologies. Grown epitaxially, these oxides constitute an ideal atomic scale toolbox for engineering designer photovoltaic and photochemical materials.

This project explores a variety of approaches to modulating the optical bandgap and electronic properties of perovskite oxide thin films. The work will focus on $LaCrO_3$ (LCO), as it represents an ideal starting point with an optical gap at the far end of the visual light spectrum. LCO is a perovskite oxide of particular interest for both its catalytic and optical properties and is commonly used in solid oxide fuel cells as a cathode for catalytic reactions. The use of LCO in fuel cells is fairly well understood, but prior work that studies its optical and electronic properties for solar cell applications has been limited. By replacing some Cr ions with various dopants, straining the films with different substrates, and growing multilayer structures, we will explore ways to modulate the optical bandgap. The altered optical and electronic properties of these multilayers will be characterized to evaluate their potential for use in oxide photovoltaics and solar hydrogen catalysis applications. These results may enable future alternative energy technologies that employ the novel properties observed in complex oxide materials.

With a fiscal year-end start, this project has shown promising initial results in its first months. Initial research has focused on $LaCr_xFe_{1-x}O_3$ (LCFO) epitaxial films and multilayers. We have observed new optical properties in the LCFO films that do not occur in either (LCO) or $LaFeO_3$ (LFO) films. Initial results on multilayer films are also promising, as we have successfully fabricated films comprised of alternating LCO and LFO atomic layers. This demonstration opens avenues for future research into superlattice films and interfaces.

Future work will focus on an in-depth analysis of multilayer and superlattice LCO and LFO films using the results of this fiscal year as a launching point. We will explore the properties of additional materials, such as $SrCrO_3$, which may have novel properties when strained as an epitaxial film. Other avenues of interest include doping of LCO and LFO films with Ti ions to inject free carriers into the system. This approach may lead to photocatalysis and solar cell research.



Top: Valence band X-ray photoelectron spectroscopy data for $La(Cr,Fe)O_3$ alloys, demonstrating that the alloys have broadened valence band features that may enhance optical absorption in the material and reduce the optical band gap; Bottom: Confirms the aforementioned hypothesis, showing optical absorption data which indicates that the alloy samples have a lower bandgap than either the $LaCrO_3$ or $LaFeO_3$ end members.

PN13100/2581

Optically Stimulated Luminescence Data Storage

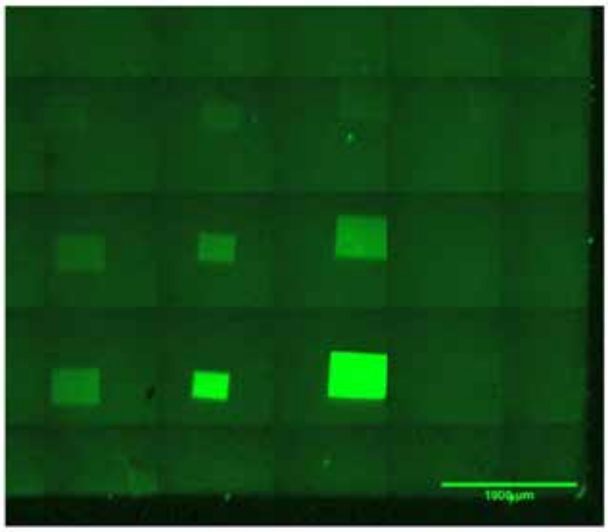
Jun Cui

We are developing a new approach using optically stimulated luminescence (OSL) that can be scaled to high data densities with excellent data lifetime and integrity.

To meet the ever-increasing data storage need for cloud computing, scientists and engineers are constantly pushing the limit of data storage density. Hard disk drive (HDD) has been the dominant computing data storage device for the past 50 years, with over 800 million units sold in 2012. Until recently, HDD density increased by 100% every year, but advancing technology has decreased this rate to 25%. Solid-state hard drive (flash) is the storage method based ferroelectricity, where direct wiring enables immediate access to the bit. However, flash suffers memory wear and thus can be used only as short-term storage. Although CD, DVD, and Blu-ray are popular optical data storage methods, they have a significantly lower density (under 10% of HDD). Still another recent advanced storage method is holography, a promising technology because of its large density and fast data access rate, reading a page each time instead of a bit (like flash). However, precision system setup and high recording medium cost make commercialization of this technology painfully slow and ineffectual.

We propose to develop a new data approach that is capable of making a paradigm change in storage technology. OSL has been in use for many years as a radiation dosimetry technology. OSL media has an extraordinary linear optical response (over seven orders of magnitude for LiF) that will enable ternary (higher-level) encoding of data in a single bit, thus storing more data in each defined location. The objective of this project is to prove the concept of OSL data storage by demonstrating multi-values data encoding. This information is a primary concern for potential development partners and will be key in identifying the potential market opportunities for this new technology.

Initially, we felt that OSL might be best suited for archival storage (currently magnetic tapes), where stability is a primary concern and where data are write once/read many (WORM) times. The "read" scheme for OSL storage is similar to current BluRay technology readouts, but the "write" requires use of ionizing radiation such as scanning ion or electron beams similar to CRT televisions. Because the readout is similar to Blu-ray, OSL is anticipated to achieve similar bit densities readily. In addition, with the blue laser spot size, the maximum areal density is ~12.5 Gb/in^2, a factor of 6 times higher than current magnetic tape (2.1 Gb/in^2). Adding multi-bit encoding ability, this level could increase to 50 or 100Gb/in^2, exceeding extrapolated predictions for magnetic tape for at least the next decade and possibly beyond theoretical stability limits of flexible magnetic tape technology. Further improvements in read technology, including near field optics to reduce the laser spot size, will enable additional storage density increases.



Multi-level encoding using SEM to write patterned boxes: nine distinct levels observed.

To evaluate the optical response range of OSL, we measured the optical emission from LiF/polymer dosimetry tags that had been exposed to a range of different ionizing radiation types and intensities. The samples were illuminated using light between 400 and 450 nm, and the emission spectra was measured from 500 to 1000 nm. A linear optical response was clearly observed over a dose range that varied by 20 times with a stable, repeatable intensity variation.

To understand the ability to pattern small areas, LiF crystals and films were exposed to e- and Ga-beam patterning using scanning electron microscopy (SEM) and a focused ion beam (FIB) system. The resulting patterned materials were then measured optically using standard fluorescence optical and confocal microscopes. We were able to pattern using e- and Ga-beams to a high resolution beyond the confocal microscope measurement limit (estimated between 100 and 200 nm) at the wavelengths of interest. This combination of linear response and high resolution patterning is a key for future work in developing this new approach to data storage.

For FY 2014, we will focus on understanding the commercial barriers to entry for a data storage technology based on OSL and develop a prototype system that demonstrates the potential of this approach.

PN13099/2580

Perfecting Atom Counting: 100% Efficient Mass Spectrometry

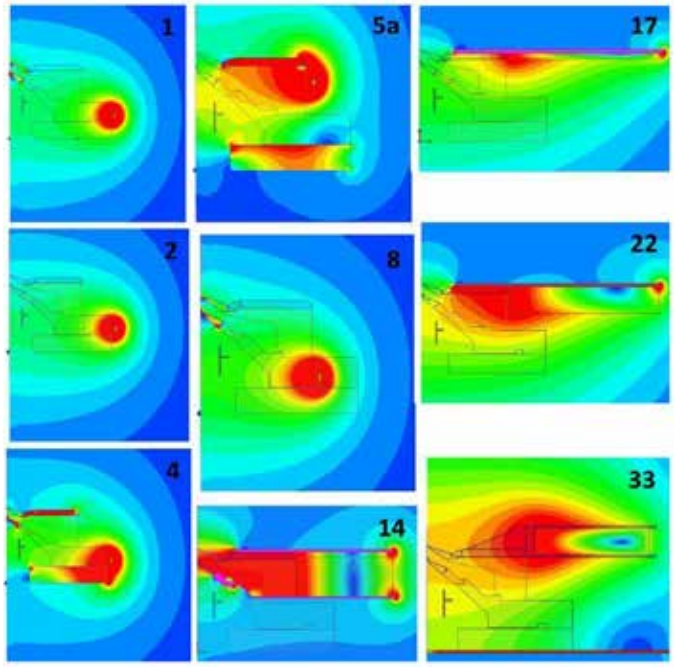
Gregory C. Eiden

The main focus in the early phases of the research is the sensitivity of fixed-laboratory instruments, a key PNNL and DOE interest.

Measurement of the elemental and isotopic composition of materials is critical to semiconductor manufacturing, geology, characterization and monitoring of the nuclear fuel cycle, and many other fields. The performance of such analytical methods can be determined by several factors; sample throughput and cost of analysis, data quality, and sensitivity are typically of high interest. This project seeks to improve dramatically the efficiency of such methods, thereby enabling sensitivity to be more broadly traded off against other design/performance metrics of interest: instrument size, power or consumables (fieldable vs. fixed laboratory issues), or both additional sensitive laboratory-based instruments and improved field-deployable instruments to be developed.

Our approach to increasing instrument efficiency is adapting methods known in high energy physics to plasma ion sources used for chemical analysis and materials processing (e.g., in an inductively coupled plasma mass spectrometry [ICPMS] or a semiconductor wafer ion implanter). Specifically in the ICPMS application, we are characterizing the effects of magnetic fields on a plasma as it flows from a high pressure discharge region through several stages of differential pumping into the high vacuum section of the instrument where the mass analyzer and detector are located. The research involves designing magnetic and electric field shapes that can achieve the desired high efficiency transport of plasma or ions. We are working in a region of pressure and plasma density where a rigorous analysis from first principles is not possible. Thus, our approach is using knowledge of plasma behavior and extraction to create likely geometries for efficient plasma sampling and ion beam formation and transport, model these proposed geometries, fabricate the best candidates, and observe how the fields interact with the plasma being sampled into vacuum.

A late FY 2013 start, our project began with a literature review on plasma physics and magnetic lens design and performance, performed magnetic field modeling for ~40 different candidate design cases for our test instrument, down selected and fabricated a first prototype magnetic lens, installed the prototype lens onto a quadrupole ICPMS, and began characterizing the lens performance. Results to date include measurement of instrument response as a function of analyte mass, magnetic field strength and properties of key components in the plasma sampling interface and, most recently, pressure effects in the first vacuum stage of the instrument. In portions of the data, we observe signals to vary with the square of the magnetic field strength, as predicted in our earlier work.



False color map of the magnetic field strength for nine of the ~40 designs modeled, illustrating different choices of solenoid size, location, and incorporation of magnetic and non-magnetic materials for key parts that affect field shape along the of the ion beam optics centerline. 2D models are run initially as a survey of likely candidates followed by more time-consuming modeling in 3D for the most promising candidates. The last image shown for "case 33" (lower right corner) is the solenoid design fabricated as part of the first prototype lens.

Thus far, our most important finding is that different ions have completely opposite dependencies on magnetic field strength. Specifically, the argon dimer cation (Ar_2^+) shows an initial decrease in measured signal with increasing magnetic field; however, after reaching a critical field strength, the signal increases rapidly. By contrast, analyte signals decrease with an increasing magnetic field but do not increase at higher fields. The behavior of this magnetic lens design compared with the lens we built in 2002 demonstrates the complexity of these experiments: the mass resolved signal measured at the back end of a mass

PN13093/2574

spectrometer has many "long" lever arms in the front end ion optics that influence observed signals in complex ways. While the magnetic field in one region of the instrument might have a beneficial effect, the overall effect might be dominated by the effect of a field in an adjacent region.

The magnetic field in our prototype lens exists in three distinct regions of the interface: in the atmospheric pressure plasma, the first vacuum region where the plasma is expanding into a vacuum of ~1 Torr, and the second vacuum region where the expanding plasma has been "skimmed" and is formed into an ion beam. The effect of the magnetic field on the flowing plasma or ions is distinctly different in these three regions. Other important insights concern a better understanding of methods for shaping the magnetic field in the critical regions of the ICP vacuum interface by choice of the geometry of the field forming components as well as their magnetic properties.

In FY 2014, we will complete characterization of the first lens, review and analyze the results alongside a re-assessment of the 2002 project's results, and re-consider any relevant literature. This analysis will lead to the design of a second magnetic lens or device. In particular, we will "divide and conquer" by attempting to separate magnetic field effects in each of the three interface regions (ICP, expansion chamber, and downstream of the skimmer). Our goal is to obtain a better understanding of magnetic field effects on charged particle and plasma transport, mostly via experimental observations. This improved understanding will then be used to design an improved lens and/or vacuum interface.

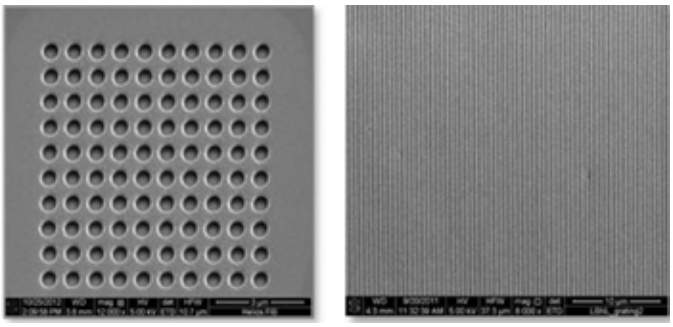
Photocathode Development for Next-Generation Light Sources

Wayne P. Hess

Our project goal is to produce new photocathode designs to enable construction of fourth-generation light sources at significantly reduced cost and enable development of dynamic transmission electron microscopy.

The international light source community recognizes the great need for a more scientific approach to new photocathode development and is in the initial stages of addressing this issue. Whether based on free electron laser (FEL) or storage ring designs, the ultimate output characteristics and cost of next-generation ultraviolet (UV) or X-ray sources are heavily dependent upon photocathode brightness and emittance characteristics. Novel photocathode designs could potentially reduce light source construction costs dramatically by significantly simplifying downstream accelerator or FEL design. The fields of materials science, solid-state photochemistry, and surface science can and should make immediate and timely contributions to this essential activity.

To meet the challenge presented above, this project will develop robust photocathode sources with intense electron yields and ultralow thermal emittance. We expect to create several new photocathode sources using inventive semiconductor, mixed-metal, composite, and plasmonic designs. Specifically, we will directly address photocathode durability by using robust metal designs with or without protective insulator coatings. The project will first develop innovative photocathode materials and designs, such as nanograting or nanohole arrays using novel metal and metal insulator hybrid materials. Second, we will demonstrate high electron bunch intensity, low thermal emittance, and operational durability required by next-generation light sources.



Nanohole array FIB milled in 50 nm silver thin film on mica substrate (left) and a gold nanograting produced by nanolithography techniques (right).

We made considerable progress in our project over the last year. Specifically, we published measurements about photoemission enhancement from a novel gold nanograting and constructed a silver nanohole array using focused ion beam milling. We integrated a continuous wave 266 nm laser with the ultraviolet photoelectron spectrometer (UPS) and acquired nanolithography software for a helium ion microscope (HIM) located in EMSL. The HIM nanolithography capability will allow us to construct nanoplasmonic photocathode designs at will and in-house. In addition, we measured the properties of candidate photocathodes, including electron yield (bunch intensity) and emittance (electron angular distribution). Two plasmonic photocathode designs were created by nanolithography and focused ion beam patterning. We also conducted a study of the growth of a hybrid MgO/Ag(001) photocathode using He I and laser UV and angle resolved photoemission spectroscopy to complete a project milestone.

In late FY 2013, we confirmed the hypothesis that the quantum efficiency enhancement of metal alkali halide hybrid (Cu:CsBr) is due to work function lowering and defect formation in the alkali halide thin film. There is no indication that that Cs metal forms on the photocathode surface during activation or contributes to the enhanced quantum efficiency. Previous work did not use Cu substrates that were cleaned to surface science standards, and we found that scrupulously clean Cu has an enhanced QE compared with “technically clean” Cu photocathodes. We also find that the work functions reduced from that of clean Cu by over 1 eV, and very large electron yields were produced under 1 mW cw laser irradiation. The UPS results confirmed the predicted enhancement in quantum yield. Additionally, theoretical modeling of Cu:CsBr began in earnest. The electronic properties and energies of the hybrid systems will be calculated using embedded cluster density functional theory (DFT) and time-resolved DFT. The theory can predict the kinetic energy maximum of Br atoms desorbed following UV excitation. A comparison of both the theoretical and experimental results will benchmark the accuracy of the modeling of electronic properties of these candidate photocathode systems.

To validate our thin film growth results, experiments continue to be performed in FY 2014 on Cu:KBr. Using angle-resolved UPS, we will measure the thermal emittance to demonstrate the utility of hybrid materials. We will additionally use modeling and photoemission measurements to determine the role of intraband defects in metal-supported alkali halide and metal-supported MgO systems.

PN11061/2383

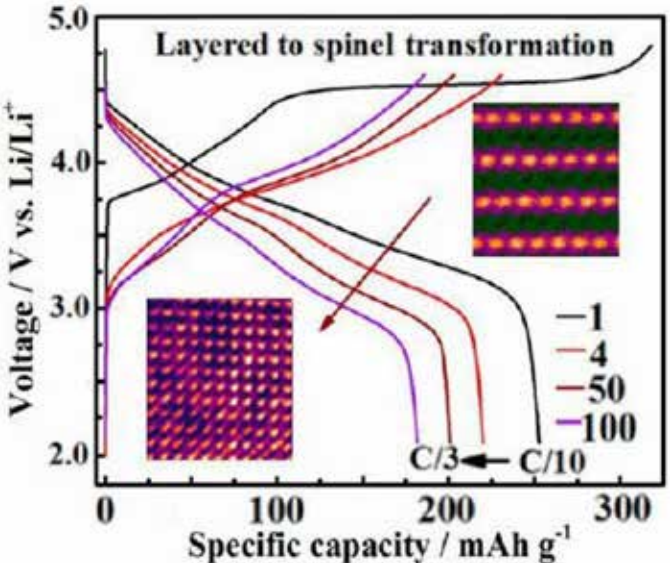
Probing Structure-Property Relationship of Energy Storage Materials Using *Ex Situ* and *In Situ* Dynamic Microscopy and Spectroscopy with High Spatial and Fast Temporal Resolution

Chongmin Wang

The specific aim of this project is to probe the structure-property relationship of next-generation energy materials to enable accelerated discovery of new energy materials for energy security and sustainability.

Electrochemical energy storage devices such as Li-ion batteries are complex, multi-component systems that incorporate widely dissimilar materials and phases in physical and electrical contact. Despite the success of lithium (Li)-ion batteries, there are at least three areas where new research and development progress is needed: discovering new materials that can be used as electrodes (both for cathodes and anodes) with higher storage capacity; tailoring the microstructure of electrodes and the overall architecture of the battery cell to enable high rate operation (fast charging); and resolving the capacity fading of the battery (i.e., increasing the cycle life). Research is progressing in several areas to improve Li-ion battery performance.

The main objective of our research is to develop and use a combination imaging, spectroscopy, and tomography methods to probe into the structural and chemical evolution of energy storage materials with atomic spatial and fast temporal resolution to determine the structure-property relationship and its correlation with charge and ion transport in energy materials. In particular, these chemical imaging methods will explore the new chemistry and nanostructured materials (such as Li-S and Li-Air) for energy storage with high capacity, high power, longer cyclability, and safe operation. This goal will be accomplished by developing concerted multi-scale chemical imaging methodologies based on *ex situ* and *in situ* TEM, STEM, DTEM, electron energy loss spectroscopy (EELS), energy-dispersed X-ray spectroscopy (EDS), and STXM and X-ray tomography.



Electrochemical properties of lithium-rich nickel manganese oxide and their correlation with structural phase transformation upon battery cycling.

In 2013, we achieved several project milestones. We extended our open cell *in situ* capability for the study of materials beyond the Li-ion battery. Typically, we have an *in situ* capability for studying materials for the sodium and magnesium ion-based energy storage concept. This capability is now the key component that enables the research effort of the DOE-sponsored Joint Center for Energy Storage Research (JCESR). Beyond the open cell *in situ* capability, we successfully implanted the wet electrochemical cell, allowing us to use real battery relevant liquid electrolyte and obtain information normally unavailable using the open cell configuration such as the SEI formation, structural and chemical evolution, and their effect on battery properties. We used this liquid cell for the first time to study the charge and discharge behavior of silicon.

In accordance with our high spatial resolution instrumentation capability, we are also developing fast temporal resolution instrumentation: dynamic TEM. This capability will enable nanosecond temporal resolution and is currently under testing stage. We teamed with FEI, a leading TEM company, to demonstrate the power of the four detector based EDS system for 3D chemical imaging of a single nanoparticle. This work is published in *Ultramicroscopy*, is one of most downloaded papers, and is highlighted on the cover of the journal. At the same time, we enhanced the capability that enables imaging of a sample across the platform of FIB, APT, S/TEM, and photon source capability, which is key for imaging a specified site of a sample across different instrument platform and yields complementary information. These new capabilities enable us to

PN13010/2491

probe energy materials under the reaction conditions and capture complementary information across different instrument platform.

With our *in situ* and *ex situ* capabilities, we probed the scientific questions related to anode and cathode materials of the Li-ion battery. In the anode, we focused on silicon-based material, intending to determine why silicon failed and how can we smartly design high capacity material base on silicon. In theory, silicon has an energy storage capacity of ~15 times of the graphite that is currently used in rechargeable batteries for electronic devices. However, when inserting lithium, silicon will expand ~300%, which will cause the short life cycle of battery. Using our *in situ* TEM capability, we studied the lithiation behavior of silicon. We are collaborating research activities with leading industries to understand the fundamental properties of the materials they are developing for the battery of electric vehicles.

For cathode materials, we are collaborating with Argonne National Laboratory to probe the capacity and voltage fading mechanism of lithium rich nickel manganese oxide materials. $Li_{1.2}Ni_{0.2}Mn_{0.6}O_2$ is a high voltage and high capacity cathode material for Li-ion battery. However, a major challenge of this material is the capacity and voltage fading upon cycling of the battery. We used high resolution imaging and chemical elemental mapping to identify the cause of the capacity fading, which yielded three publications.

In the new chemistry for batteries, we studied the Li-S system in detail. The Li-S and Li-PS flow-type batteries are promising candidates for portable and large-scale energy storage, respectively. However, the poor stability of polysulfide species during battery cycling is a major challenge that severely hampers this technology. This instability leads to "shuttle phenomenon," where the polysulfide molecules dissolve and diffuse to the Li^- electrode and causes dendrite growth through parasitic reaction. The difficulty in addressing this issue mainly originates from the significant gap between battery material synthesis efforts and actual knowledge about the molecular structure and stability of dissolved polysulfide species. To unravel the structure and possible stability related reactions of polysulfide species, a combination of density functional theory (DFT), magnetic resonance, and X-ray spectroscopy studies were performed for the Li-S system. A paper based on this work will be soon published.

In 2013, our work covered by this project has leaded to five peer-reviewed journal publications, three invited talks, and five conference presentations. We were the leading organizer for a chemical imaging symposium, and we received monies to study cathode and anode materials in 2014. At the same time, the lead of this project will be heavily involved in the focused research group activity of high capacity anode material based on Si.

Looking into 2014, we will focus on using *in situ* and *ex situ* imaging methods to study the energy storage materials such as Li-ion battery, S-ion battery, multi-valence ion batteries, andLi2Sn, which are crucial in the cycling performance of the battery. We will closely integrate different capabilities and theoretical calculations and collaborate with the research effort of JCESR, other projects, and industrial organizations to push the research effort on energy storage to a high level for scientific and industrial applications.

Rare Earth-Free Phosphors for Lighting Applications

Carlos A. Fernandez

The replacement of down conversion rare earth (RE)-based phosphors will reduce market demand and U.S. dependence on foreign sources, decrease environmental impact, and enable the widespread adoption of solid-state lighting products.

Nitrides and oxynitrides have remarkable thermal and chemical stability and good fracture toughness while exhibiting unusual photoluminescence (PL) when activated with RE ions. As for the application of RE-free based nitrides such as Mn^{2+}-doped phosphors, most of these investigations have focused on Eu^{2+} and Mn^{2+} co-activated materials. In this regard, we are conducting synthetic and modeling work to target low-cost, stable RE-free phosphor materials based on alkaline earth (Ba, Ca, Sr, Mg) Mn-substituted aluminum nitrides and silicon nitrides efficient down converters. Ultimately, the full replacement of RE in phosphor materials will positively impact the U.S. energy capacity by accelerating the production and application of inexpensive energy efficient lighting products.

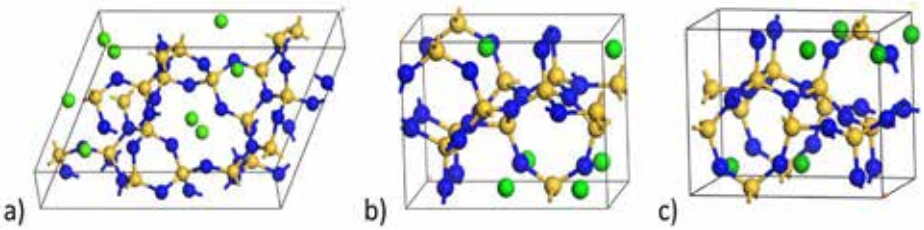
The outcome of this project is to establish a set of basic tools to correlate composition-structure/function relationship followed by demonstrating the materials prospective to replace RE-based phosphors. For FY 2013, our work focused on two tasks: performing molecular modeling to understand the relationship between phosphors structure/composition and their optical properties on Mn-doped SiN- and AlN-based phosphors; and performing structure/composition and PL properties analysis of best candidate phosphor materials discovered in Phase I. The main goal behind this approach was to demonstrate the prospect of the aforementioned materials to replace RE-based phosphors.

Down-conversion phosphors consists of a host matrix that absorbs light in the near UV-blue region of the spectrum via an energy transfer processes to a dopant present in the crystal structure to emit visible light. The color (or wavelength) of the emission mainly depend on the host matrix composition, the nature of the dopant, and its concentration.

By means of molecular modeling, the crystal structures of three SiN-based host materials were calculated. Specifically using hybrid functionals (HSE) with screened exchange (as these functionals have been demonstrated to yield band gaps in better agreement with the experiment), we measured the theoretical and experimental optical band gaps for the parent $Ca_2Si_5N_8$, $Sr_2Si_5N_8$, and $Ba_2Si_5N_8$ materials in eV showing to be in agreement with experimental work. The valence band in these materials is dominated by the N-2p states, while the conduction band is dominated by the p states of the alkaline earth elements (Ca, Sr, Ba).

To simulate the low concentrations (~6%) of the dopant Mn^{2+} ions in the host lattices, we used large unit cells (120 atoms). The optimized structures for Mn^{2+} replace the alkaline earth metal of parent compounds $Ca_2Si_5N_8$, $Sr_2Si_5N_8$, and $Ba_2Si_5N_8$ with a mole percent of 6%. The theoretical optical bandgaps were compiled for the three compounds with the Mn^{2+} dopant (6%). The lowering of the band gap compared with the pure host material is from the presence of Mn 3d states and hybridized N-2p and Mn-3d states in the gap. With the predicted crystal structure for the host and doped materials, these detailed band structure calculations will be published. Additionally, we calculated the optical activity of AlN-based phosphor material (α-$Ca_3Al_2N_4$). The band gap of this parent material is calculated to be ~3.3 eV (376 nm), which is more favorable compared with the aforementioned SiN-based systems from an optical activity standpoint.

Mn^{2+} doping of this material at 17% lowers the band gap or absorption edge to ~2.5 eV (500 nm), while doping with 9% lowers the absorption edge to ~2.6 eV (477 nm). The lowering of bandgap in this Mn^{+2}-doped AlN phosphor material is (as in the case of the formerly described SiN-based phosphors) due to the appearance of new hybridized N-2p and Mn-3d states in the band gap. As part of the next phase of this project, we will further characterize this system experimentally and theoretically and study the optical activity at various Mn^{2+} dopant concentrations.



Crystal structures of a) $Ca_2Si_5N_8$, b) $Sr_2Si_5N_8$, and c) $Ba_2Si_5N_8$

Considering theoretical considerations, we chose to substitute silicon in the moderately emissive Ca_3N_2:Si_3N_4:Mn system with Al. Synthesis conditions optimized for the Al-based system enabled formation of emissive species that performs significantly better than both the literature

PN13085/2566

precedents of alternative materials to RE-based phosphors and our own previous efforts.

The novel compound obtained by the reaction of Ca_3N_2 and AlN with large excess of Mn (25 mol %) can be obtained by heating at 1800°C for 5 hours and exhibits two orders of magnitude stronger luminescence signal than AlON:Mg:Mn system, which was one of our best phosphors synthesized in the previous year. It is a strong orange emission with a maximum at 595 nm and full width at half maximum (FWHM) of 38 nm. The features observed in the excitation spectrum prove that the obtained compound has its emission properties closely linked to Mn doping. However, the exact Mn concentration in the sample has not yet been determined but is estimated to be low because analytical methods such as ICP-MS or EDX could not detect its presence in the analyte. Though these are preliminary results that indicate that further advances in achieving high purity phases and composition control in the target materials will be required, the materials developed in this project already demonstrate the potential of the chosen approach to discover inorganic non-RE highly emissive materials that consist of abundant inexpensive and non-toxic elements.

Finally, it has been reported in the open literature that materials such as $M_2Si_5N_8:Eu^{2+}$, which possess similar host matrix composition while doped with the rare element Eu^{2+} show excitation in the blue to green spectral region showing excellent temperature stability with remarkably high PLQY (above 90%) even at temperature as high as 200°C. Based on our calculation estimates, we expect the new material (α-$Ca_3Al_2N_4$) with Mn^{2+} doping to have about 65–70% greater PLQY compared with reported state-of-the-art RE-based phosphor such as $M_2Si_5N_8:Eu^{2+}$. Mn-doped α-$Ca_3Al_2N_4$ was also synthesized, and its emission spectrum measured showed significantly high PL, orders of magnitude higher than any other RE-free phosphor developed in this project and by others.

For FY 2014, we will modify the synthesis setup to achieve better control over reaction conditions. From the reaction product composition standpoint, with the aid from theoretical predictions and based on our recent findings, we will be able to discover other materials with enhanced emission properties. The final outcome for next year will be the development of a library of RE-Free phosphors of various compositions and PL properties demonstrating the color tunability of the new materials proposed.

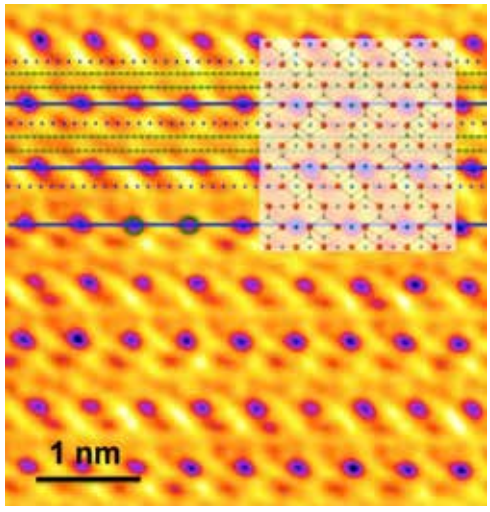
Site Specific Atomic Resolution Probing of Structure-Property Relationship Under Dynamic and/or Operando Conditions Using *In Situ* and *Ex Situ* Chemical Imaging Based on Multi-Instrument Approach

Libor Kovarik

The goal of this project is to develop chemical imaging capabilities that enable an understanding of the structural and morphological evolution of catalytic materials under* in situ *and* in operando *conditions.

Catalysts and catalytic technologies are of great interest to the applied and basic scientific research due to their enormous economic and technological importance. To advance current catalytic materials and develop new catalytic technologies, it is essential to develop basic principles that relate the structure property relationship and form a basis for theory informed design. New capabilities that can capture and visualize the working catalysts under elevated temperature and environmental dynamic operating conditions are critical in describing the basic principles that govern the microstructures-property relationship, and thus play a critical role in the overall research effort.

Transmission electron microscopy (TEM) is one of few techniques that can directly visualize catalytic materials at the atomic level. With recent advancements in aberration-correction instrumentation, the TEM has proven especially indispensable for gaining increasing insight into the atomic nature, with the resolution and sensitivity of visualizing individual atoms. However, being traditionally a high vacuum technique, TEM studies have been often limited to work under *ex situ* conditions. Performing TEM experiments under realistic conditions in recent years led to advancements with *in situ* TEM capabilities, and new *in situ* holders and environmental microscopes have been developed. Despite this progress, challenges remain to improve, customize, and develop tools that push environmental capabilities and resolution limits under dynamic operating conditions. In addition, it is important to enable cross-platform capabilities, thereby providing complementary information.



Atomic level depiction of γ-Al_2O_3 after *in situ* heat treatment. In this HAADF image, the dark tones represent a high value of projected potential. The derived crystal structure of δ_1-Al_2O_3 is overlaid on the image.

The goal of this project is to develop TEM-based imaging capabilities that raise atomic-level imaging and spectroscopic analysis of catalytic materials under elevated temperature and environmental conditions. This will lead to a new knowledge that can be applied to much needed development of new materials for heterogeneous catalysis. The goal would be realized by extending the state-of-the-art aberration-corrected ETEM and STEM complementary microscopy and spectroscopic imaging capabilities to a new level, enabling atomic-resolution imaging and spectroscopic analysis under dynamic operating conditions.

In particular, an important part of our research in FY 2013 focused on establishing a new *in situ* heating capability for atomic-level imaging at high temperatures. As part of this effort, we installed a MEMS-based heating holder with a double-tilt capabilities and heating up to 1200°C. The double-tilt is essential for achieving a precise orientation of catalysts and catalytic materials for atomic imaging and represents an upgrade from our previously installed single-tilt heating unit from FY 2012. We additionally developed new procedures and methods for transfer of catalytic materials on thin membranes of the heating chips.

Specifically, our work showed that thermal treatment leads to complex intergrowth of several phases. We developed a crystallographic approach that enabled us to derive unambiguously the crystallographic nature of two main ingrowth variants of Al_2O_3. In combination with DFT methods, we established how energetic degeneracy leads to the structural complexity in transition aluminas. As a result, it can now be justified how high temperature exposure leads to disorder and intergrowth among Al_2O_3 phases and

PN11070/2392

alumina. This work, which represents an extension from our prior work in FY 2011 and FY 2012, has important implications for understanding thermodynamic stability, surface properties, and catalytic behavior, and suggests a modified view for transformation processes in transition Al_2O_3 phases.

Another substantial part of this project focused on developing a new capability for atomic-level imaging with the environmental transmission electron microscope (ETEM). This work represents a continuation from last year's effort, where we successfully installed the ETEM and demonstrated high resolution imaging of catalytic and energy related materials under gas pressures up to 20 mbar. In FY 2013, we developed and built a gas control unit for ETEM that could deliver accurately controlled gas mixtures and allow seamless switching between gases while maintaining atomic resolution imaging. In the existing setup, switching between gasses leads to variation in gas pressure and therefore a drift on the sample and loss of atomic resolution. The new unit incorporates pressure and mass flow controllers for mixing up to 10 gasses, nine of which are permanently attached and one that is an inter-changeable specialty gas. The unit is fully computer controlled and achieves accurate mixing of gasses and seamless switching while maintaining a constant gas pressure.

Employing atomic-level in situ investigation at 400–600°C and O_2 pressure of several mbars, we identified steps in the early stages of Pd oxidation. We found that oxidation is initially associated with a strong faceting, followed by the formation of a 2–3 layer surface oxide. Under oxidizing conditions, the layers appear to be highly stable; experimental evidence shows that layers are consistent with the PdO-like phase formed in an unforeseen orientation. Switching between oxidizing and reducing gasses enabled study of the hysteresis effect associated with oxidation and reduction cycles. We found that the PdO surface oxide shows a different structural nature under reduction environment, thus clarifying the known hysteresis of oxidation and catalytic performance associated with Pd atoms. The ETEM was also applied to study Pd clusters support by high surface area alumina. At the scale of only few nms, we identified a different oxidation mechanism with complete dissolution of Pd clusters, leading to the formation of highly dispersed Pd clusters.

In parallel, we performed an ex situ study focusing on understanding of thermal stability of noble metals supported by $MgAl_2O_4$. We employed a multimodal approach of atomic resolution STEM imaging and beam-line EXAFS measurements (performed at APS). The combined results enable us to quantify how these heterogeneous metal catalyst species, ranging from single atoms to nano-scale clusters, bond with surface of the substrate. The multimodal approach provides new insight on the unusual high thermal stability of Ir on the spinel support and provides guidelines for a design of calcination and reduction treatments.

We also used electron tomography combined with state of the art reconstruction algorithms to understand a catalyst before and after reduction. This was accomplished by using an *ex situ* reduction chamber and a vacuum transfer holder. The catalyst imaged was a Co based Fischer Tropsch (FT) catalyst. We were able to show in 3D that upon reduction, the cobalt oxide moves out of the interior pores and out towards the surface of the alumina support. The quantification of the tomograms shows that the surface area remains constant for the alumina support after reduction (which we used as a control since it should remain constant), but that the cobalt oxide surface area reduces slightly. The reasons for that are because of the mass loss during the reduction, as well as the redistribution of the cobalt to agglomerate with larger particles near the surface of the alumina support. To our knowledge, this was the first direct 3D imaging of the redistribution of the oxidized catalyst upon reduction and imaging of the catalyst in its active metal state.

In summary, the work of this project established important new capabilities for catalytic research under high temperature and environmental conditions, and the application provides a new level of understating for a number of heterogeneous catalyst materials. In total, the work produced four peer-reviewed journal publications, one invited talk, and 11 conference presentations. Several other publications are currently being prepared for submission. For the future, the newly developed expertise and capabilities will be applied to address stability of catalytic surfaces and their interaction with noble metal clusters and nanoparticles under another project.

Towards Understanding Interfacial Chemistry at Reactive Solid/Liquid Interphases

Maria L. Sushko

This project develops a new computational technique for a deep understanding of the interfacial processes and designing new materials for energy technologies, including safer, longer lasting batteries.

Reactive solid/liquid interfaces are ubiquitous in nature and energy technologies. Mineral surface CO_2 reactions are central for carbon sequestration, heterogeneous catalytic processes in mesoporous membranes that drive energy production in fuel cells, and electrochemical processes at the electrode/electrolyte interface that determine battery performance. These processes are inherently pathway dependent and strongly influenced by confinement, flow conditions, temperature, and external fields. Quantum processes such as bond breaking/formation and electron transfer at the reaction site are directly influenced by the liquid phase structure and the densities and diffusivities of charged species at hundreds of nanometers. However, there are currently no truly multiscale models describing reactivity at solid/liquid interfaces from atomistic to micron scales.

This project aims to develop the first hierarchical multiscale code for theoretical modeling of pathway dependent reactions at solid/liquid interfaces. Our goal is to produce a comprehensive open source software compatible with popular codes for quantum mechanical simulations with sufficient functionality to be applied to a wide range of reactive solid/liquid interfaces. The range of problems for which this computational method is relevant include electrochemistry at solid/liquid interfaces, reactive transport through the mesoporous medium, morphology control of crystals through electrochemical synthesis, and crystal growth through self-assembly of nanoparticles or prenucleation clusters.

Our FY 2013 research activities were conducted along three main directions: developing a mathematical formulation for linking quantum and mesoscopic length scales; optimizing the performance of the mesoscopic classical Density Functional Theory/Poisson-Nernst-Planck (cDFT/PNP) code; and applying software to understand the mechanism of electrodeposition of Li^+ onto Li metal electrode from mixed salt solutions.

First, we developed a formulation for seamlessly linking quantum Kohn-Sham (KS) DFT and cDFT based on an embedding approach for classical charge density, providing an embedding electrostatic potential for the KS equation for electron density of the quantum region, which in turn provides an external electrostatic potential for the classical region. The resulting system of coupled KS and cDFT equations can be solved using freeze-and-thaw.

We developed efficient numerical algorithms for solving 3D steady-state PNP equations with excess chemical potentials described by cDFT. The equations are discretized by a finite difference scheme and solved iteratively using the Gummel method with relaxation; the Nernst-Planck are converted to Laplace through the Slotboom transformation. The algebraic multigrid method is then applied to solve the Poisson and the transformed Nernst-Planck equations. A novel strategy for calculating excess chemical potentials through fast Fourier transforms reduces computational complexity from $O(N^2)$ to $O(N \log N)$, where N is the number of grid points. Integrals involving the Dirac delta function are evaluated directly by coordinate transformation, which yields more accurate results compared to applying numerical quadrature to an approximated delta function.

We applied the software to study the mechanism of Li^+ electrodeposition onto Li metal electrode in the presence of additives. The aim was to understand how additives can protect the electrode from dendrite growth, which jeopardizes battery safety. The fundamental issue we addressed was understanding the specific ion effect on the balance of weak interactions between the solvated cations and anions at the rough electrode surface. We studied the energetics in mixed solutions of $LiPF_6$ and X^+Y^- salts with $X^+ = Na^+$, K^+, Rb^+, Cs^+ and $Y^- = PF_6^-$, BF_6^-, ClO_4^-, TFSI-. We showed that cations with weaker electrostatic correlation interactions (effective repulsion between cations) preferentially accumulate near the protrusions, preventing Li^+ deposition at the protrusion sites. These cations induce Li^+ deposition next to the protrusion resulting in smoother electrode surface. Simulations show that small concentrations of Cs^+ and Rb^+ can efficiently protect electrode surface from dendrite growth while Na^+ and K^+ do not, in good agreement with experimental data. The reason larger cations are preferentially adsorbed at high potential protrusion sites is in the inverse scaling of electrostatic correlation interactions with ionic diameter. Anion size also influences the effective repulsion between cations: smaller anions screen cation charge, more efficiently reducing the cation-cation electrostatic correlation energy. These simulations not only revealed the origin of specific ion effect in electrodeposition but also demonstrated the applicability of the developed computational approach to complex processes at reactive solid/liquid interfaces.

PN13044/2525

Understanding and Control of SEI Layer (Solid-Electrolyte Interface and Interphases) in Multivalent Energy Storage Systems

Guosheng Li

By demonstrating elegant electrochemical performances of new electrolytes, this project will achieve a breakthrough of magnesium (Mg) battery for energy storage applications.

Key benefits of multi- over mono-valent (i.e., lithium [Li] or soldium [Na]) systems for viable low cost energy storage include increased electrons available per molecule, which significantly impacts energy density characteristics. Potential candidates include Mg (which can lose two electrons) and aluminum (can lose three electrons) coupled with increased abundance and decreased production cost. The coulombic efficiency of Mg electrochemistry can be as high as 100%, which further improves the practical energy storage capacity compared with the Li/Na system. For cost, Mg metal is 24 times less expensive than metallic Li, along with its safety characteristics is particularly attractive as a candidate for cost sensitive energy storage applications. No dendritic formation was reported on Mg metal anodes during the battery operation process, which is critical for applying metal anode instead of metal carbon composites. Li batteries using Li metal anodes suffer from thermal runaway as well as cell failure due to the internal shortage.

Extensive electrochemical and spectroscopic studies have been conducted to understand the role of solvents for the Mg reversible plating and stripping. For example, diglyme (DG, tridentate ligand) could be a donating ligand for Mg to improve the coulombic efficiency of Mg plating/stripping. More importantly, the $Mg(BH_4)_2$/DG mixture is a good model system in understanding the molecular structures in Mg complex electrolytes. Thus, developing a novel Mg electrolyte, which facilitates 100% coulombic efficiency and demonstrates its performance in coin cell configuration, are two primary project objectives.

In FY 2013, a new, safer electrolyte based on $Mg(BH_4)_2$ and DG was developed. $LiBH_4$ was employed as an additive because it has been shown to increase performance in the case of $Mg(BH_4)_2$/DME. Extensive cyclic voltammetry (CV) measurement has been performed to identify coulombic efficiency (CE) for reversible Mg plating/stripping. 100% of CE was determined for the electrolytes, which consist of $Mg(BH_4)_2$/DG with by adding appropriate amount of $LiBH_4$ as the assisting electrolyte. The XPS spectra recorded after Mg plating on Pt electrode clearly show plated Mg, and Mg completely disappeared after stripping. XRD and EDS results also confirm that only Mg is plated with no sign of Li deposition and no carbon element, indicating no electrolyte decomposition.

To understand coordination structures of $Mg(BH_4)_2$ in DG and correlate them with performance, THF, DME, and DG solvated $Mg(BH_4)_2$ complexes were isolated and characterized by ^{1}H NMR and ^{11}B NMR spectroscopies in non-coordinating CD_2Cl_2, which does not interrupt the structures of the solvated $Mg(BH_4)_2$. To identify the composition of $Mg(BH_4)_2$ in various solvents, extensive NMR studies were performed. In the ^{1}H NMR spectrum, the solvated $Mg(BH_4)_2$ species shows proton resonances at 3.96 ppm (CH_2), 3.85 ppm (CH_2) and 3.61 ppm (CH_3) for the coordinating DG. As expected for the shielding effect of DG metalation, proton resonances of the coordinating DG are downfield shifted compared with free DG in CD_2Cl_2 (3.56, 3.49 and 3.33 ppm, respectively). Due to different coupling interactions with ^{11}B (I=3/2) and ^{10}B (I=3) isotopes, two sets of hydride signals, a major quartet (J_{BH}= 80 Hz) and a minor septet (J_{BH}= 30 Hz) were observed at the same chemical shift (-0.36 ppm). Integrals of proton resonances give a 1:2 ratio of DG vs. BH_4^-, which leads to a ratio of $Mg(BH_4)_2$: DG=1:1. Thus, $Mg(BH_4)_2$ in DG is formulated as seven coordinated $Mg(BH_4)_2$DG, a great difference with the composition in DME and THF. Mg complexes exist as a dimer composition $Mg_2(BH_4)_4(DME)_3$ and a monomer composition $Mg(BH_4)_2THF_3$ in DME and THF solvents, respectively. Clearly, the lower CE observed for THF and DME compared with DG is from the different Mg complexes from the complicated characteristic coordination in each environment.

A rechargeable Mg battery with a Chevrel phase Mo_6S_8 cathode, an Mg metal anode, and $Mg(BH_4)_2$-$LiBH_4$-DG electrolyte is assembled. The cell delivers an initial capacity of 99.5mAh/g (Mo_6S_8 only) at a C/10 rate (Mo_6S_8 theoretical capacity is 122 mAh/g); the capacity drops slightly for the first a few cycles and is stabilized for the remaining cycles with ~89.7% capacity retention for 300 cycles. Two patents related to this technology are in process.

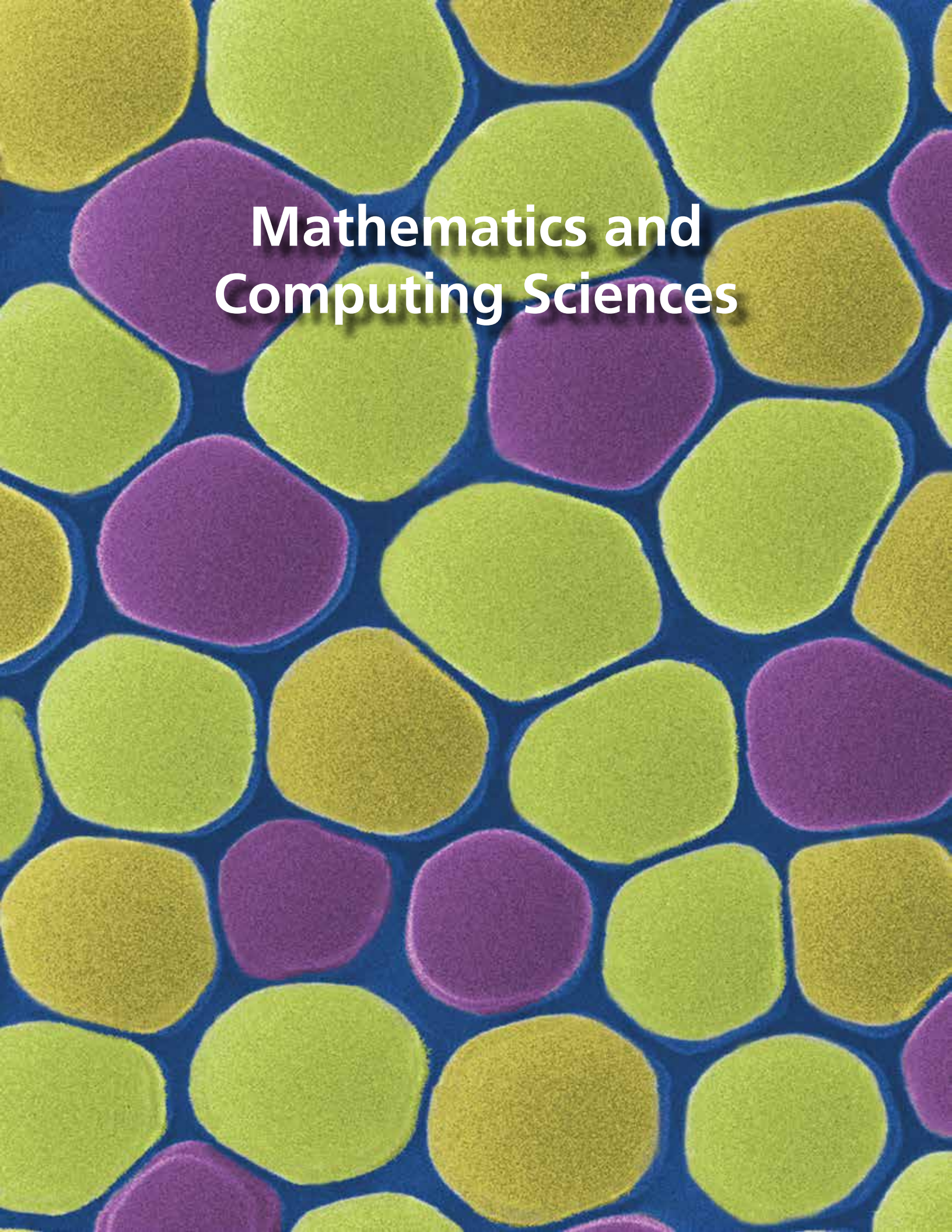
In parallel with the research effort for a $Mg(BH_4)_2$-based Mg electrolyte, the $MgCl_2$-based electrolyte was studied and developed in FY 2013. The preliminary works on $MgCl_2$ validated the rationalized synthetic approach, termed a mono-

PN13050/2531

chloride abstraction, to yielding a series of the Mg^{2+} dimer electrolytes characteristic of exceptional oxidation stability (up to 3.4 V vs. Mg), sulfur compatibility, high current density (up to 32.7 mA/cm^2), and reversible Mg^{2+} ion plating and stripping (up to 100% coulombic efficiency). A patent for this technology is currently pending.

The most important findings of FY 2013 are the two Mg electrolytes that can facilitate reversible Mg plating/stripping on Mg metal electrode. $Mg(BH_4)_2$/DG electrolyte was demonstrated in a coin cell using Mo_6S_8 as cathode. We believe that the Mg battery consisting of $Mg(BH_4)_2$/DG/Mo_6S_8 has a great potential for the large-scale energy storage application. The new $MgCl_2$-based electrolyte has a wide electrochemical window (up to 3.4 V vs Mg), which is essential for developing higher energy density Mg battery for portable device and transportation vehicle applications.

For FY 2014, we will develop novel rechargeable Mg batteries using $MgCl_2$ based electrolyte, which will present higher electrochemical and chemical stability compared to the current state-of-art. The proposed cathode materials include various transition metal oxides, which have a higher capacity, cell voltage, and energy density (>400mW/g). The success of this project will lay the groundwork for the commercial implication of an Mg-based batteries system electric vehicle application.

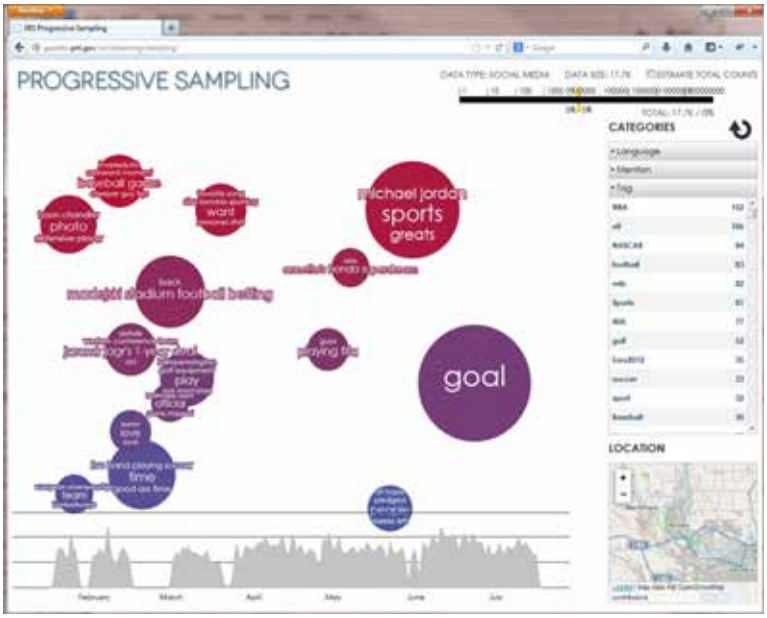

Mathematics and
Computing Sciences

A Compressive Sampling Framework for Interactive Visualization of Massive Datasets

Scott T. Dowson

Identifying key artifacts, patterns, and trends in massive datasets through a visually interactive platform can be used to prevent and counter terrorism.

In combatting terrorism and protecting our nation's infrastructure where the volume or rate of data to be analyzed is only growing, alternative methodologies must be developed that facilitate the analysis of all information, not just subsets due to time or capability restrictions. The object of this project is to develop and demonstrate an effect approach that can quickly analyze massive datasets to identify and visualize significant concepts, trends and patterns in the data. The result of this endeavor will be an extended software framework capable of integrating with massive data stores with dynamic models that analyze continuously streams of sampled data and interactive visualizations that update reflecting the current state of the models, providing the means to assess and explore the content of the data collections.



The prototype dynamic web application showing estimated contents of the sampled dataset.

One key aim for this research was not just to test and evaluate sample-based analytic models and their effectiveness in identifying salient features within massive datasets, but to accomplish this through a reusable, extensible framework that could serve as the basis for future research as well as application development. We accomplished this task by leveraging an existing visual analytic framework, the Scalable Reasoning System (SRS) and extending it to meet the demands for a streaming sample-based analytics platform.

To integrate generically with a massive data provider, classes of interfaces were designed to provide a continuous stream of sampled data to the analytics engine. For our test and evaluation, two datasets were obtained, collections of up to 1.04 billion social media records and 514 million cyber network logs. Each collection was managed on a file system and a streaming data adapter was developed to provide random samples of data, with replacement, from each dataset. Future adapters could be easily developed to sample data from a database or web application.

Sampled data streaming into the analytics framework are routed to various analytic models. Classes of interfaces were designed to accept a continuous stream of sampled data. These models are designed to dynamically update with each sample and on request to provide a current snapshot of the state of the model that can then be represented in interactive visual environment. Three models were developed to demonstrate and evaluate the effectiveness of sampling: a temporal model capturing the distribution of data and features, a categorical model identifying major facets and attributes, and a concept text model that extracted significant keywords from unstructured text and identified themes in the data consisting of sets of highly correlated keywords. For the temporal and categorical models, the data was completely analyzed and modeled providing ground truth as to the true distribution of the temporal and categorical data. Successive snapshots of the models were then taken at various orders of magnitude of sampled data, from 10^3 to 10^7 samples, ultimately representing approximately 2 percent of the overall a 515 million social media evaluation dataset. The estimated distributions of data across the temporal bins as well as instances of various categorical data were evaluated at these various snapshots.

The estimated temporal distribution was calculated across 215 temporal bins spanning 15 weeks derived from a preliminary sampling of the data to estimate the temporal extends of the dataset. The estimated distribution was calculated based on sampled counts for each bin based

PN13095/2576

on the proportion of the data sampled. The error for the distribution was based on the difference between the estimated and actual proportion for each bin on the perceived differences in a visual representation of the distribution. The overall average bin error was calculated at successive sampling snapshots and was shown to converge to a sub-pixel (indiscernible) level after 1 million samples (0.2% sampling rate).

The estimated distribution of major categorical properties were calculated for the top 10,000 most frequently occurring values based on sampled counts and proportion of data sampled. A 95% confidence interval was calculated based on the standard error of a proportion for each categorical value. The 95% confidence interval fell below 1% error after sampling 10 million samples (2%) for the most frequently categorical value (occurring in 0.7% of dataset records).

Major topical themes in the sampled data were derived first by extracting significant keywords from the sampled records based on the Rapid Keyword Extraction (RAKE) algorithm. A graph was then modeled with the nodes representing each extracted keyword from all sampled data, and edges connecting keywords when they co-occurred within a sampled document. At the time, a snapshot was requested for the top themes (a set of highly correlated keywords), with the most significant keyword pairs taken from the graph and keywords hierarchically clustered based their similarity as computed by the Jaccard similarity coefficient. Top clusters of keywords are the estimated top topical themes in the dataset.

Each model's estimates and uncertainty are visualized in a web-based application. New visual metaphors were designed to not only convey each model's estimation of the contents of the massive dataset, but also the calculated uncertainty of those estimates. From the rate at which data was being sampled and modeled (~ 300,000 samples/min) and the rate successive snapshots were generated (every 5 sec), the visualization components were designed to support continuous streams of model snapshots. To demonstrate the viability of this approach, the components were integrated into a prototype web application that dynamically updates to convey the estimated contents of the massive dataset.

The main conclusion of this project was the finding that with sampling a very small fraction of the overall dataset, meaningful and accurate estimates of the temporal distribution, major categorical attributes, and significant topical themes can be identified. Further, the extensible framework designed, developed, and evaluated on a high-end desktop platform demonstrated that these insights can be discovered in massive datasets in near real-time without the need for capital investments into distributed computational resources.

A Distributed Systems Architecture for the Power Grid

Adam S. Wynne

This project demonstrates the efficiency of a distributed system architecture for decentralizing power models in a large-scale power grid. The architecture eliminates the bottleneck of a centralized location for an overloaded data managed monolithic computation, an increasing burden to fast-growing data rates and quantity.

The substantial growth of high quality sensors provides a comprehensive view of the entire connection of the power grid. With this sensor data and high performance computing (HPC), a revolution in the operations of the power grid, namely real-time operations, becomes possible. The key to leverage these benefits is keeping up with the pace of high frequency samples generated from these sensors. Fast distributed monitoring and analysis applications are increasingly important to power system operation and control. A distributed systems architecture testbed for evaluating power system distributed applications is critical to plan the deployment of HPC platforms to connect their decentralized electric power applications and form the guideline of distributing power applications to meet real-time requirements of critical operations and control in the electric power system.

Our previous implementation has the peer-to-peer data exchange through TCP sockets. Therefore, the data communication logic has been hard-coded in the power application. This mechanism has limitations for validating distributed power models (such as the state estimation algorithm), as the study of algorithm behavior requires that different partitions of the same power grid network be tested. The partition may change how the sub-areas are connected, hence peer-to-peer interactions. Hardcoding the data communication in the state estimation code is not sufficiently flexible to represent the procedural steps of the distributed state estimation.

Our research project further enhances the distributed systems architecture that we developed in FY 2012 and introduces a service-oriented layer that contains a workflow environment for composing distributed procedural steps, observing data exchanges individually, automatically launching remote jobs running on HPCs. This layer also links to the coordination and fault-tolerant middleware, as coordination logic is developed in the middleware. The architecture is implemented as a middleware platform that coordinates distributed procedural behavior and adapts to runtime computing conditions to improve reliability, maintain causality, and guarantee responsiveness. The middleware is equipped with a modeling tool to estimate how the distributed power model is affected by factors of data communication, geographical locations, and computing resource demands. As a result, this research facilitates the design of cost-effective situational awareness applications within a large-scale power grid.

The deliverables of this project form a testbed to configure decentralized power applications in an evolving distributed power gird for real-time analysis. Our major accomplishment from is a software demonstration deployed on the PowerNet Lab testbed that encompasses:

- scalable middleware that, beyond connecting distributed code, coordinates the dynamic data communication according to various synchronization criteria
- an orchestration framework that monitors a step-by-step data exchange
- a prototype of distributed state estimation with PMU data
- a testbed environment leveraging the laboratory
- rigorous simulation/analysis tool on end-to-end delay estimation.

In FY 2013, the service layer was improved and made production-ready by leveraging and incorporating the Laboratory Integration Framework and Toolset (LIFT), an integration framework and service-oriented platform that was used to augment the GridOPTICS project with the ability to share data sources and execute analytics across organizational boundaries. A demonstration was produced in conjunction with the GridOPTICS and SharedPerspectives projects that demonstrated the sharing contingency analysis data across two simulated organizations. These data were modeled using two separate instances of DSA and GridOPTICS feeding event information to two instances of the SharedPerspectives user interface.

Additionally in FY 2013, work in previous years on semantic data discovery was extended by incorporating standards for data types, a standard for representing services, and a demonstration illustrating the ability to discovery data across power organizations. Specifically, the Common Information Model (CIM) was used to represent data types, and the Unified Service Description Language (USDL) was used to represent web service properties. A demonstration was implemented that illustrated the ability to issue a complex semantic query to locate data in adjacent regions.

PN11002/2324

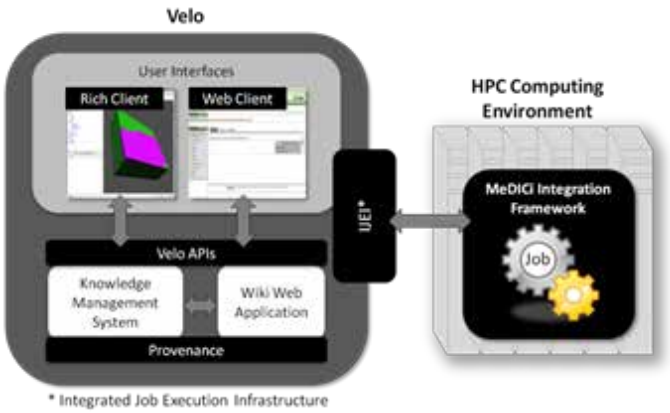
A Multi-Modal Integration Framework for Chemical Imaging

Kerstin Kleese van Dam

This project delivers an open multi-modal analysis framework for chemical imaging to allow the fast and precise analysis, comparison, and integration of experimental results from several imaging technologies against the background of exponentially growing data volumes, acquisition rates, and data complexity with the increasing need for real-time analysis.

Existing approaches today favor one off, customized solutions to data acquisition and analysis. This unfortunately requires extensive resources to develop and maintain new tools in a fast evolving technology domain, as it prohibits sharing of development efforts and tools among communities. PNNL is proposing a new approach consisting of highly customizable infrastructure and reusable analytical software components, enabling communities to create their optimized analysis environments with significantly reduced effort levels, while utilizing highly optimized solution approaches.

The core components of the new PNNL experimental analysis environment are Velo (collaborative data management and analysis infrastructure), MeDICi (data intensive workflow system) and REXAN (component library for Rapid Experimental Analysis). The framework is being developed to support multi-modal imaging with particular emphasis on real-time analysis to allow experimental steering and correlation of results across different modalities. The project is building on existing efforts such as capabilities for data acquisition, annotation, storage (ICAT, Velo, and MeDICi), and extending and augmenting these for real-time analysis of large data volumes through the use of high performance computing (HPC) facilities.



Chemical imaging architecture combines MeDICi Integration Framework and Velo infrastructure.

Velo is a reusable, domain independent, eScience environment and knowledge management infrastructure for modeling and simulation. Velo leverages, integrates, and extends web-based open source collaborative and data management technologies to create a scalable core platform that can be tailored to specific science domains. Velo helps scientists to organize experiments, computations, and large and complex datasets used in scientific modeling. Currently, it is deployed in domains as diverse as carbon sequestration, climate modeling, chemical imaging, and environmental remediation.

The new architecture that we are developing integrates Velo with the MeDICi workflow engine for running complex pipelines to process and transform instrument data into visualizable datasets. MeDICi is a middleware platform to integrate data- and compute-intensive processing components with data sources and visualization tools across multiple scientific domains. MeDICi is connected to Velo through the use of Velo's Integrated Job Execution Framework (IJEF), which seamlessly connects MeDICi running on an HPC server to the knowledge management server and user interface. It automatically stages all required input files, submits the MeDICi workflow job, monitors the running workflow – updating status information in real time – and performs simulation-specific post-processing operations such as filtering results or streaming results (or remote links to large data sets) back to the knowledge management system for integration with analysis tools.

Key to our project is the development of library of reusable analytical components. We contended that it is possible to describe most experimental analytical pipelines through a limited set of conceptual workflows that have a range of recurring components that can be reused across a range of conceptual workflows. Therefore, a library containing highly optimized versions of those reusable components would help with the accelerated development of new analytical pipelines as well as raise the overall speed, scalability and performance of the resulting pipelines. In FY 2011, the project identified the initial core components for such analytical pipelines in collaboration with the initiative researchers and stakeholders and began developing two test analysis capabilities. During FY 2012, we analyzed the results of our initial analytical tools and created the

PN11001/2323

initial REXAN library of reusable analytical components. The integrated Velo/REXAN system was then used to create a range of initial analytical pipelines and tools now used in production by the scientists:

- PNNL High Resolution Mass Spectrometry – reduces analysis times from hours to seconds and enables much larger data samples (100 KB to 40 GB) at the same time
- ALS X-ray tomography – reduces analysis times of combined STXM and EM data collected at the ALS from weeks to minutes, decreases manual work, and increases data volumes to be analyzed in a single step
- Multi-modal nano-scale STXM and TEM data analysis – provides a semi-automated process for particle detection.

In FY 2013, the project expanded both its infrastructure and analytical library capabilities:

- The **MSIQuickView** tool was created for interactive real time analysis and post-hoc analysis of mass spectrometry data. New capabilities such as a new user interface that allows users to visualize and compare different data sets at the same time as well as easy, memory efficient browsing of thousands of images in tandem. New image clustering capabilities were added to aid the exploration of large volumes of images in addition to 3D visualization capabilities for 2D images for the same purpose.
- The **BioFilmViewer** tool was created and extended to enable the manipulation and analysis of X-ray images of biological samples. The tool leverages existing capabilities in other tools such as ImageJ, Chimera, MATLAB, Cell Profiler, Drishti, Vaa3D, Icy, BioImageXD, and Paraview into one easy-to-use environment.
- **Remote near real-time analysis capabilities** were created using Velo, BioFilmViewer, and GlobusOnline. The new remote analysis capability allows data to be streamed directly from the instrument back to the scientists home organization, allowing use of their own analytical pipeline to reconstruct and analyze the results as the experiment is progressing. This setup enabled users to make critical decisions about the experimental set up while present at the facility, increasing the experimental success rate and the quality of the resulting experimental data.
- **Community–based collaborative analysis** infrastructure prototype was created for the PREMIER transition electron microscope (TEM) network. The system allows different sites with TEM capabilities to share their imaging data and their analytical approaches, enabling more in-depth analysis of common components through the availability of more varied data on the material or biological substance. The infrastructure will also be used to provide more advanced analytical capabilities over time based on the REXAN library.
- **A new approach to multi-modal analysis** was researched as part of the project, investigating the use of semantic descriptions for analytical components to enable the automated inference of valid analytical pathways for specific data integration challenges. The project extended the existing VisKo system for this purpose and demonstrated initial success in describing and sharing new visualization and analysis methods.

The project increased its outreach activities in FY 2013 and had notable success in promoting its concepts to the DOE funding agencies through invited presentations (BES Facilities Directors meeting and participation in two ASCR program development activities for experimental facilities and accelerated scientific knowledge discovery). The project was also successful in becoming the only pilot collaboration to apply advanced computing capabilities to high resolution coherent imaging of energy materials at light source facilities. Additional proposals have been submitted to the NIH and EPA utilizing the REXAN capability to build advanced analytical capabilities. In addition, project members contributed to four papers, one book, and two online journal articles in FY 2013.

A New Modeling Approach for Biology: Combining Natural Selection and Thermodynamics for Biodesign and Natural Systems

William Cannon

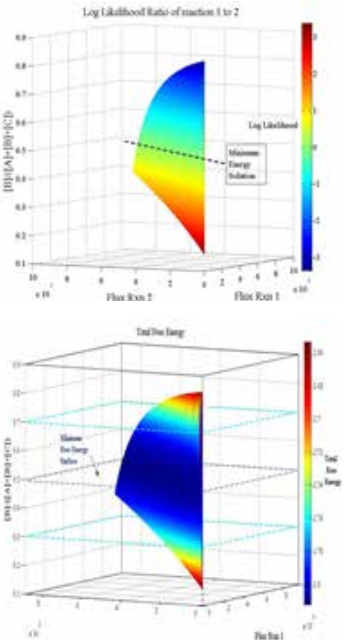
This research is using thermodynamics to understand natural selection and biological adaptation so that changes in the environment can be predicted and national policies can be developed to mitigate the economic consequences of these changes.

Ideally, the modeling of metabolism would use kinetic simulations, but these models require knowledge of thousands of rate parameters, the measurement of which is very labor intensive. The best option currently is to model reaction fluxes only, without regard to the law of mass action. However, this convenience also limits the predictive power of the methods. Not only does the range of solutions span many orders of magnitude, but also metabolite levels and energetics are next to impossible to predict from flux values.

We are developing an alternative approach based on rigorous statistical physics. Previous work in this area has focused on irreversible processes and small model systems of between two and four reactions. The objective of this project is to apply the statistical physics of reversible processes to microbes and to evolve the system by selecting of the most-fit microbes based on free energy production. Our ultimate goal is to integrate statistical physics and physical chemistry with the study of biological processes, thereby enabling much more predictive simulations that can inform research decisions and strategies in the laboratory.

To accomplish the goals of this project, we considered in detail the relationship between the probability of a change of state in a system of coupled reactions according to kinetic formulations and compared these kinetic probabilities to the probability of a change of state according to statistical thermodynamics. Both approaches are based on the law of mass action. The fundamental difference is this: kinetics describes the *difference* between the rates (likelihood) of individual reactions, while thermodynamics describes the *ratio* of the rates (likelihood) of reactions. This distinction has important consequences for the probability of a change of state due to a chemical reaction: the calculation of the probability of a change of state due to kinetics requires a recursive solution to a master equation, while the calculation of the probability of a change of state due to thermodynamics does not. In fact, the Markov chain resulting from the thermodynamic formulation is a Markov Random Field, which has special desirable properties that make computation of the state dynamics much easier.

When using any modeling approach, it is important to understand the assumptions involved in the model. We have shown that the probabilities of changes of state calculated from statistical thermodynamics assumes that all chemical species are distributed according to a Boltzmann distribution, which is the same assumption used in Transition State Theory for modeling chemical reactions. However, when used for modeling reaction intermediates, this assumption is not as severe: unlike transition state species, chemical intermediates in a reaction pathway are stable compounds that can be isolated. The assumption of Boltzmann probabilities is tantamount to assuming that the system is not a source or sink of heat. This assumption makes the timescales over which the reactions occur statistically independent of each other.



Top: Total system free energy of the reaction system $A \rightarrow B \rightarrow C$. The minimum total free energy occurs when the compounds are distributed according to a Boltzmann distribution. Bottom: Likewise, the most thermodynamically stable steady state occurs when the distribution is Boltzmann.

We have shown that heat loss/gain from the environment can affect steady-state behavior significantly. For a system of coupled reactions such as

$$A \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} B \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} C$$

a steady state occurs when the net flux of $A \rightarrow B$ equals the net flux of the reaction $B \rightarrow C$. When the flux of each reaction is plotted, the steady state solution is a line along the diagonal. However, many

PN13056/2537

steady-state solutions are possible, depending on the value of reaction rate constants. For example, rate constants can be varied for each reaction above from 1 s^{-1} to 10,000 s^{-1}; subsequently, the flux for each equation can be calculated by solving the appropriate differential equations. Different rate constants result in different steady-state concentrations of the intermediate, *B*. Because concentrations of the initial reactant *A* and final product *C* are constant, the differing concentrations of the intermediate correspond to different system pressures, ranging low to high. Only one solution corresponds to ambient pressure, and this is the thermodynamically reversible process that involves no heat sources or sinks. The solution that corresponds to the thermodynamically reversible process is the minimum in the free energy surface. In other words, coupled reactions that occur on significantly different timescales result in thermodynamically irreversible processes, thus losing or gain heat from the environment and consequently are not at a thermodynamic minimum. In other words, reactions that are sources or sinks of heat do not represent thermodynamically stable systems. Although they can occur under ambient conditions, reactions require fluctuations from the most stable state to proceed.

The above knowledge is important because first, reaction rates in biological species are not rigidly fixed but are mutable and subject to natural selection or "survival of the fittest." The work required to maintain a metabolic state is likely a significant contribution to fitness: the less work required for maintaining a steady state the more work that is available for reproduction. Clearly, reactions in metabolism have evolved to occur on the same timescale; the degree to which this has approached reversibility is a matter for future studies. However, the assumption of no sources or sinks of heat – use of Boltzmann probabilities in simulations – may be quite appropriate for modeling biological systems.

Our initial work developing the mathematics for non-equilibrium simulations of state for biological systems has been submitted to the *Biophysical Journal*, and the mathematical approaches have been implemented in Matlab. Further, these methods are being implemented in high performance computing code Boltzmann on this and other projects. We expect to publish an application note for the methods and code in FY 2014. In addition, this work was orally presented at two conferences during the last year. Funding on this project has also supported the use of high performance computer code that to run millions of Boltzmann simulations in parallel. The HPC code Biocellion will be used for optimizing metabolism and modeling natural selection. A report describing the application of Biocellion to modeling microbial systems is in preparation and was carried out in collaboration with the Shou Group at the Fred Hutchinson Cancer Research Center in Seattle.

We will continue to pursue our long-term goal of using the principles of thermodynamics to understand natural selection and biological adaptation. Work during FY 2014 will result in several publications that have been in the development this year. We will be submitting an additional applied math paper on the modeling methods, a modeling application on yeast interactions using Biocellion and an application note for Boltzmann. In addition, we will perform the first combined use of the Boltzmann and Biocellion codes to model simple representations of microbes and further investigate the quantized energy level phenomena. We expect the latter two studies to also result in publications in FY 2014.

PN13056/2537

Actionable Visualization Tools For Power Grid Situation Awareness

Luke J. Gosink

This project addresses the fundamental need for greater situation awareness (SA) within the grid domain through the design of actionable visualization strategies that increase inter-organizational communication and collaboration.

A sustainable, efficient power grid operation relies heavily upon real-time information transparency and wide-area SA for all organizations in the grid. Without real-time, system-wide awareness, there is a significant potential for small local power instabilities to escalate and cascade into large-scale blackout events. While many organizations have developed methods to support internal SA capabilities (i.e., awareness of state within the organization), there is little evidence of successful research that can help organizations see across organizational seams to increase external SA.

This project supports the need for external SA by designing a new visual communication mechanism that will allow organizations to share information securely across organizational boundaries. We will extend this mechanism with new visual analytic strategies that incorporate information outside of the power domain (e.g., weather, political/social, cyber, and other entities) so that organizations have a better awareness of events that can potentially impact grid stability. We will support access to these functionalities through a new application, Shared Perspectives, that will allow power system operators to achieve broader a SA of events both within and outside of the power grid. Project progress is based on two fundamental objectives: the creation of a master specification document for fostering SA that identifies use-cases, challenges, and needs of operators when they attempt to share information during critical events; and the design and implementation of the Shared Perspectives visualization application that directly addresses challenges identified within the master specification document.

During FY 2012, significant progress was made in compiling the master specification document. We conducted numerous interviews with organizations (including representative cases of utilities, balancing authorities, and reliability coordinators) to identify the need for external SA. Additionally, we attended several workshops and seminars to identify current approaches used to generate this awareness.

Based on our initial research, we identified that there are some organizations lacking in methods that support external SA. Operators communicating across organizations often depend exclusively on telephones to create a foundation of common understanding. For non-critical situations, this medium is sufficient; however, growing evidence indicates that during critical

Shared Perspectives Application

complex events, relying only on phone communication creates a significant opportunity for miscommunication with potentially disastrous consequences. To address this challenge, our research identifies the need to broaden the communication conduit through a common tool that supports audio and a common visualization perspective where organizations can "push" and share information to help build common understanding.

From our findings, we initiated design work for Shared Perspectives based on the need to build common understanding flexibly and securely between organizations through information sharing. The application is based on a scalable web architecture that supports a highly interactive collection of visualization components. This set of components enables the rapid design of custom applications for industry use-cases. During FY 2013, we completed the design and implementation of the Shared Perspectives application and demonstrated it to subject matter experts from several utilities. The demonstration depicted a use-case that involved four organizations that used the application to resolve conflicts caused by a downed intertie. The new components of Shared Perspective include a new visualization strategy for power data that depicts data streams as an interactive "theme river" view that consolidates data from different sources and highlights critical trends for interactive exploration and collaborative problem solving.

In FY 2014, we will integrate Shared Perspectives with the PNNL's Grid Optics software suite. We will continue industry engagement to further our research on the master specification document. In addition, we will conduct a user study to assess project benefits that will be officiated by North American Electric Reliability Corporation-compliant trainers who will use Shared Perspectives (under a newly issued copyright) to assess the importance of shared visualization in collaborative problem solving. A final task will be to design and perform a validation study on the effectiveness of Shared Perspectives to reduce time-to-resolution for various events.

PN12003/2404

Advanced Optimizations for Extreme-Scale Homogeneous Systems

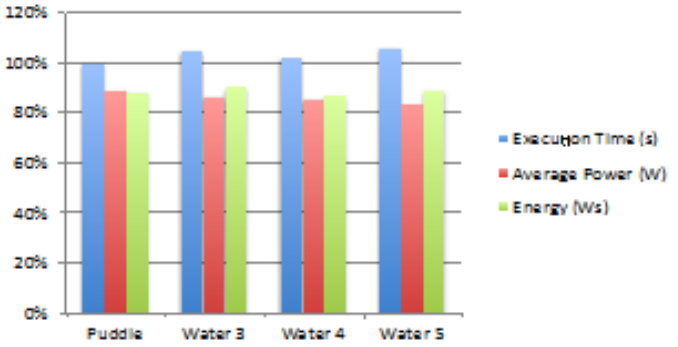
Darren J. Kerbyson

In this project, we bring together techniques that are aimed at tackling performance, power, and reliability on the latest supercomputers that contain a massive number of homogeneous processor cores.

Extreme scale systems are rapidly increasing in their capacity to tackle increasingly challenging scientific computational simulations. At the same time, their complexities are growing multifold in terms of their concurrency to provide performance that also impacts on their reliability and their electrical power requirements. Thus, extreme scale computing requires the combination of advanced techniques to achieve reliable, high performance at acceptable energy efficiencies. It is critical that these qualities are considered in combination for future supercomputers in the context of applications of interest to obtain high computational rates and hence achieve the most science on the available systems.

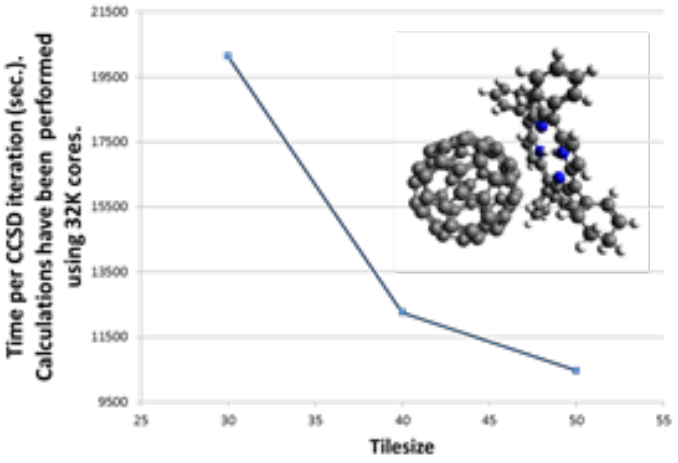
Within this project, we explore the use of several previously developed techniques in the context of large-scale systems, particularly those that comprise homogeneous processor-cores. Specifically, we are developing and integrating capabilities in performance diagnostics, energy efficiency optimizations, and software fault tolerant techniques with a particular emphasis on data movement. Last year, we developed performance diagnostics applicable to Global Arrays applications that allow detailed summary profiles to assist in bottleneck identification and integrated a novel Energy Template approach for a molecular dynamics application code to optimize energy use by up to 13%. We also explored optimized implementations on the latest advanced homogeneous systems, notably Blue Gene/Q, and developed data centric optimizations applied to molecular science applications.

The focus of FY 2013 work was incorporating computer science techniques into other applications. First, we incorporated dynamic energy optimization into the ARGOS molecular dynamics application that used application-specific behavioral information in conjunction with intelligent runtime software capable of utilizing power-saving mechanisms provided by underlying system hardware. Our Energy Template approach is based on a separation of concerns: for the application-specific policy identifying periods during execution, it is beneficial for each processor core to conserve power separated from the platform-dependent mechanism that defines the tools used to implement power savings. Making use of application-specific models of future application behavior combined with runtime data collection, we can identify points during ARGOS execution in which it is advantageous to make use of low-power states to conserve power without negatively impacting overall application performance. Experimental results show energy savings of up to 13%. A paper on this work is currently in preparation.



Relative execution time, power consumption, and energy consumption for ARGOS on four input desks.

In addition, we developed a new implementation of the coupled cluster model with singles and doubles (CCSD) method based on the integrated task pool for single and double equations. Currently, only one global task pool is employed instead two separate ones for equations corresponding to single and double projections. This leads to a significantly better load balancing. In addition, over 90% of the nxtask calls have been eliminated from the CCSD and equation-of-motion (EOM) CCSD codes. Scalability tests have been performed on Titan XK6 system at Oak Ridge National Laboratory showing over 70% parallel efficiency in runs



Impact of tilesize on the time per CCSD iteration for the C60-prophyrin complex (1036 basis set functions).

PN12006/2407

involving 12,000 and 24,000 cores (tests were performed for a large molecular system described by 998 basis set functions). We also demonstrated that the improved task scheduling provides better utilization of large tilesize, a basic parameter that describes the granularity of the Tensor Contraction Engine-generated CC codes. Using 32,000 cores, we observed over twice the reduction in time per CCSD iteration when moving from tilesize 30 to 50.

Using improved iterative implementations, we performed the largest up-to-date simulations for open-shell systems (triplet excited states of β-carotene: 216 correlated electrons and 1032 basis set functions). In the near future, we will perform even larger simulations for molecular systems composed of 1200–1300 orbitals. The results of our simulations are being submitted to the *Journal of Chemical Theory and Computation*.

Analytic Framework: Signature Discovery Workbench

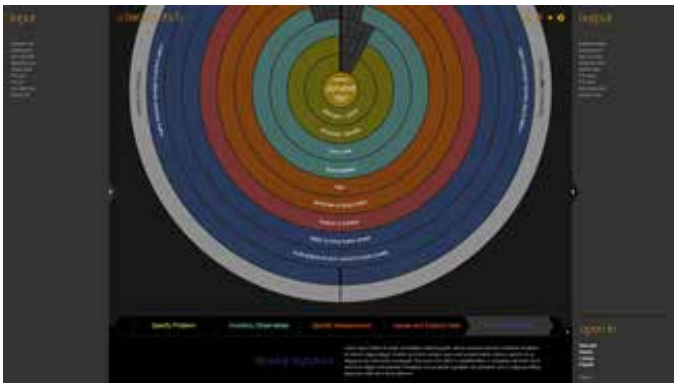
Russ Burtner

The Signature Discovery Workbench provides a graphical user interface that intuitively leads the user through the Signature Discovery Process using visual metaphors for status, context, and flow through a novel "query by example" recommendation system.

The Signature Discovery Workbench provides a facing application to allow users to create workflows quickly and efficiently that detect, develop, and validate information signatures from disparate data sources across multiple domains. The system will allow rapid comparisons between new and existing signatures as they are being developed, enabling informed decisions about workflow construction based on accuracy, utility, risk and cost. This rapid development and decision support will improve the signature discovery process and generated workflows through a radial decision tree, recommendation engine, and interactive visualizations.

Several design issues arise related to our research effort. Initially, we will examine the best way to visualize a signature discovery workflow in a way that does not hinder advanced users' ability to create new workflows but also aids novice users through the process. In this way, we hope to determine the best interaction methods needed to help users create a signature simply and efficiently, regardless of subject matter expertise and skill set. We will determine whether there is a single, understandable visualization method or technique that can be used to show accuracy, utility, risk, and cost for signature validations. In doing so, we will explore whether this method could also be used to make decisions when comparing recommended signatures.

During FY 2013, we made good progress toward our project goals. First, we interviewed 16 subject matter experts in the areas of biology, nuclear, cyber, and data science. Through the analysis of these interviews, we developed user-centered requirements for the prototype user interface and received validation and feedback regarding the signature discovery methodology and approach in general. In addition, we developed a set of design guidelines for the development of the analytic workbench. These are documented in a series of design specification documents that drives our development of the prototype within this project as well as documents the interface for how the workbench integrates into other projects.



The Signature Discovery Workbench enables users to create signatures visualized as multi-tier radial sequences of analytic steps. The interface allows for the creation, exploration, and execution of signatures.

Our team has made progress toward developing a proof-of-concept prototype to demonstrate the functionality and user experience of the workbench. At the end of FY 2013, the prototype is operational and has limited integration with the server to retrieve services and signatures. Specifically, we have a demonstration for a sequence alignment signature. As a result of our efforts, we published one journal paper discussing the concept of user interaction based on inferring signatures from user actions, wrote three whitepapers for clients, wrote one draft conference paper, supported two summer interns, and attended a seminar to develop a research agenda for the program.

In FY 2014, we will leverage spatializations of the domain-specific data to steer and augment the signatures. Based on the user research across the three domain areas (biology, nuclear, and cyber), spatialization visualizations are leveraged in each of these domains as visual metaphors to gain an overview of classification and clustering of information as well as a means for exploring relationships between the information. We plan to extend the workbench to include a spatialization of the user's information that includes inferring the analytical reasoning associated with specific interactions afforded in the visual metaphor, whereas workflows and signatures are implicitly created on behalf of the user. Thus, we will shift user focus from actively building a workflow to analyzing and synthesizing information spatially while the workflow is inferred and implicitly adapted. In addition, we will introduce "attractors" for information based on individual signatures. These will seed the clusters of information and thus drive the meaning behind areas of the spatialization. A radial signature developed in FY 2013 represents each of these clusters.

PN13020/2501

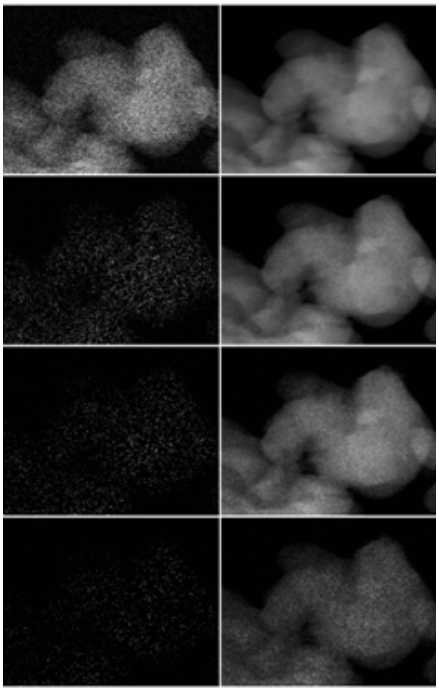
Compressive Sensing for Threat Detection

Andrew J. Stevens

This research improves the ability to detect and predict national security threats. Using compressive sensing algorithms, we worked to improve the efficiency and robustness of signatures that indicate threats in a noisy environment or during instances when these threats are purposefully obscured.

Compressive sensing is an information theory paradigm that provides methods for reconstructing observations or measurements from incomplete data streams. Typically, this gap is from a designed undersampling procedure, but it is possible to use compressive sensing on noisy or intentionally obfuscated data. While compressive sensing is a highly active research area and has provided some groundbreaking results, it has not been fully leveraged in the area of sensor applications and signature detection. Moreover, compressive discrimination – signatures that operate directly on the compressed measurements – is a relatively recent development with numerous applications.

Our research quantifies the effectiveness of reconstruction techniques in the context of signature detection. Specifically, in FY 2011, we designed and implemented a software infrastructure to run degradation/reconstruction/signature experiments that created and evaluated test cases to determine the optimal parameter configuration for the compressive sensing algorithm that would optimize signature detection performance. For our work in FY 2012, we augmented the software tool by supplying additional algorithm capabilities to improve signature detection performance and increased the speed at which a signature detection could be performed. We applied and experimented with our methods using a variety of observational data types, including multiple types of image and spectral data.



Reconstruction of STEM image data for a zeolite sample. The left side column is the sample, and the right column is the reconstruction. From top to bottom, the sampling levels are 100%, 20%, 10%, and 5%.

In the culmination of this project in FY 2013, our research developed capabilities in compressive discrimination and new sensor applications as follows:

Infrared spectroscopy. Using data from a sensor designed to detect airborne chemicals, we developed a method to discriminate among a set of known chemicals using a compressed signature extracted from the sensor data. Our approach reduced the number of false positive detections when compared to standard approaches such as partial least squares and matched filters.

Remote hyperspectral sensing. Using compressive sensing techniques, we were able to discriminate among multiple hyperspectral signatures in data compressively reconstructed from only 2% of the original data.

Scanning Transmission Electron Microscopy (STEM). We applied compressive sensing to STEM, and we were able to discern salient features in images reconstructed from 5% of the original data. This capability allows STEM to operate on materials that are too delicate for a complete scan. We are also working to apply this capability to other chemical imaging sensors as part of other projects and issues.

To apply compressive sensing in multiple domains, our originally created software tool was used to conduct compressive sensing generically. As part of the tool, we implemented procedures for automatically tuning various algorithmic parameters and validating the accuracy and robustness of signatures and reconstructions. The created tool has been integrated into the signature discovery toolsuite.

Although this concludes our project here, we have sought out internal applications for several other additional projects. Externally, we have applied for federal funding, and we have ongoing collaborations with Duke University and the University of Utah Scientific Computing and Imaging Institute.

PN11013/2335

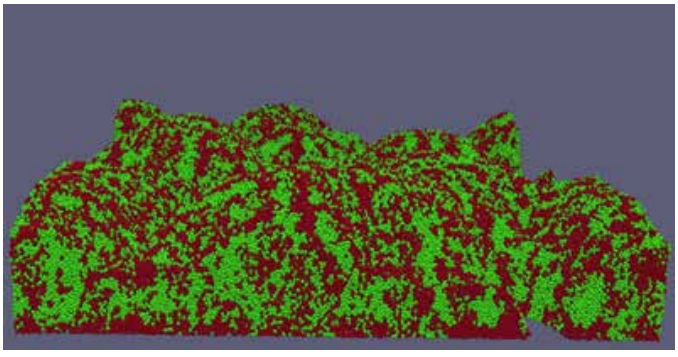
Computational Optimization and Predictive Simulation for Synthetic Biology

William Cannon

This project addresses challenges in predicting the response of biological systems to their environment with the development of a new simulation technology.

Despite the existence of computational optimization methods such as genetic algorithms based on biological evolution, it has been difficult to use these methods to study biological adaptation. Current methods for predicting biological responses using modeling and simulation make only a limited use of physical laws. We are developing a multi-scale, optimized, predictive simulation capability for biology that is similar in spirit to multi-physics codes. Based on physical principles and coupling metabolic reactions of microorganism with their environment, this simulation capability allows for a dynamic optimization of metabolic processes to the available nutrients and environment. To model biological systems on a large scale using physics principles, high performance computing methods have been used. While these methods can be used to simulate any biological system, our focus has been on applications for synthetic biology and natural settings.

The result of this project has been the development of two independent but synergistic computing programs: Boltzmann for simulating individual cells and Biocellion for large-scale simulations of millions to trillions of cells. Boltzmann is designed for more physically realistic single-cell simulations. Unlike current flux-based approaches that ignore thermodynamics and the law of mass action, our approach models the states of cells using thermodynamic equations of state, or a mathematical implementation of the law of mass action. These equations describe concentrations of all the chemical species and energy levels that the system can obtain. While stochastic simulation methods have been in place for modeling kinetics of metabolic processes in biology, these approaches are very difficult to scale because all of the parameters needed for a kinetic simulation (rate constants) are difficult to obtain, where the parameters needed for a state-based simulation (the standard free energies of reaction) are relatively easy to obtain. In fact, we built into the code a method for calculating these parameters at different pHs, ionic strengths, and temperatures.



A snapshot of a yeast colony growth simulation using Biocellion. Cell color is based on their metabolic activity state. The physic-chemical interaction between the cells and their environment, an agar plate, was also modeled.

Overall, the code was designed with ease of scaling in mind. Boltzmann was presented at a conference and will be submitted to a journal for publication. Boltzmann consists of 20K lines of ANSI C code that can be run on any computer either as a stand-alone executable or as a component of a larger program. Other programs can call Boltzmann using a programming interface that allows the state of the system to be preserved and communicated to a master program that provides a high-performance computing framework for simultaneously running millions to billions of individual Boltzmann simulations.

To satisfy the need for a high-performance computing framework that allows the simulation of millions or more microbes, we developed Biocellion. This computer program is a scalable simulation framework for multicellular biological systems that significantly improves the software development cost-benefit ratio when using powerful parallel computers. Biocellion consists of several hundreds of thousands of lines of computer code designed for large distributed clusters with multicore compute nodes. The program encodes three functional capabilities: update cell states, evaluate the mechanical interactions between cells, and track molecule concentrations in the extracellular space to model the indirect interaction via diffusible molecules. The interfaces for these capabilities have been abstracted so that Biocellion can be coupled with any agent-based modeling program such as Boltzmann or other programs such as Copasi.

To make Biocellion maximize extreme scale computing resources, algorithms are implemented using state-of-the-art parallel computing libraries and models: Partitioned Global Address Space (PGAS) Global Arrays, MPI, and Intel Thread Building Block (TBB). While Global Arrays allow users to create arrays spanning multiple compute nodes

PN12016/2417

without in an automated fashion, TBB takes advantage of users expert knowledge of irregular parallelism and allows them to create parallelization points. Additionally, TBB maps the nested irregular parallelism to the multiple cores available on a shared memory compute node.

The code involves two different types of inter-node communication in exploiting distributed memory parallelism: irregular, small, and infrequent data accesses to update load information for load balancing and more regular, large, and frequent communication for ghost data exchanges. We use Global Arrays for the irregular but infrequent communication to lower programming complexity, and we use MPI for ghost data exchanges to achieve high performance. An initial paper describing Biocellion has been written and will be submitted for publication when the user manual is completed. A conference presentation on Biocellion is also upcoming.

Already, Biocellion and Boltzmann are being used on multiple projects at PNNL. In addition, the Institute for Systems Biology and at the Fred Hutchinson Cancer Research Center are using Biocellion to address problems of how cells grow and self-organize. Future work will allow us to couple Boltzmann and Biocellion, which will provide an unprecedented predictive modeling capability for biology.

Cyber Security Testbed and Dataset Generation

Sam Clements

As tools and techniques change, enhancement of a cyber security testbed will allow research to keep current with the ever-changing field of cyber security. We will be able to model changes, create quality datasets, and devise new ways to thwart cyber attacks.

A prevailing issue with dataset collection is the method and process by which it is generated and collected. The experiment scenario setup requires repeatability and reproducibility with different levels of fidelity dependent upon the research being performed. Most existing cyber security datasets are collected from one-time incidents. These datasets are either burdened with limited distribution requirements or have been modified to protect the data provider or users information within the dataset. Often, both requirements occur, which limits the availability and usability of the dataset. There have been few efforts to generate expansive and realistic datasets for the research community, and it is often difficult to collect data from real cyber security events because of the potential liability and reputation impact.

This project is developing a testbed on which scientifically sound experiments can be run, whereas most cyber security experiments done today are conducted ad-hoc using networks and datasets that are incomplete, anonymized, unstructured or otherwise lack rigor resulting in results that are difficult to replicate, validate, or compare with other experiments. We will use a number of technologies and other tools to generate datasets of realistic computing environments. First, we seek to build out a representative enterprise network where variables can be controlled to allow for rigorous scientific experimentation. Second, these experiments will be captured into annotated datasets that can be freely distributed to researchers. Our primary goal is to understand the ways in which enterprise testbeds can be built and instrumented such that they provide a valuable platform on which to conduct other experimentation and activities. This research is an important core of and the basis for furthering other research and development in cyber security at the enterprise scale.

We started developing the CyberNET testbed using Citrix Xenserver hypervisors. With this platform we built out the Kritikos network and learned the inner workings, challenges and limitations of Xenserver. The built-in management tools for Xenserver do not include the capabilities to easily and dynamically create the networks that are a part of our project plan. As a result, we began investigating other tools and settled on Microsoft System Center Virtual Machine Manager (SCVMM) as it has a plug-in for Xenserver and HyperV along with many options for creating and deploying various network and system configurations. Using the HyperV cluster, we were able to create and deploy various small network configurations from SCVMM ranging from 5 to 60 machines. These networks consist of domain controllers, file servers, and various desktop agents.

The question of what the networks should look like and how to design them has been a question throughout the year. We held various discussions and collaborations along these lines with staff from the University of Washington, the U.S. Cyber Command, and the Software Engineering Institute/Computer Emergency Response Team. One of the primary challenges that always emerged was benign white noise generation. It is quite challenging to create realistic datasets with adequate benign data.

The most important findings we discovered this year are the lack of benign network data generators and the necessity of having the right hardware and software for the intended application. Developing CyberNET has exposed numerous hardware and software limitations to simulating a midsized enterprise. For example, HyperV does not support many of the operating systems used for collecting data (e.g., several Linux distros, BSD, and Open BSD). Additionally, the initial hard drives were too slow to handle the number of VMs that were required to meet our target goals. We also identified network biases from our virtualized network environment.

Understanding these challenges and applying appropriate mitigations are keys to long-term success. All of these challenges helped to prepare for success as we begin the efforts to collect useful datasets for other researchers. We have found solutions to many of the challenges and are set to begin integrating the solutions in the next year. The primary follow-on activities will expand CyberNET capabilities by removing the hard drive bottleneck, optimizing data collection techniques, and developing a network noise generator. Together, these enhancements will collectively push the envelope of the possible in-network simulation, dataset generation, and utility to other researchers and provide a platform that improves other research activities.

PN13051/2532

Data-Intensive Algorithms for Bioinformatics-Inspired Signal Detection

Christopher S. Oehmen

This project is developing a general purpose method for discovering sequence-based signatures for use in solutions to cyber and network security challenges.

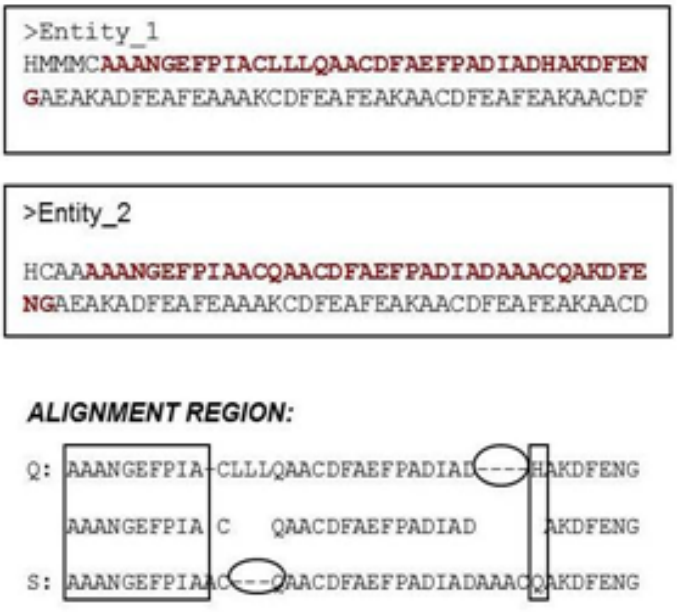
Many cyber security challenges result from a combination of scale and complexity. For example, understanding the intent of network transactions to defend the DOE complex requires analysis of an enormous, ever-growing body of digital information to identify increasingly sophisticated attempts at infiltration and exploitation. To make matters worse, in this environment the vast majority of transactions are benign. One limitation of rule-based detection schemes is that we must first know what to look for, which is a challenge in the face of constantly evolving exploit methodologies.

This project is directed at developing capabilities that enable the discovery and analysis of signatures in cyber and network security. This method is most impactful when a sequence of events or behaviors is necessary to discriminate between true and false identifications. The root of the method is derived from bioinformatics-based techniques for discovering sequences of text that are more similar than one would expect by chance. We extend this concept, originally applied to gene and protein sequences, to arbitrary sequences of events. This allows us to discover patterns of interest in string-based information (such as network transactions or legacy source code) by using biological theory to capture the inexact and evolving relationships between text strings associated with digital information and use machine-learning principles to extract patterns without a priori knowledge of those patterns.

Success would have significant operational impact by enabling pattern-based, data-driven identification schemes grounded in a theoretical framework to augment the current rule-based approaches to cyber security. The intent was to extend existing algorithmic and computational infrastructure so that it can be applied to a wider space of potential applications in cyber security and other domains. In FY 2012, we explored more complex representations of the signatures of interest, including graphs and profile-based models. In FY 2013, we investigated the ability of our algorithms to detect and/or quantify signature changes over time or evolution. We also supplied our set of tools to be the first demonstration of the Analytical Framework and the Semantic Workflow process.

Signature Drift Experiment. We investigated the applicability of identifying signatures of evolving systems and attempted to understand the nature of the evolution of the signature as it related to the evolution of the data set. We created an experiment within which we took a number of variations (versions) of a common code library and processed them through our pipeline. We generated signatures for the first 8 versions (of 16 total) as a basis for comparing the later versions. The functional change was minor in this particular dataset and evolution was difficult to see. Other more diverse datasets should prove easier to analyze. We also generated our own evolution of a simple code base using standard sort algorithms by creating several versions based on real-world types of development. The intent was to use this as a control set to match with the library code discussed above.

The two datasets proved to be dissimilar enough to make comparison unwieldy in the timeframe of this project. In the end, we determined that our code libraries were too similar from beginning to end to have much drift in their signature and our control code was too small to prove the method at hand. However, we believe that further analysis of this data will provide insights to the ability to detect the evolution of our signatures.



Bottom: Alignment between the two highly conserved regions (bold) in the two entities above. Boxes show blocks of exact matches (left), a mismatch (right), and circles show insertions necessary to align the two sequences.

Integration of the Analytic Framework. As the Analytic Framework started to take shape, they called upon our project to the be the prototype for integrating tools into the framework, creating the user interface for the framework and development of the semantic backbone of the framework. We worked closely with all three teams to create an end-to-end demonstration of our method using the semantics and mechanics of the Analytical Framework, which was demonstrated successfully.

PN11019/2341

Developing Functionality and Performance Enhancements to the Global Array Toolkit

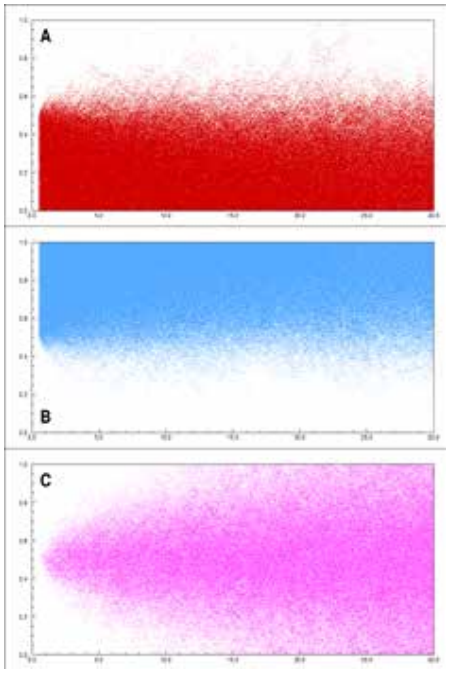
Bruce J. Palmer

We are determining what performance and functionality bottlenecks need to be addressed in global arrays (GA) to support applications that are expected to run on exascale platforms. Our project will help implement proposed solutions in target applications and evaluate their effectiveness.

GA is a partitioned global address space programming model developed in conjunction with the NWChem chemistry suite at PNNL. Currently, GA is the premier data-centric programming model that has been used on most of the large-scale supercomputers deployed to date. Its programming approach provides simple one-sided access to data arrays that are globally visible from any process in the system. Despite its success in applications that use it, the GA application base remains relatively small. Existing applications include PNNL's NWChem Chemistry code, the STOMP subsurface code, bioinformatics (Scalablast), and conventional computational fluid dynamics (TETHYS). More recently, it was used as the communication layer in the GridPACK™ framework for modeling the electric power grid.

This limited GA base is a concern because further development of the application would benefit from a more broad range of requirements. Thus, this project incorporates GA into several applications that represent a large diverse group of computing problems with the goal of using these implementations to understand the scaling behavior of GA within multiple scenarios. This project has also completed work on an extension of GA, Global Pointer arrays (GP), which expands the GA concept to arrays of arbitrary array elements.

Target areas in this project include sparse matrix operations, discrete particle tracking combined with unstructured continuum fluid flow solvers, large-scale semi-empirical electronic calculations, and lattice Boltzmann simulations of continuum flow using structured grids. These problems will help us to define new functionality needed by application developers to incorporate GA into their codes. In FY 2013, progress was made in all four application areas, as detailed below.



Simulation of simple A+B→C reaction occurring on discrete particles using TETHYS.

Sparse matrix operations. We previously developed code that implements a sparse matrix-vector multiplication using the GA library and compared it with a comparable operation using the MPI-based Hypre solver library. We modified the previous implementation so that the GA-based algorithm approximates the data layout of the Hypre implementation and reduces communication. Testing on Infiniband systems indicates that a fully optimized Hypre implementation is faster than the GA implementation but still within a factor of two or three in performance. The communication in GA, however, is much simpler to implement. The GP arrays were used to implement a sparse matrix transpose operation, and results indicate that core parts of the algorithm scale well. The portions of the algorithm with poor scaling characteristics correspond to operations that potentially only need to be performed once in an iterative context and therefore might not be a serious bottleneck in actual applications.

Semi-empirical electronic structure code. The coulomb interaction contribution to the Kohn-Sham matrix in the CP2K electronic structure code has been converted from its original replicated data implementation using MPI to a fully distributed algorithm using GA. The original algorithm was modified so that the calculation was decomposed into pair interactions between spatial blocks to reduce latency costs in communication. Additionally, the algorithm was verified to give numerical correctness and competitive performance with the original MPI code. Currently, we are working on scaling the algorithm to large core counts, for which we recently demonstrated that the algorithm works.

Reactive chemistry and particle tracking. The particle tracking algorithm has been parallelized and incorporated into the TETHYS continuum fluid dynamics code. More recently, unimolecular and bimolecular reactions have been incorporated into the code to model combinations of particle convection and chemistry. Efforts are currently underway to combine this capability with the bacterial communities micromodel.

Lattice boltzmann algorithm. The PARAFLOW lattice Boltzmann simulation code was converted to use GA for the ghost cell updates that are a core part of the algorithm. For FY 2013, we continued to investigate the use of non-blocking ghost cell updates operations combined with a two-phase update of the LB cells to see if we could improve scalability by overlapping computation and communication in the lattice Boltzmann algorithm. Unfortunately, no significant performance benefits were obtained.

PN11021/2343

Development of Parallel Multi-Reference Coupled Cluster Capabilities

Hubertus J.J. van Dam

This project will bring out extreme levels of scalability available in principle in multi-reference coupled cluster methods. The resulting implementations will address important problems in areas such as transition metal catalysis and chemically reactive potential energy surfaces for which multi-reference methods are essential.

Single-reference coupled cluster methods have proven to be scientifically useful for describing correlation effects in electronic structure theory. Previous projects have demonstrated that these methods can be implemented in a highly scalable way. However, single-reference methods can degrade if the chemical system exhibits near-degenerate electronic states. In those cases, there is essentially no single electronic configuration that stands out as a suitable candidate for the reference configuration. In such a situation, the only solution is to select a set of dominant configurations and use all of them as reference configurations. This is the basis of all multi-reference coupled cluster methods.

In practice, many multi-reference approaches are developed following different philosophies addressing the multi-reference problem and emphasizing various characteristics. These approaches have shown the high value of multi-reference methods for challenging problems. A limitation of current implementations of multi-reference coupled cluster methods is that they have been focused on theory proof-of-concept aspects instead of scalability. Interestingly, multi-reference methods are very promising candidates for parallelization, following from the structure of the equations that allow for two levels of parallelism. One level of parallelism comes from equations associated with a single reference configuration that we have shown can be made to scale well. The second level derives from the fact that all reference configurations can be treated concurrently.

The objective of this project is to express the parallelism inherent in the multi-reference methods for a variety of methods. Our research involves a thorough analysis of the coupling terms between equations for the different reference configurations, as it will dictate how the parallelism can be expressed. Also, the importance of couplings of different types will be considered in relation to their impact on the achievable parallelism. Ultimately, the project will deliver extremely scalable and scientifically versatile capabilities for studying a comprehensive set of chemical systems.

In FY 2011, we focused on the Brillouin-Wigner multi-reference coupled cluster method, which uses particularly simple coupling terms but has the basic infrastructure for a multi-reference coupled cluster approach. We implemented improvements to handling integral files and the use of task pools, and we demonstrated the basic parallelism this method can achieve. We found that even at this level, code scales to at least an order of magnitude more processors than previously reported in literature. We also found that triples corrections are important for achieving highly accurate results, even in the multi-reference case. An analysis of perturbative correction terms was performed focusing on both accuracy and achievable scalability.

In FY 2012, we addressed three main issues. First, the code developed in FY 2011 was modified to exploit the reference level parallelism. For this purpose, the full set of processors is partitioned using processor groups and the reference configurations are distributed across the processor groups. We found that both the Brillouin-Wigner and Mukherjee formulations of multi-reference coupled cluster scale well using such an approach. Second, we added perturbative triples corrections to the single-double code. In this approach, we built on the infrastructure created for the reference level parallelism. As expected, the triples correction improved markedly, and we demonstrated scalability to 80,000 cores on Titan. Third, we considered a number of size-extensivity corrections, important for the Brillouin-Wigner formulation which by itself is not strictly size-extensive. We developed a universal correction for application to multi-reference formulations and showed that this correction significantly improves Brillouin-Wigner formulation results. Finally, we implemented an approach that allows core-level excitations to be treated with multi-reference coupled cluster that offers a new application domain important for the characterization of materials. We presented our work at two conferences, published three papers, and have submitted another paper to a journal.

In FY 2013, we addressed two remaining bottlenecks. The integral transformation preceding the coupled cluster component was made to scale much better. In addition, we extended the coupled cluster code to exploit accelerators that provide the major part of the compute power of current architectures. Finally, we demonstrated the capabilities of the code on a variety of science problems, including the fate of uranium in subsurface processes, the chemistry of aerosols, and biological light harvesting systems. This work was published in three papers and two conference presentations and was honored with an award for outstanding postdoctoral achievement.

PN11030/2352

Discrete Mathematical Foundations for Cyber Systems Analysis

Cliff A. Joslyn

Our goal is to develop a modeling formalism for representing state and change of state in general cyber systems with hierarchically interacting discrete mathematical models.

Cyber systems are currently modeled piecemeal and ad hoc, and unifying methodologies are elusive. Traditional modeling methods are inappropriate because in digital systems, there is the absence of space, ordinal time, energy, conservation laws, and natural metrics involving continuity or contiguity. Thus, the Asymmetric Resilient Cybersecurity (ARC) group needs to advance a comprehensive formal framework for cyber analytics built on sound modeling methods founded in discrete mathematics. The repertoire of state variables validated as significant will form the foundation for robust multi-scale models to serve analysts and decision-makers.

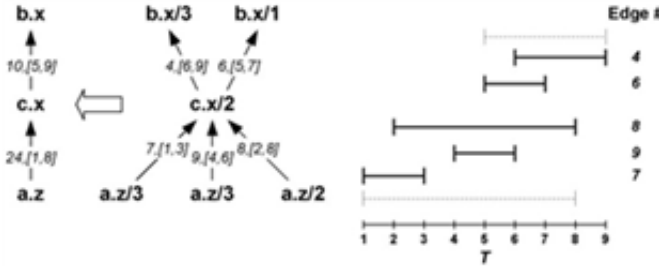
We seek to represent cyber systems as discrete mathematical objects interacting across hierarchically scalar levels, each distinctly validated but interacting, similar to hybrid modeling and qualitative physics but discrete all the way down. Expressible as complex attributed graphs and networks, our approach draws on broad techniques from combinatorics and discrete mathematics to develop metrics for resiliency in scalar levels of cyber systems. All discrete math objects are ultimately representable as systems of binary relations and thus (because the most general binary relation is an arbitrary directed graph) can all be reduced to graphs in some form. Though they may appear to be merely graphs, they are decidedly complex, even arbitrary objects involving additional structured attributes, features, and constraints. While discrete representations for cyber will be graph-based or -reducible, they will have special mathematical properties appropriate for the constraints they are representing. In turn, these images imply novel metrics based on those mathematics. We will build from the foundations of discrete math that are directed graphs but embellished with further combinatorial complexity in the form of complex labels, hypercubic data structures, partial orders, and partitions.

Our initial focus is on Netflow data structured as a graph with complex IP:Port labels, time interval edge attributes, and payload information. This inherently multi-scalar position allows drill-down to packets or roll-up to whole-network representations. We prove our approach by performing Netflow analytics with the three new methods described below.

Graph cubes. The complex node attributes support the ability to do hierarchical roll-up over different aspects of node structure, port numbers and IP octets. The resulting graph contractions provide reduced network views and the ability to find patterns at multi-scalar coarsened levels.

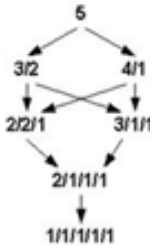
Example of graph cube

Time interval analysis. Netflow records are attributed with [begin,end] time intervals that can be ordered in three different ways by measuring the presence, kind, and amount of overlap. Series connections demand certain order relations, while the edge contractions supplied by the graph cubes require and respect different orderings.



Example of time interval analysis

Distributional statistics. Each IP:Port node stands maps to a set of downstream IP:Ports in a Netflow graph, and the grouping of that distribution determines an integer partition. Statistical information measures on those partitions allow us to gauge the amount of grouping among the IP:Port node pairs, identifying outliers, and unusual patterns. Previously established information theoretical combinatorial measures of the compression and steepness of labeled degree distributions are complemented by appropriate correlation and rank correlation measures.



Example of distributional statistics

Modeling at this level of ambition and generality requires firm grounding in specific applications and data sets to avoid vacuity or uselessness. With the rest of ARC, we have identified and are developing a range of test databases and canonical use cases, initially at the Netflow scalar level. Practitioners and engineers on our team interact with other domain experts to identify and develop standardized test suites.

PN13038/2519

Enabling the Meaningful Exploitation of Integrated Regional Earth Systems Data

Kerstin Kleese van Dam

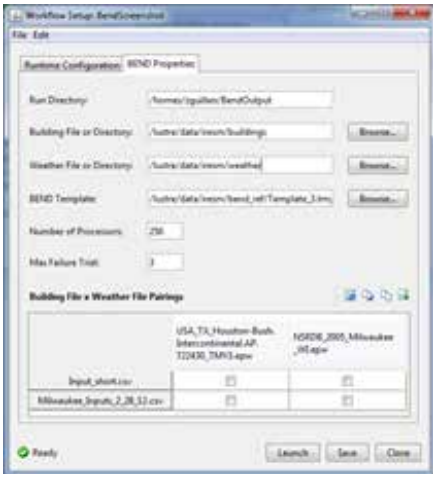
This project developed a flexible coupling software framework for multi-scale, multi-disciplinary integrated regional earth systems simulation models as well as designing and implementing a supporting research infrastructure to enable the collaborative development, exercise (including uncertainty characterization), and result analysis of the modeling framework.

The fundamental goal of the Platform for Regional Integrated Modeling and Analysis (PRIMA) is to develop a framework that enables the critical analyses of the tradeoffs and consequences of decision and policy making on the background of integrated human and environmental systems, combining a wide range of different scientific and economic processes. While initially focused on modeling interactions of climate, socio-economic, crop, and energy systems, further inclusion of other science domains such as biogeochemistry, subsurface flow, and medicine are envisaged for the future. Coupling models from different disciplines and backgrounds presents a range of key scientific, computational, and data challenges, such as

- consolidating and integrating overlapping processes consistently across models
- defining scientific relevant exchanges and feedback between models
- dealing with significant variability in spatial and temporal scales
- differing simulation models, implementation approaches, programming languages, and computational requirements
- differing data formats and variable names
- differing data structure, representation, and semantic issues between the unique scientific domains involved.

Thus, barriers are significant scientific, technical, syntactic, semantic, and communication challenges to enable the successful integrated model framework development, including model coupling, evaluation, and numerical experiments to address the relevant scientific questions. For the collaborative creation of new scientific insights from this innovative modeling capability, it was further necessary to facilitate information exchange, evaluation, and use of complex modeling results through the rigorous data model that integrates the data, information, and knowledge provided by each into a form in which different aspects can be flexibly represented to the appropriate user communities. It also became clear that the advanced modeling framework enabled the scientists to run their models in ensembles of unprecedented scales, thus producing results orders of magnitude larger than ever before. This drastic increase in data volume overwhelmed existing data analysis provisions and the project responded by developing additional, highly scalable analytical methods and visual analysis capabilities.

In FY 2011, we designed a decentralized knowledge model that could leverage existing community specific standards and vocabulary for data and metadata (such as NetCDF CF), which were then integrated via ontologies that created "crosswalks" between the disciplines where overlaps existed. This model was the foundation for the data exchange between models and subsequent integrated analysis of multi-model results. For FY 2012, we built the flexible model coupling framework and supporting collaborative infrastructure to enable planned modeling experiments, which extended the MeDICi workflow system and applied the developed components to create an initial set of complex scientific workflows. These components and workflows were successfully used in FY 2013 to create and harden required workflows and support model experiment execution. We used existing PNNL-developed Velo software as its core data and knowledge management, analysis, and sharing infrastructure. To serve community needs, the following new components were developed.



PRIMA BEND model set-up interface

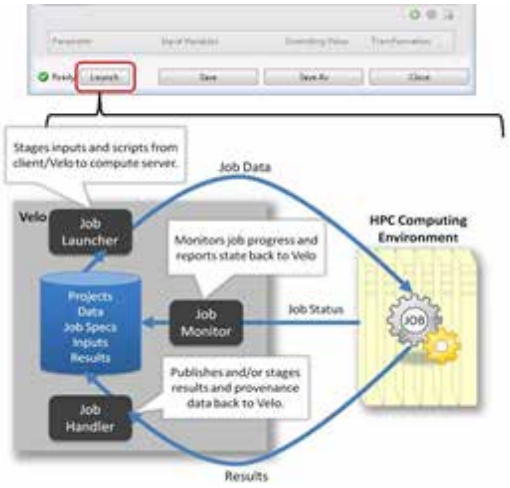
PRIMA data model. This content model defines PRIMA-specific objects and their relationships as they will be stored in the Velo knowledge base, which has three primary sub-models: the project model, registry, and provenance model.

PN11032/2354

Project model (numerical experiments). Projects are data objects that help scientists to organize their work and encapsulate related data. PRIMA created a specific project subtype for numerical experiments that encapsulates all required components such as workflows, referential data, and any results from simulation runs.

Registry. The registry defines key data types that PRIMA added to greatly enhance the provenance graph provided by Velo: datasets, tools, and machines. Each of these types can be versioned and are included in the provenance graph.

Provenance model. Provenance provides a historical record of what transpired in a numerical experiment and is implemented in PRIMA by a combination of metadata saved to objects and a graph of relationships connecting them. When a workflow instance is created inside a numerical experiment, Velo relationships are added, connecting the workflow instance to the workflow definition. As workflows run, status information is conveyed in near real-time and stored as part of the workflow instance metadata. Relationships are added connecting all the inputs and outputs to the workflow instance. A relationship is also added to connect the workflow to the machine on which it ran. Because the provenance model is a part of the overall PRIMA data model, any referenced datasets, tools, and machines can be associated with the provenance providing a powerful cross reference search capability, referred to by the provenance community as knowledge provenance.



Velo Job Handler and PRIMA BEND Job Launcher

PRIMA job execution components. PRIMA streamlined the job execution process for all of the different models and couplings (and upcoming uncertainty quantifications) by invoking all simulations as different MeDICi workflows; therefore, a MeDICi Job Handler was created that was associated with all workflow instances. The MeDICi Job Handler invokes, monitors, and post-processes all running workflows in a consistent manner. Workflows are initiated by remotely dropping a MeDICi workflow bundle file to a monitored folder that serves as a pipeline endpoint. The MeDICi Job Handler parses the MeDICi status log as it is running to update the workflow instance metadata with the current run status and creates separate folders under the workflow instance's Output folder for each run. At the completion of each MeDICi pipeline, the Handler automatically copies the MeDICi status file both to Velo as well as small output files and images. The entire output file set is automatically registered and archived via the Publisher component described below, and a remote link is created inside the workflow outputs folder pointing to the remote data. Any additional post-processing steps used for analysis (such as combining the results of multiple files into a single table) are handled by the pipeline itself as a final step.

PRIMA provenance components. In addition to adding the Registry object definitions and automatically capturing provenance these new provenance relationships during workflow instance creation and execution, PRIMA also developed an Archive component that resides on the server. The Archive is a protected area where data (such as simulation outputs) are copied so that they cannot be inadvertently deleted. All remote links created in the knowledge management system point to an Archive location to avoid dangling pointers. The Archive also promotes reproducibility of results in that users can guarantee that they can find the results of a given simulation. It additionally improves performance, as it prevents large data from being transferred to the knowledge management server.

PRIMA contains a corresponding data Publisher to publish data to the Archive. The Publisher is implemented as an MeDICi workflow so that extremely large datasets (such as PRIMA) can be copied in parallel to the Archive, accelerating the process. Every new remote dataset registered via the PRIMA Registry View invokes Publisher to transfer data to the Archive. The Publisher is remotely invoked as other MeDICi workflows by creating an SSH tunnel and placing a MeDICi configuration file in the file system endpoint folder.

The Velo infrastructure and the MeDICi system were used to develop three basic workflows for specific PRIMA models: BEND, RGCAM, and EPIC. For BEND, the project created a more complex high-throughput modeling and analysis environment to demonstrate the capabilities of the infrastructure. The BEND model quantifies energy demands and consumption of residential and commercial buildings based on the characteristics of thousands of different buildings across a geographic area. Representative building type portfolios are created for small geographical regions, within which an energy demand profile is calculated for each building type combining building usage and characteristics with the weather for that geographic area. As a result, scientists have to run the model multiple times to create the energy profile for a larger region.

Using the PRIMA infrastructure and workflow systems, we created an environment where scientists could easily set

PN11032/2354

up, execute, and analyze complex scenarios for BEND. In addition to the simplified set up and execution of the modeling runs, the infrastructure has enabled significant efficiency gains through consistent verification of model results and error recovery procedures for problematic runs. With the new workflow set up, BEND was run 2.5 million times in a few weeks to model the building energy demand in the eastern interconnect (EIC) portion of the United States. A "simple" BEND model experiment generated over 600,000 annual profiles to represent commercial and residential hourly energy demand for 3 years. These profiles covered 100 unique climate regions, 16 unique building categories (e.g., offices, single family residential), and multiple vintages of buildings as well as different HVAC systems. The resulting dataset was 400 Gb in size. The larger PRIMA project will explore the EIC at a higher degree of spatial resolution, and the analysis will be done on a 5-year time step covering a large portion of the modeled century. This will result in the potential of the BEND model to generating data sets in the tens of terabytes.

For additional FY 2013 work, we developed analytical capabilities of the infrastructure and framework to match increased data volumes generated by simulations such as those performed by the BEND team. This work solidified the Medici and Velo architecture and improved the BEND model integration with PIC. With improved integration and process handling, additional tools were needed for data discovery and analytics; thus, we developed scalable data handling tools for data manipulation and analysis. Our deep data analysis utilizes a new R implementation specifically developed for the large-scale data analysis platform Hadoop, a framework that allows for the distributed processing of large data sets across clusters of computers using a simple divide and recombine programming model "Map and Reduce." It is designed to scale up from single servers to thousands of nodes, each offering local computation and storage.

The latest development growing in the big data domain is R on Hadoop, which provides unlimited scalability. R is a well-known open source library of statistics algorithms with a large following in the many different data analysis communities. It allows the easy generation of basic or complex statistical analysis scripts based on a common language. R has a well-maintained and tested core of functions and a rich set of contributed libraries from the statistics community (several thousand libraries) that are open source and well documented.

RHIPE is an R package that integrates the R environment for statistics and data analysis with the Hadoop distributed computing framework. Using the RHIPE environment, we designed and implemented a range of data handling and exploration functionalities. For BEND, the benefits of RHIPE are largely associated with the ease to scale the data dramatically and the speed at which results can be compiled. This functionality allows a traditional R analyst the ability to do detailed exploratory data analyses with little wait time (seconds/minutes as compared to hours/days).

While the more intense data handling and statistical insight is anticipated to continue being developed within the R framework in FY 2014, PNNL teamed with the Scientific Computing and Imaging Institute at the University of Utah to develop a data interactive software. The visualization software is intended for BEND building scientists to perform exploratory data analysis on the precompiled hadoop "mappings" of the data in a simple-to-use graphical user interface. The software allows the user to compile a linked dataset into varied summary tables and graphic visualization.

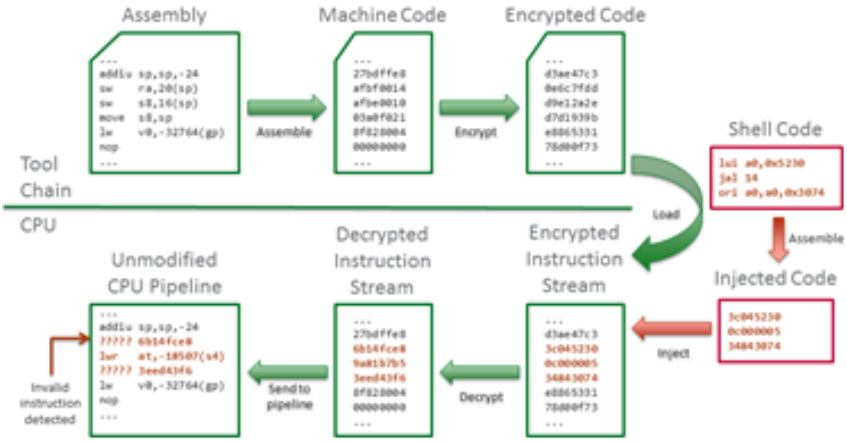
Encrypted CPU Instruction Stream

Richard L. Griswold

This project seeks to render several types of cyber attacks ineffective by using instruction stream and program counter encryption to create unique, dynamic system architectures, thus denying our adversaries a key asymmetric advantage.

While instruction stream encryption is not new a concept with references dating over 20 years, it has not yet been implemented at the hardware layer. Previous studies in this area have demonstrated the effectiveness of the technique, but research did not progress beyond implementation in software emulators or operating system components. These systems suffered from significantly reduced performance, and the software components themselves were vulnerable to attacks. We initiated this project to research methods for implementing this technique at the hardware layer to increase performance and reduce the attack surface. During the course of this project, we investigated program counter encryption as a complimentary technique to provide extra protection against additional attack vectors in the form of data-injection attacks.

This project has two primary objectives: first, to demonstrate that instruction stream and program counter encryption can effectively defeat several attack vectors, and second, to implement these techniques at the hardware layer. We focused our efforts on embedded systems instead of general-purpose workstations, which gives us a larger variety of processor architectures to work with and much greater hardware configurability. We implemented both instruction stream encryption and program counter encryption in hardware using an FPGA for our prototype. This has allowed us to measure the performance, resiliency, and effectiveness of these techniques.



Code injection attack attempt with instruction stream encryption enabled

During the first year of the project, we selected Plasma, an MIPS-compatible soft-core processor, as the target for our prototype and successfully implemented instruction stream and program counter encryption in a software emulator. We modified the GNU tool chain (compiler, assembler, linker, and associated utilities) to support creating binaries with encrypted instruction streams and creating additional instructions needed for program counter encryption. The software emulator and modified tool chain validated our designs through test programs and an operating system, including a network stack and web server, with the entire instruction stream encrypted and program counter encryption enabled. The tools also supported test caching techniques for improving performance and measuring effectiveness against both code and data injection attack vectors.

Our tests demonstrated that instruction stream and program counter encryption are able to successfully defeat code-injection attacks that seek to introduce new, malicious code into the system and data-injection attacks that can compromise system security by hijacking the execution flow. The tests also revealed that using a decrypted instruction cache can keep the impact to system performance very low without compromising security.

During FY 2013, we implemented and demonstrated instruction stream encryption and program counter encryption in hardware using an FPGA. For the hardware prototype, we implemented the modifications to the Plasma processor in a Xilinx Spartan-3E FPGA clocked at 25 MHz. We used AES in counter mode (AES-CTR) for the encryption algorithm and based the decryption engine around an optimized Avalon AES ECB encryption core with a 128-bit key length. Tests with this configuration show that it effectively blocks code-injection and data-injection attacks as well as attempts to execute unauthorized binary code. The performance penalty for our hardware prototype is less than a 5%, which is at least an order of magnitude better than the previous software-based approaches.

We filed two invention reports in FY 2013, one for instruction stream encryption and one for program counter encryption. Additionally, we filed a patent application for instruction stream encryption and submitted a conference paper discussing our instruction stream encryption implementation. We have now completed the original goals for the Encrypted CPU Instruction Stream projects and are currently pursuing commercialization opportunities for these technologies.

PN12026/2467

Establishing a Strategic Goods Testbed to Disrupt Illicit Nuclear Trafficking

Luke E. Erikson

This project focuses on establishing a strategic goods testbed to disrupt illicit nuclear trafficking.

Accessing data and understanding the state-of-the-art has proven to be a frequent challenge for researchers in the area of the disruption of illicit nuclear trafficking. The barriers for entry include security concerns, policy and logistical hurdles, and cost. One of the most promising approaches to combat these barriers, however, is to leverage current and pending investments in analytics into this complex, data rich domain. It is believed that applying advanced analytical approaches would improve the robustness of the day-to-day tactical analysis frequently conducted at PNNL and other national laboratories. Additionally, there is good reason to expect that analytical tools could empower analysts to operate more strategically either by assisting in new analyses or streamlining routine tasks. This project is intended to lower the bar and give easy and appropriate access to data, schema, examples, exercises, and domain summaries.

To meet the above needs and challenges, a strategic goods test bed has been created that focuses on analyzing illicit nuclear trafficking. This system is intended to include a framework that is representative of the real world operational resources used today at PNNL and other locations for non-proliferation tasks. An accurate framework is critical as an array of policy, security, and logistical issues present in this domain are a challenge to many traditional analytical approaches. In addition to the framework, a cache of realistic data (either sampled from actual data sources or accurate proxies) would be needed to provide accurate analysis challenges. In addition, the success or failure of the prospective work associated with this and related projects depends explicitly on access to relevant data.

Our project has been expanded in scope, and this new direction was resolved following discussions with analysts, sector contacts, and researchers in one-on-one discussions and a workshop. Specifically in FY 2013, we are evaluating and documenting the existing multicultural name matching of algorithmic work by processing trial data and submitting scores to the MITRE Challenge™ multicultural name matching database. While early MITRE results have indicated the existing algorithm has some value (placing well above the average score in the challenge), pursuing this course of action would likely require a significant investment to improve to the point of creating a substantial advancing of the state-of-the-art. Following this task, we are identifying the data sources that are relevant to disrupting the illicit trafficking of strategic goods and their owners.

Additionally during FY 2013, we are in the process of identifying existing analytical tools appropriate for the domain. To complete this effort, a total of 57 different tools were evaluated and ranked. Using sample data, we are also researching the suitable models and data structures with particular attention the distributed analytics needs that are specific to this domain.

Our project plans for FY 2014 include evaluating the compatibility for resulting data models and existing and pending analytical approaches. We will also be researching various data structures and models to handle the security-based divisions of data. These domain-specific models and sample datasets are likely to be a critical resource for downstream project work (including pending projects throughout PNNL) and would be of great interest to the wider non-proliferation community for developing new analytical approaches.

PN13028/2509

Experts Inundated with Data: the Biomarker Problem

Jason E. McDermott

We are developing approaches to improve the accuracy, robustness, and interpretability of signatures generated from high-throughput data by integrating statistical approaches with expert knowledge.

The 'omics revolution in biology has provided investigators with a plethora of complex datasets based on comprehensive molecular analyses with the potential to solve a large number of important problems including treatments for diseases, bioenergy production, and detection of bioterror agents. The field of biomarker identification has been plagued by the problem that either a purely data-driven approach is taken or a largely expert-driven approach is taken. The biomarkers identified from these analyses do not generalize well to larger sets of data, may not be interpretable to a domain expert, and may not show sufficient accuracy for their application.

Our project integrates statistical approaches with expert knowledge, derived from large-scale community annotation efforts and explicit input from a domain expert during the process. This is an important area of research that will allow advancement in this difficult and highly impactful area. We will developed a pipeline for integration of data sources that will allow the generation of better signatures based on incorporation of expert knowledge. Additionally, we will develop tools for classification and expert feedback on statistical integration processes. For example, we identified more robust signatures of processes for chronic obstructive pulmonary disease (COPD), which served as an application during our development process, resulting in a published paper.

In FY 2012, we developed the aforementioned pipeline for combining statistical analysis with expert knowledge to identify biosignatures from proteomic data. These biosignatures are capable of classifying samples accurately based on their disease state. We performed data pre-processing by designing a dilution series for mouse plasma proteomics, ran these proteomic samples, and completed an initial quality analysis on the resulting dataset that will serve as the first high-quality gold standard dataset generated specifically to evaluate the protein quantification problem. Additionally, we completed work on a proteomics data pre-processing pipeline, including quality assessment, normalization, and protein quantification using trend analysis. Finally, we created a method for principal component analysis on data with non-random missing values.

For FY 2013, we refined our pipeline for combining statistical analysis with expert knowledge to identify biosignatures and applied it to proteomic data from a mouse model of COPD. We further showed that these signatures were more robust than traditionally derived signatures by validation in a human COPD dataset. Additionally, we developed methods for statistical analysis of data with non-random missing values and published a paper describing analysis of the dilution series proteomics dataset. We helped to develop the data exploration tool Trelliscope, which allows experts to interact with the signature identification process by visualizing, ordering, grouping, and filtering individual features used in classification in a domain-agnostic fashion.

Additionally within this past year, we began applying ISIC to a complex and important dataset from a large-scale study of human ovarian cancer in the Cancer Genome Atlas. We expect that the application of our tool to this dataset will yield new biological insight and improved signatures of response to the standard platinum-based therapy for this disease. Additionally, we are expanding the tool to allow integration of multiple sources of 'omics data, including transcriptomics, proteomics, DNA methylation, copy number variation, and single nucleotide variants. This growth will allow us to make better use of the comprehensive data available in this dataset. We continued to improve our proteomics approaches and tools to the identification of signatures and to the investigation of better methods for statistical analysis and experimental design overall. This has resulted in the publication of several papers, including one in the journal *Nature*. Additionally, we have multiple proposals currently under review that make use of the tools and approaches developed under this project.

Our plans for FY 2014 include working with another similar project to apply our approaches to their specific problems. Preliminary discussions between our groups have suggested that this will prove valuable to further their goals of identification of community-based signatures of environmental processes. Also, we will continue to utilize the tools developed under this project in our separately funded cancer project.

PN12028/2429

Fishing for Features: Discovering Signatures when the Underlying Phenomenon is Poorly Understood

Kristin H. Jarman

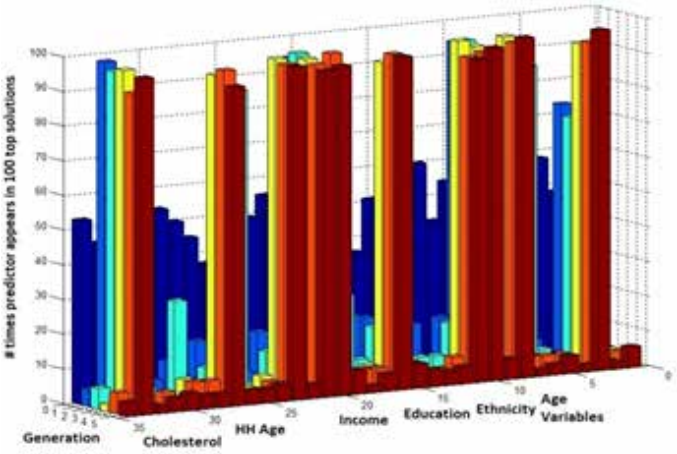
We are uncovering methods for finding and validating signature features for high dimensional, poorly understood phenomena are being developed. This work is relevant to high throughput and big data applications.

We live in the age of big data, and it touches virtually every aspect of our lives. Scientists can quickly collect terabytes of information on everything from a single atom to a complex living organism. Search engines and social media contain huge volumes of personal data on individuals. Online databases contain scientific information, consumer preferences, business performance, economic indicators, and so on. As data availability increases, researchers continue looking for ways to use this information to model increasingly complex phenomena. A great deal of work has been done and continues to be done on extracting signature features from large stores of data. Feature extraction (or variable selection) techniques have been developed for virtually all data types, from simple sensor measurements to text, images, and combinations thereof.

The state-of-the-art progresses, however, currently feature extraction methods that suffer from a similar weakness: the assumption that all potentially relevant information has been captured. This presents no problems for traditional scientific studies, but it limits our ability to develop signatures in situations where relevant data are missing, where the underlying physical or conceptual model is weak, or where available information is poorly catalogued. Methods for constructing signatures in such cases are needed, and this is the goal of the current project.

We are developing a two-stage framework for discovering signature features of complex or poorly characterized phenomena. In the first stage, a greedy strategy will be used to discover new features that are potentially relevant to the underlying phenomenon, the source of such features being internet databases, scientific and government reports, and scientific measurements among others. This greedy strategy is likely to identify spurious correlations and produce false discovery. Therefore, in the second stage, features will be pruned based on the strength of their association with the underlying model. The approach will be applied to two problems: identifying exposure risk factors for inflammation and other health conditions; and linking genes to enzyme activity for processes relevant to biofuel production by *Bacillus subtilis*. If successful, this research will produce a robust capability for discovering signature features in poorly understood, big data applications.

Our first year of research in FY 2013 focused on the development of a robust feature pruning methodology that can find and validate relevant variables or features in high dimensional datasets. This work addresses the following important issues: 1) selecting features in the presence of mixed data types; 2) identifying the best features when the classes are imbalanced (or the phenomenon is rare); and 3) ranking features by importance and assessing the significance of each.



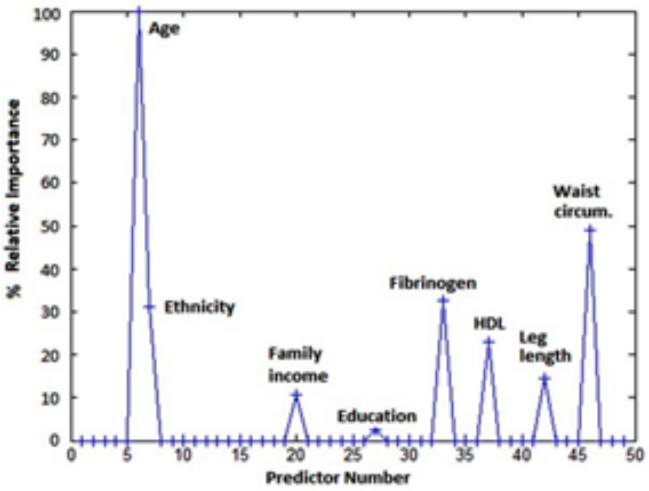
Proportion of predictors present in the population of solutions (z-axis) as a function of generation number (y-axis) with the last generation in the forefront.

The developed feature selection methodology can be applied to datasets containing a mixture of continuous and categorical variables. Using a Genetic Algorithm (GA), the methodology explores a dataset and selects a small set of features relevant for the prediction of a binary (1/0) response. Binary classification trees and an objective function based on conditional probabilities are used to measure the fitness of a given subset of features. To date, the work has primarily employed the U.S. Centers for Disease Control's National Health and Nutrition Examination Survey (NHANES) database, which has been used to show how the current feature selection methodology can be used to identify risk factors for diabetes. Results show that our

PN13030/2511

algorithm is capable of narrowing the set of predictors to about eight factors that can be validated using reputable medical and public health resources. This work has been presented at an annual conference and was published in the conference proceedings.

For FY 2014, our plans include applying the feature pruning methodology again to the NHANES dataset to identify environmental risk factors related to C-reactive protein (an indicator of inflammation), work that we intend to publish in a peer reviewed journal. Additionally, work will begin on the greedy approach to identifying new signature features from online databases. In particular, existing genomic, proteomic, and metabolomic datasets related to isoprene synthesis by *B. subtilis* will be used for empirical data. Online databases will be used to link genes, proteins, and metabolites in terms of their functional similarities. A greedy strategy will be used on these linkages to target potential variables for inclusion into a signature model. This greedy feature identification process will then be iterated with the existing feature pruning methodology to construct a signature for isoprene production by *B. subtilis*.



Relative importance of the predictors present in a GA solution. The y-axis represents the proportion of individuals in the testing set directed down the tree using the variables shown inside the box, where HDL refers to HDL cholesterol. According to this analysis, age has by far the largest impact.

Geological Sequestration Software Suite Framework

Chandrika Sivaramakrishnan

We are designing, developing, and deploying a software framework with effective capabilities to evaluate, model and manage sites for geological carbon storage.

Advanced multiphase flow and transport modeling codes coupling physical, mechanical, chemical, and biological processes are expected to play a crucial role in understanding and evaluating the feasibility and long-term effects of sequestering CO_2 in large-scale deep geologic reservoirs. However, the process of managing and interpreting raw data, building the subsurface domain conceptual model, and transforming/scaling into the numerical model for evaluation by these types of simulation codes is a tremendous challenge even for experienced modelers. The number of data sets and number of tools applied under varying conditions is extensive and difficult to manage. The process is also iterative, leading to tens and thousands of simulations run for a single site. Organizing, managing, and tracking data at each stage are critical to analysis integrity, verifiability, and repeatability yet is ad hoc, time-consuming, and error prone.

In this project, we are designing and creating novel provenance capture and analysis, and tool integration frameworks for carbon sequestration modeling. These capabilities will provide the ability to track data and assumptions in a given model, and make available a broad toolset for modelers. We will integrate them with the Geological Sequestration Software Suite (GS^3) and demonstrate their capabilities in support of uncertainty quantification and visualization. GS^3 is a set of tools integrated through a software framework that supports the modeling process used to evaluate and monitor CO_2 sequestration sites.

In FY 2010, we developed the initial version of GS^3 as a software framework specifically for the carbon sequestration domain. We combined version control system with web-based Semantic Mediawiki to provide robust versioned data storage along with an interface to annotate, semantically tag, and search data. We developed metadata extractors to parse geologic well logs and present them visually using plotting tools and Google map interface. During FY 2011, we refactored GS^3 into Velo, a generic scientific knowledge management framework, and deployed GS^3 as a carbon sequestration specific customization of Velo. This refactoring enabled the use of Velo across several scientific domains as many face challenges similar to carbon sequestration. For example, many need the ability to parse large volume of initial data, document missing data and assumptions, and track of data

Velo architecture

used throughout the project life cycle. Finally, in FY 2012, we updated to newer version of underlying GS^3 technologies (Mediawiki and Semantic Mediawiki) and improved many core functionalities such as tree-based file navigation, copy/paste functionality through web interface, rich text editor for user annotation, and editor for ASCII files within the wiki.

During FY 2013, we continued to add new capabilities like support for running parallel jobs on PNNL Institutional Computing clusters, enabling us to integrate and demonstrate parallel uncertainty quantification (UQ) runs developed under the UQ project, a new interactive provenance graph interface through which users can walk the provenance graph to explore various captured relationships and creation of the STOMP Input Advisor tool. This scenario provides users with appropriate reference documents and code snippets to guide them through STOMP input file generation.

The developments of this past year were conducted in parallel with the deployment of GS^3 for different projects, giving the opportunity to test the platform and its different tools on real cases. In particular, we completed the contract with the U.S. Environmental Protection Agency (EPA) Region 5 to help them process UIC Class VI injection well permit applications, and we initiated a new contract with EPA Headquarters to build their CO_2 storage data system, which will continue in 2014. Another success has been the adoption of GS^3 by the FutureGen 2.0 project as its main platform to manage data and simulations for the years to come.

Lastly, FY 2013 was devoted to making sure that GS^3 has a solid foundation for the future: six invention disclosure reports have been submitted on Velo-GS^3 and related products, and the open source licensing and quality assessment processes have been initiated.

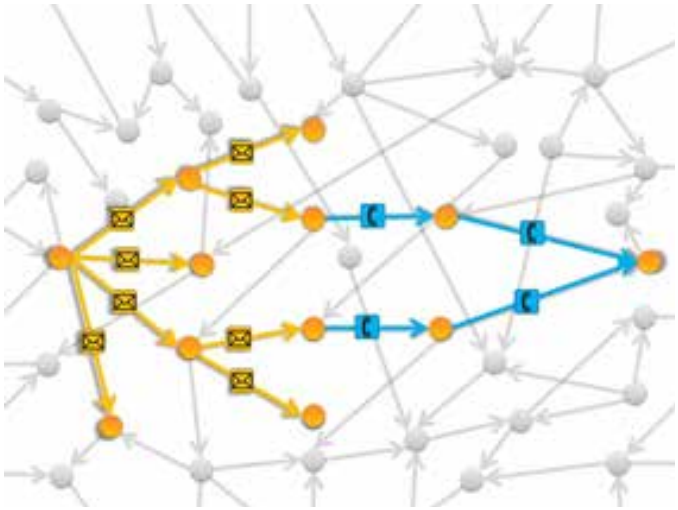
PN11036/2358

GRADIENT: Graph Analytic Approach for Discovering Irregular Events - Nascent and Temporal

Emilie A. Hogan

We are detecting time-evolving signatures in large volumes of raw data using one or a few passes. Our approach is reducing data size by reformulating the problem in a graph theoretic or sparse matrix framework and leveraging novel high performance computing platforms to mitigate the increase in complexity.

The past several years have seen a renaissance in the study of graphs – or networks – and the development of graph analytics spurred by increased interest in social networks and the growing importance of engineered networks. Innovative new technologies such as the Cray XMT supercomputer have been designed explicitly to support graph analytic computation and are effective on irregular problems found in graph analytics. In addition to naturally occurring graphs, graph approaches can solve challenging scientific problems in very different fields ranging from graph-based 3D flow field visual analysis to bi-simulation for discrete, continuous systems. With the continuing explosion of data across the scientific and analytic communities, new approaches are required that can find appropriate, compact representations of underlying data to support discovery and analysis.



Example of collection of paths in communication graph

Our research will advance the current state-of-the-art by using a graph analytic approach to support signature detection and analysis on extremely large, dynamic, streaming data sets across a broad range of domains. We will use large-scale graph analytic techniques to enable dynamic signature detection in massive data sets. The uniqueness in our research is reducing the raw data by transforming the signature detection problem into a graph representation and using large-scale graph analytic techniques and specialized, high-performance computing platforms to tackle this problem.

In FY 2013, we continued our work on the pass-the-hash application from the previous year while also focusing on a new data source, communication data, for our graph analytic signature detection. We obtained access to telephone and email records within the PNNL organization for 9 months. In both cases, we have the caller (sender), person called (recipient), and time of call (email). The data totaled 26.6 GB in over 84,896,837 records, which created 86,013,804 edges and 789,284 nodes. We anonymized the data so that names, email addresses, and phone numbers are not retrievable. Our goal was to identify patterns of escalation in communication from lower levels of the organization to higher. This information can look like a chain of email forwards with each successive email in the chain going from a lower level to a higher level. We observed that this pattern is identical to that of the pass-the-hash cyber attack on which we began work in FY 2012, where instead of people, the vertices corresponded with IP addresses, and we have the credential level instead of organizational level.

Using the above data, our approach was two-fold: obtaining path detection within large graphs and examining change characterization between two graphs. Through path detection, we find the pattern of interest described in the previous paragraph. We implemented code on the Cray XMT, which ingests a graph (in which edges are labeled with a timestamp), a target vertex, and a path length. The output is a list of paths that terminate at the target vertex, have length at most of the given path length, and respects time in that the time increase as the edges are traversed. We have not yet implemented the escalation of communication, but it is simply a matter of adding an "if" statement once the data are available.

In change characterization, we have a series of graphs and a function that assigns penalties to vertex changes (i.e., if the job code of a vertex changes from 71 to 72, determine how much change should we assign, as the person at that vertex may have gotten a promotion, or it may be a different person all together). We implemented the traditional lower and upper bounds for graph edit distance as well as applied a novel approach for combining these bounds to obtain a more accurate edit distance. Collectively, the results from the path-finding work this year were presented at two conferences, two workshops, and an invited talk. A paper on edit distance is in progress.

PN12037/2438

Hierarchical Signature Detection in High-Throughput Environments

Luke J. Gosink

This project will perform the fundamental research required to support signature detection tasks in hierarchical, bandwidth-limited systems typified by multi-type data.

Many phenomena of interest are detectable only by measuring large hierarchical systems of distributed multi-type data. As these systems continue to grow in size and complexity, rates of data generation too are accelerating. This combination of system size and complexity, along with accelerated data-generation rates presents a huge, compelling disruption for signature detection tasks. More specifically, detection tasks in hierarchical systems are increasingly failing not because of a signature's inability to detect an event accurately, but because of the inability to process data and detect the event within critical time requirements. Conventional analytic systems fail to address this problem because they use algorithms that require data to be collected, aggregated, and then processed at a central location. In this context, data bandwidth limitations become a bottleneck that increase the time it takes to process and detect events of interest.

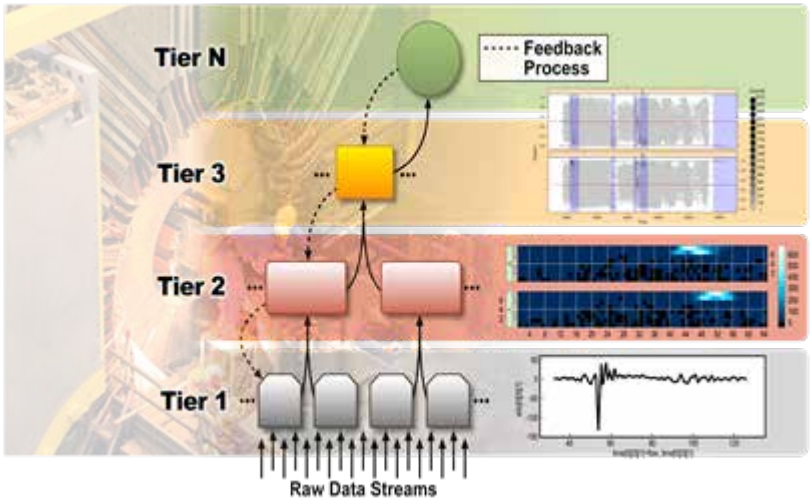


Illustration of the hierarchical signature detection statistical framework

This project presents a statistical framework for signature detection based on three methodologies that will readily scale to hierarchical systems. First, we use genetic classifiers to detect precursor events in data at individual nodes in the hierarchical system. These precursor events are the key processes or activities likely to (or must) occur prior to the primary event. Second, we extend these classifiers with information fusion approaches that can combine and propagate the evidence generated by these classifiers to higher level nodes in the hierarchy. Finally, we use this evidence to drive a predictive capability based on conditional random fields (CRFs). With this new framework, our work presents a method to be performed in parallel as a set of smaller, intermediate detection tasks. This predictive component leverages all existing evidence to assess the probability that the primary event will occur, thereby creating early situation awareness that will cue contingency and mitigation actions in the hierarchy.

In FY 2013, we completed the design and implementation of all three components in our statistical framework, including the realization of the theory as working code. Both the framework and its implementation were validated with tests that used synthetic data to assess the ability of our approach to: 1) detect precursor events across nodes in the hierarchy based on genetic classification; 2) aggregate and propagate classifier-based evidence throughout the hierarchical system; and 3) leverage propagated evidence to predict successfully that an event will (or is beginning to) occur. The test results indicate that the framework can detect events much faster than conventional systems: the framework reduced the mean detection time for events by approximately 86% compared to conventional detection methods.

In addition to the development and validation of our statistical framework, we began to extend this framework to the Belle II high-intensity physics experiment. The objective of this application of the framework is to help detect and differentiate between different subatomic particles in the experiment's iTOP detector; this detector generates data at a rate of 10 TB per second. Toward this effort, we developed and successfully validated two levels of hierarchical signatures for the experiment's iTOP detector that can discern between two important types of subatomic particles.

For FY 2014, we will continue to test our statistical framework and continue integrating our framework into the Belle II experiment's iTOP detector. To demonstrate the framework's generalizability and robustness, we will also extend our statistical framework to a cyber application, where we will gather evidence across a network and use this evidence to provide early situation awareness for specific cyber events of interest.

PN13014/2495

Immense Social Media Stream Analytics

Courtney D. Corley

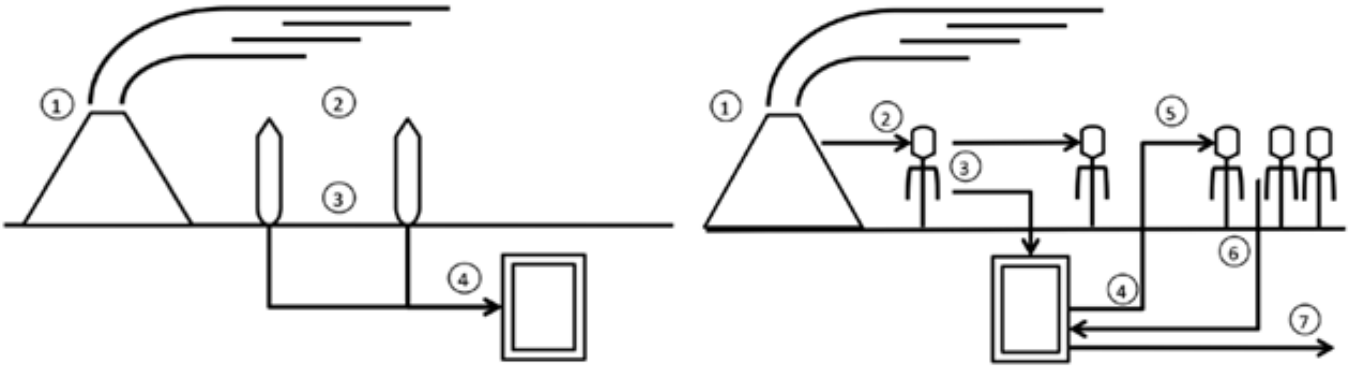
Knowing how social media is used during and in response to anticipated and unanticipated events (such as natural disasters, speeches, and crises) enables more accurate measurement of the potential effect of the events and informs planning and response decisions.

Today's analysts need methods for efficiently identifying and monitoring significant events and measuring their effects as expressed through social media. Existing analytic methodologies do not accommodate the innate characteristics of social media; they are limited by the inability of existing computational methods and systems to process the immense volume of data in real-time and the inability to cope with intentional or unintentional message attenuation such as spoofing, damping, and thematic and lexical drifts. These types of analysis methods on social media streams remain a capability gap.

The objective of this project was to design a system for interrogating immense social media streams through analytical methodologies that characterize topics and events critical to tactical and strategic planning. First, we created a conceptual framework for interpreting social media as a sensor network. Next, we addressed two scientific challenges: to understand, quantify, and baseline phenomenology of social media at scale; and to develop analytical methodologies to detect and investigate events of interest. The ability to process billions of social media posts per week over a period of years enables the identification of patterns and predictors of tactical and strategic concerns at an unprecedented rate through SociAL Sensor Analytics (SALSA). Ultimately, this capability illuminates the phenomenology of social media, giving power to the analyst to inform strategic reporting and planning. The results of this research will inform the detection and characterization of emergent and anomalous behavior in social media.

SALSA clusters topics in a manner that uses the content and user as the units of analysis. SALSA implements topic clustering through using the dot product similarity metric between authors and their hashtag usage, over the course of a specified time period. Within each cluster, the sum of the individual hashtags determines the cluster volume (likewise for frequencies). There are several pitfalls: hashtags are merely a proxy for relevant message topics. Hashtags are terse (and not always literal) representations of the message topic at hand and are subject to frequent changes in usage over time. Moreover, this method of clustering is static and provides clusters of hashtag usage by author over the specified time period. This method's accuracy of classification beyond the dates specified is severely limited by the measurement of the 1,000 most-used hashtags and their use by individual authors of that time period. This method of clustering, however, serves as a successful proof-of-principle in eliciting social signals from topic categories. The clustering results provide actionable information in spite of the usage of a naïve brute-force algorithm.

To develop our algorithms and framework, we utilize over 10 billion social media posts, including tweets, blogs, microblogs, mainstream news articles, forums, and classified ads. From these, 60% are in English and the other 40% are distributed among at least 60 languages. Since the project's beginning, it has made use of over 2.5 million



Left: A traditional sensor network event detection: 1) something happens, 2) sensor acquires measurement, 3) sensor records measurement, 4) system stores measurement. Right: Social sensor network event response: 1) something happens, 2) person receives stimulus, 3) person communicates response, 4) system routes message, 5) people receive message, 6) people communicate response, and 7) system routes message. Users of social media are akin to physical sensors, creating a global network of measurement capabilities, including all the inherent problems of physics-based sensors with the added complexity of spontaneous human behavior.

PN12042/2443

core hours and occupies 170 TB of space. The time-series representations of social signals are modeled to study social signals and inform phenomenology. The features may be studied individually or in an ensemble to calibrate the social signal that may be attenuated by exogenous factors. There are myriad potential signals that could be measured from the data stream (e.g., posts per hour directed from user to user @s and mentions), highest volume users per hour, or most influential users per day. Determining the markers for these social signal aberrations such that we can anticipate emergent event types—not predicting that a meteor is about to crash but predicting whether a signal aberration—is the result of a random event or an information campaign.

As with physical sensors, SALSA must quantify a baseline from the social sensor measurements. Once a baseline is calculated, it is used to detect aberrations in the sensor data. For this research, we define an aberration as an observed value that is significantly different than expected. The method of determining the expected value provides a means of quantifying the degree or significance of the aberration. As with baselines, the value of interest is the volume of social signals that match specific criteria at a particular point in time. SALSA implements this as a brute-force approach in which potentially millions of distinct criteria are regularly evaluated (hourly or daily) with time-series models to quantify and rank aberrations. This allows an analyst to characterize and quantify the types of aberrations that occur during events of interest.

Previous work that created the ability to ingest and search rapidly over large amounts of near-streaming web data, along with the development of nearly 50 descriptive analytics, has allowed the team to develop a battery of aberration and pattern detection in web social media data. Over the past year, statistical time series and Fourier analysis has allowed for the generation of patterns of event evolution, both in the direct measurement of the near-streaming data along with indirect measurement of these events' diffusive properties over large networks of people. We are now in an extremely unique position to observe, study, and report on the emergent behavior of large numbers of interacting people in large social networks. This work has contributed significantly to the characterization of the overall behavior of web social media, giving us a highly sought after eagle-eye-view of a very noisy environment.

Moreover, practical and rapid analyses of social media response to real-world events have been made possible to the scientists at PNNL. Examples include the analysis of aggregate sentiment during the events of the Boston Marathon Bombing. The research paradigm of the "social sensor" also resulted in creating a foundation for biosurveillance applications in both citizen generated data, social media, and mobile decision platforms. Our team will continue to develop these analytics into a useable tool as well as continue fundamental research in improving this application of data centric modeling.

PN12042/2443

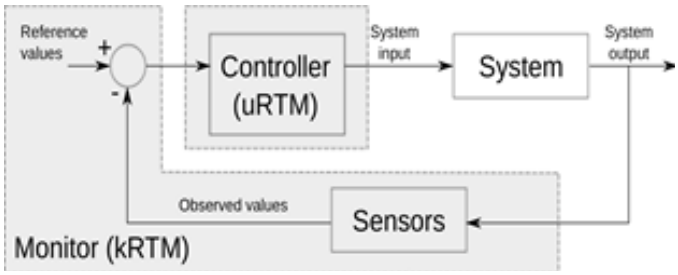
Integrating Advanced Optimizations for Extreme Scale Systems

Adolfy Hoisie

Achieving exascale performance with a limited power budget is essential to ensure scientific progress is many areas, including biology, chemistry and national security. To this extent, an automated and concerted approach between hardware and software is necessary to ensure the most efficient use of resources.

Limitations on energy consumption, uncertain technology, poor reliability, and programming difficulty form a daunting challenge for exascale systems. Currently, these problems are addressed manually by programmers or system administrators. However, burdening programmers with low-level details is counterproductive and overlooks the benefits and capabilities of the exascale system. An automated system that collects information at runtime accurately and provides actionable tools to improve system's steer performance and power consumption is necessary to ensure the most efficient use of resources.

This project focuses on one essential aspect of the automation of system management so that an exascale system will operate at the most efficient level. Specifically, we developed performance and power steering capabilities that collect information pertinent of program behavior at runtime; generate a dynamic model of power and performance based on this information; and verify, also at runtime, whether the application's performance and energy use expectations are within desirable range. The research advances enabled by this work are multiple. We designed and implemented a low-overhead, high-precision runtime monitor that collects system's and application's characteristics and reacts to changes in the execution environment caused by power or resiliency issues. The complete system works like a closed-loop control system.

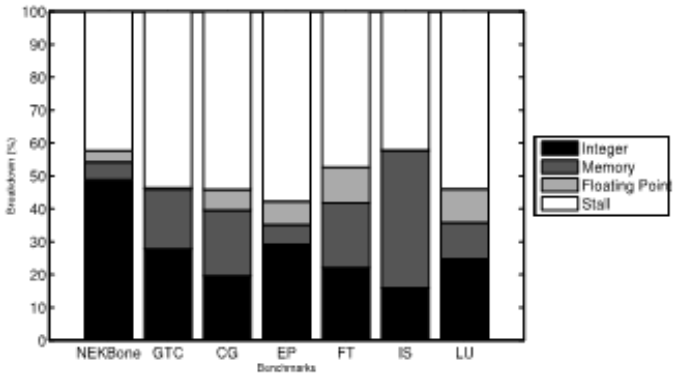


Closed-loop control system

In FY 2012, we designed and developed two basic system layers: the system monitor interface (SMI) and the runtime monitoring system (RTM). The SMI specifies a common interface between the low-level system component (hardware and operating system) that provides raw information and high-level components (runtime system and programming language) that use the information provided. The RTM constantly monitors the execution of applications and collects information and statistics used to develop a dynamic model of the applications and to verify whether the application's performance and energy consumption meet the expectations provided by the dynamic model. The RTM consists of a specific kernel module designed to perform monitoring effectively with high frequency and low overhead.

In FY 2013, we studied the power efficiency of MPI+OpenMP applications that when varying the number of MPI tasks and OpenMP threads per MPI task. There is an optimal operational point, but one of the configurations is not trivial to determine manually. Our monitoring system understands the characteristics of several scientific applications (e.g., Nekbone) and automatically determines the optimal number of OpenMP threads per MPI process to improve performance and power efficiency on modern multi-threaded architectures.

Current state-of-the art systems do not provide fine-grained, per-core power information, which severely limits the design and development of power- and energy-aware software algorithms. To this end, we designed a generic methodology to develop per-core proxy power sensor models that provide the power consumption of a single compute core in system. This information is essential to understand power-saving opportunities and their impact on performance and power consumption. We integrated this per-core power sensor models in the monitoring system developed in FY 2012, hence our system can now monitor and power and energy consumption besides performance. We used this system to characterize the power consumption of DOE and other parallel applications and analyze the power consumption of the main system components. The figure shows the system power breakdown for the tested applications.



Power breakdown for scientific applications

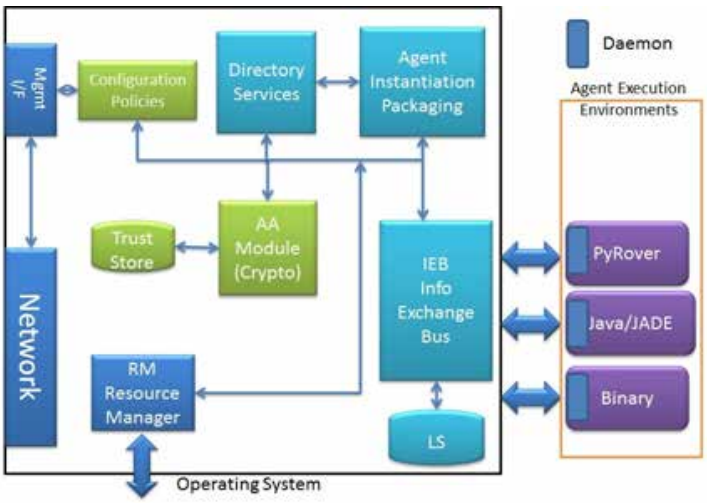
PN12043/2444

Intelligent Networked Sensors Capable of Autonomous, Adaptive Operations in the Electric Power System

Jereme N. Haack

Our project creates a framework that adds "intelligence" to command-control driven sensors and actuators used in today's electric power system and forms the basis of a new distributed operations and control paradigm.

The number of intelligent devices (including sensors) connected to the electric power system is expected to increase several orders of magnitude by the year 2020. The devices include synchro-phasor sensors deployed throughout the transmission system as well as smart meters in the distribution system. Newly installed sensors in the power grid are expected to generate terabytes of data every day. Most of the measurements performed by sensors need to be sent and received in real-time. Information networks that are part of the electric power system, however, are not equipped to handle data in these types of volumes and do not provide reliable or real-time delivery assurances. The power systems of today are also not capable of managing two-way power flow to accommodate distributed generation capabilities due to lack of system flexibility and concerns about system stability. The information collected by sensors can also be used to detect cyber attacks directed at the power grid. These challenges hinder our ability to utilize the intelligent devices to their fullest abilities.



The Intelligent Device Platform provides a secure, robust, resilient framework that allows software agents to be deployed for control and monitoring in the electric power system.

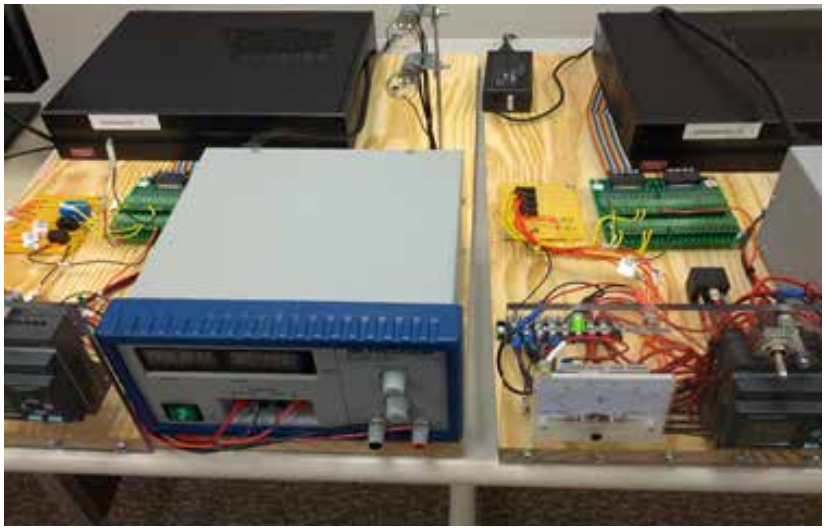
Our research project creates a framework to add intelligence to the sensors, which manifests itself in the ability to adjust measurement methods and in the functionality of performing measurements in multiple dimensions, such as in information networks and the electric power system simultaneously. The ability to observe and correlate measurements from multiple domains will significantly improve the cyber security of control systems used in the electric power system. These intelligent sensors are capable of sharing information through the various layers of the electric power system to enable two-way power flow to facilitate integration of distributed generation as the power systems of today are unable to accommodate distributed generation capabilities due to lack of system flexibility and concerns about system stability. Finally, when paired with distributed autonomous controllers, sensors can form the basis of an intelligent system that supports micro-grids and islanding.

Our research will lay the information network groundwork for distributed operation and control of the electric power system, especially at the distribution layer where the end customers are being served. The intelligent sensors and the distributed control paradigms enabled by this intelligent sensor are expected to improve system reliability and support the integration of distributed power generation.

In FY 2011, we developed the design of our software platform Volttron™, which can be used to enable distributed intelligent devices in the electric power system. The primary strengths of this platform are that it is hardware architecture agnostic (requiring a modern CPU with MMU support) and it supports multiple software agent execution environments simultaneously. It provides robust security even when network connectivity is disrupted and provides distributed, decentralized directory services for location functions and agents. For FY 2012, we completed the initial version of the agent platform based on the design document from FY 2011. We assembled the hardware purchased from the previous year into a testbed and used it to showcase demonstrations of our software platform, including coordination of plug-in electric vehicle charging and cooperative load scheduling.

PN11045/2367

In FY 2013, we integrated the generator trip detection algorithm from a related project. Our hardware testbed built in FY 2012 was demonstrated at a national conference, which led to several potential collaborators. We are rebuilding the testbed using ARM-based boards after Volttron™ was changed to work on the platform. Volttron™ was integrated into the PNNL Lab Home, where it coordinates the charging of electric vehicles with household energy usage. We plan to add a GUI interface to make demonstrating this integration easier, and we are currently working on a publication based on this work for an IEEE conference.



Hardware testbed platform

As a result of this project, Volttron™ became a cornerstone of large DOE-funded project for coordinating roof top HVAC units. In addition, Volttron™ will enable controlling for the charging of the Benton Franklin Transit electric bus system throughout Benton County, Washington state. Finally, Volttron™ will back the testbed focus for several integral PNNL projects.

Kritikos: Identifying Cyber Assets and Assessing Criticality in Terms of Business Processes

Thomas E. Carroll

We are developing Kritikos, a (near) real-time enterprise introspection method for discovering cyber assets, identifying functional relationships and dependencies between assets, and assessing the importance of the assets to business processes that they serve.

Computer and network operations and cyber security focus on information technology are largely naïve of an enterprise's manifold business objectives and processes. The lack of business operations information hinders the defenders' sense making, assessment, and management of situations. Current practices for identifying assets and mapping functional relationships are generally characterized by laborious, time-consuming, and error-prone processes that are difficult and expensive to deploy in large, dynamic enterprise environments. By linking business processes with their supporting cyber assets, *Kritikos* enhances cyber defenders' awareness and improves the quality and speed of situation assessment and management that align cyber security with the objectives of the enterprise.

Kritikos works as follows. A network model of the functional associations between services is constructed by discovering recurring spatio-temporal patterns in the NetFlow record set which, with its derivatives, are embedded instrumentation for the collection, characterization, and export of network traffic flow information and statistics. These recurring patterns arise from human-initiated machine-to-machine interactions that occur in the everyday operation of the enterprise. A business model that has processes annotated with essential services is correlated to the network model to identify other process essential assets. Asset criticality is then measured as a function of business process importance.

Our focus in the past year was developing a method to build network models. While related work has constructed network models based on the existence of systems or communicating entities, *Kritikos'* goal is to discover and model the functional associations between network services. A functional association between a service A and service B arises whenever A and B are accessed. These associations arise because of graph associations observed as paths in a graph (a service A depends on a service B and communication is observed) or as temporal associations in which the graph of communication does not provide for a path between entities but does recur in time (services A and B perform function in time). Temporal associations occur because not all communications may be observed. This may be due to the placement of sensors or because of external synchronization mechanisms.

We developed our approach that uniquely discovers functional associations by first transforming the NetFlow record sets into a spectrum representation and then using a machine learning algorithm to identify recurring spatio-temporal patterns in the spectrum. Patterns recur from users initiating chains of machine-to-machine interactions. To discover these patterns, *Kritikos* creates a spectrum from a 5-sec moving window of NetFlow data. We step the window forward in time, starting at each NetFlow data point, and generate a vector of offsets to the next data points within the next 5 sec. The vectors from a record set can be divided into subsets of vectors with recurring offsets within each set. Each will then contain a similar network service, though we do not encode the network addresses or port information. A type of artificial neural network, Self-Organizing Map (SOM) ingests the vectors and generalizes interactions by spatially arranging similar interactions together. Patterns are extracted from the SOM and data mined to label elements within the patterns, such as network addresses and protocol information. At this junction, we construct graphs of services that are functionally associated with one another. In a connected component, the nodes are services, and the edges reflect associations.

We commenced verifying our approach in a controlled testbed environment. A network for a virtual enterprise was modeled and built in the testbed. Simulated user agents distributed throughout the enterprise used applications that caused network traffic to be generated. The initial results of our approach operating in this environment indicate that we are capable of identifying both graph and temporal associations. More exhaustive testing is reserved for next year.

Network model construction is only one key piece of *Kritikos*. Our goal for FY 2014 is to identify the relationship between assets and business process and to assess criticality. To this end, we will construct business processes for the testbed's virtual enterprise to be modeled in a business modeling language. Once we determine a workable representation, we will correlate the information to create a combined network-business model for the virtual enterprise. This model will reflect both the network functional associations and the relationships between business processes and services. A service's criticality is then a function of the importance of the business processes that employ it and extends through functional associations.

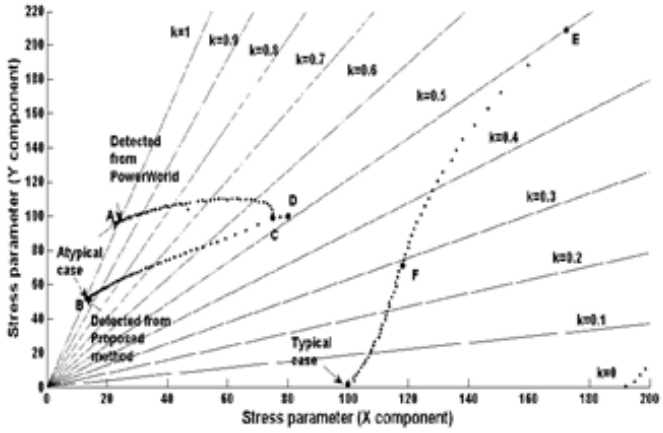
PN13039/2520

Linear Algebra Solvers and Associated Matrix-Vector Kernels for Power Grid Simulations

Barry Lee

This project researches the development of advanced linear solvers and matrix/vector kernels, key enablers for power grid modeling, through high-performance computing simulations. These advanced solvers and kernels will enhance the modeling capability for current and future power networks that use high-performance computing.

DOE's Office of Electricity has invested heavily in the investigation of more flexible, reliable, and higher capacity electric energy infrastructures that incorporate new engineering technologies (e.g., sensors and renewable energy sources), and other DOE program offices have invested in high-performance computing facilities to advance science applications through computer simulations. The purpose of these investments is that modeling and computer simulation will play a major role in the fundamental and applied research in advancing the future power grid systems. As these systems and their models increase in complexity and size, the utilization of high performance computers will be a necessity. In particular, high-performance computing will be needed to solve efficiently the mathematical equations underlying the more complex power grid models of the future. In the past, for less accurate models, simpler equations were often solved using commercial software packages and on desktop computers. For advanced models, to obtain the solutions to these mathematical equations, carefully designed parallel solution procedures that use high-performance computing will benefit the simulation capabilities.



Comparison of atypical, proposed, and PowerWorld linear solvers

The computational bottleneck of these solution procedures often arises in inefficient linear solvers and matrix-vector kernels. Thus, the purpose of this project is to develop efficient, stable linear solvers and computational kernels for power grid applications simulated on high-performance computing. This efficiency is measured by both fast turnaround time for the linear solvers and matrix-vector kernels to perform their operational tasks and their ability to utilize the computational resources on the high-performance computing (e.g., large problems solved by careful usage of more computational resources give faster turnaround time than when solved using less computational resources). This stability is a measure of the solver's ability to generate accurate solutions to the equations. Unfortunately, due to the mathematical properties of some equations, the numerical process of many solvers can dramatically amplify small errors to make the solutions totally invalid. One such relevant power grid application that leads to these undesirable properties in the mathematical equations is state estimation. Thus, the target goal of this project is to develop quality efficient and stable linear algebra software for power grid applications.

In FY 2011, we developed and implemented several computer simulators to generate some of the linear systems and matrix-vector operations arising in power grid applications. These simulators were necessary because they both generate linear systems that arise in power grid applications and expose the appropriate data structures that can better utilize high-performance computing. For FY 2012, we examined, developed, and implemented eigenvalue solvers with the appropriate transformations and preconditioners for the efficient calculations for small-signal stability problems. Succinctly, these problems analyze the stability of a power grid network under small disturbances about a stable operational point.

With this year's linear solver goal, we used the eigenvalue solvers to determine the voltage stability boundary of WECC systems with different stress loads. A possible boundary point is formulated as a generalized eigenvalue problem, which is solved using a shift-and-invert spectral transformation for an implicitly restarted Arnoldi (IIRA) method. To utilize high-performance computers, each processor simultaneously solves a different stress scenario using the generalized eigenvalue solver. To achieve efficient computation, a sparse matrix representation of the Jacobian is explicitly used, and an iterative Krylov method with a suitably designed precondition is applied in solving the linear systems at each cycle of the shift-and-invert spectral transformed IIRA.

PN11046/2368

M&Ms4Graphs: A Multi-Scale, Multi-Dimensional Graph Analytics Framework for Cyber Security

Sutanay Choudhury

This project is delivering a set of scalable, uncertainty-aware, multi-scale graph analytics tools for cyber analysts and mission defenders.

The current approach toward mission assurance and cyber defense is primarily reactive. Automatic discovery of critical missions and their relationships with users, applications, and the physical infrastructure is a major challenge. Unfortunately, chartering the response to a security incident is a manual process with high complexity and uncertainty. In addition, the cyber system is constantly evolving and generates prodigious amounts of data. Therefore, a scalable framework to provide continuous updates of a system in terms of important security metrics is absolutely critical.

A network of mathematical abstraction, a graph consists of vertices (nodes) representing entities and edges (links) that indicate relationships between vertices. A cyber system can be efficiently modeled as a graph, where entities such as computers, IP addresses, users, and software services form nodes, with relationships between these entities as edges. In turn, these diverse relationships can represent communication between machines, association of a physical machine to an IP address, or a login-event of a user on a particular machine. Despite the intuitive appeal of such a graphical model, its complexity and scale can make analytics – especially near real-time analytics – challenging. A dataset spanning only a few minutes can be so large that it does not fit in the memory of typical workstation. Computational requirements can also be a limiting factor for analysis. These challenges are further aggravated by the dynamic nature of cyber networks, where the focus of analysis is on risk evaluation.

Network dynamics refer to a constantly changing environment of a cyber network further accelerated by the proposed moving target infrastructure. The evaluation of risk involves modeling uncertainty as a fundamental feature of the analysis. We are therefore modeling key behavioral aspects of an enterprise by studying the information flow across hosts as large-scale dynamic graphs. We are taking a theoretic, data-driven approach of modeling the dynamic behavior of a cyber network system at multiple scales that range from a single machine to an entire enterprise. The

Illustration of how services (red nodes) interact with a group of users (gray nodes). The services inside the blue circle impact a large number of users.

models will provide continuous metric based updates to provide situational awareness and enable a system that is resilient by design.

For the last year of work, we initialized our adapted methods to update the graph-based model in a continuous loop with local information to enable very fast computation of essential security postures and cost/benefit metrics. By accounting for the connectivity structure of the graph (who talks to whom) and attributes of the communication (using which protocol, how often, and for how long), we are creating a comprehensive model that describes behavior ranging from micro- (host level) to macro-scale (enterprise level). The intended suite of tools will reduce the time-to-insight for cyber analysis on massive, complex networks through parallel and distributed computing algorithms that scale not only on specialized high performance computing resources but also on clusters of commodity machines.

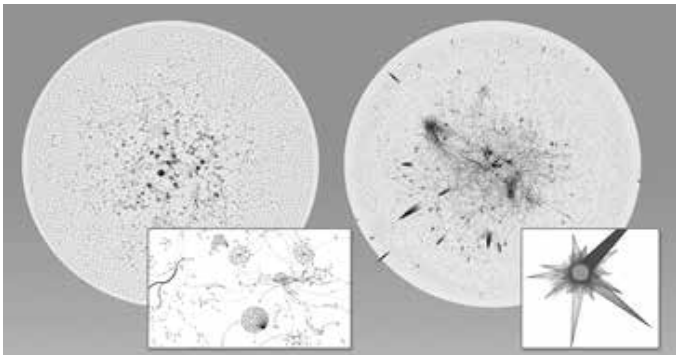
We developed a network-of-networks model to represent heterogeneous cyber datasets and motivated a multi-scale approach for cyber modeling, the results of which were presented at an IEEE conference. We established our formal foundations for multi-scale graph models, a multi-scale variant of the All Pairs Shortest Path algorithm, and created specific algorithms on probabilistic methods for reachability.

PN13037/2518

Specifically, we identified and validated selected graph theoretic metrics. Our preliminary data structures and test suite were developed, which included simple synthetic test cases and examples. We established our build framework, consisting of the unit test framework, preliminary test cases, and code scaffolding. The adaptation and prototyping of dynamic graph theoretic metrics were also made available.

Our OpenMSG version 0.5, an alpha version of the software framework for multi-scale and -dimensional graph analytics of cyber data, was released. Functionality included basic data structures and preliminary algorithms for reachability or the shortest path. We demonstrated simple functionality on a limited use case set after specific cases were identified. Collectively, our work on this project throughout the year has resulted in presentations at six conferences, two invited talks, and one workshop.

For FY 2014, we will perform a comprehensive evaluation of graph theoretic metrics for resiliency measurement and situational awareness. In addition, we will finalize the multi-scale models and algorithms to advance to performance optimization. Finally, we will release version 1.0 of OpenMSG to the open-source community. The anticipated features will include improved functionality both to graph theoretic metrics and resiliency measures that we expect to have achieved during the year.



The evolution of network traffic at multiple time-scales

Manifold Learning for Accurate Search and Locate Tasks in Image Datasets

Elizabeth R. Jurrus

We are attempting to produce an improved method for retrieving images from a data source given an example image. This is a challenging problem due to the lack of training data required for machine learning methods and the accuracy of current state-of-the-art methods.

The community at large is experiencing a rapid growth of data from airborne motion imagery and full motion video. The most accurate method for detecting objects of interest in these immense datasets is through human analysis. However, real-time detection of events in large image datasets is impossible. Therefore, there exists a significant need for automated methods to help analysts perform accurate, content-based image retrieval in large image datasets. Techniques for video analysis and image retrieval can vary depending on the task. Most algorithms in video analysis focus on continuous visual event recognition in realistic scenes. This is difficult due to the changing nature of individual's poses, appearance, and size. In addition, backgrounds for objects under surveillance can change along with lighting and occlusion. Similarly, image retrieval problems suffer from the same constraints without the time component.

We are developing a set of algorithms that will create models of image data, enabling search-and-locate tasks in image datasets (e.g., video surveillance data). Given a set of images that can be used to train our algorithms, we will build a description of the data in the images. This becomes our object model and enables searching for similar models in unseen images. Collectively, we will be creating an image processing pipeline with the following components: feature detection, selection of appropriate feature vectors (to describe the data), correspondence selection (to align shapes), shape modeling, and classification (with a manifold learning algorithm).

We believe that analyzing the data using a shape-based manifold learning algorithms to learn a low-dimensional subspace or manifold of the original feature space will better capture the variability inherent in these datasets. This learned manifold may be nonlinear, providing greater flexibility to explain complex data variations. Additionally in this framework, the image retrieval issue is reduced to a lower-dimensional search within the learned manifold, rather a search in the original high-dimensional data space. We are testing our method on the Caltech bird dataset, which contains 200 different classes of birds, and we plan to demonstrate this idea by learning a model that better distinguishes bird species.

We made the most progress on feature detection, feature extraction, and correspondence problem parts of our pipeline. We feel this puts us in a good position to generate improved classification rates over what is found in the literature. The following sections describe these areas in more detail.

We experimented with many of the current commonly used feature detection algorithms reported in the literature, including SIFT, SURF, Kadir-Brady, and Harris corner detectors. These computer vision feature detectors are based on responses to some type of image filters (typically gradients or second derivatives). We found that these techniques tend to return points of interest that contain high amounts of clutter, and we hypothesize that these contribute to the low accuracy rates in the literature. Alternatively, there has been recent work on learning features directly from images themselves, with the idea that such features will be a better match to the data. However, the features suffer from a deficiency in that they are not invariant to simple image transformations such as rotation. We formulated a dictionary learning algorithm that is invariant to rotations, which showed improved classification accuracy rates (> 51%) over state-of-the-art methods (49%). We also experimented with the feature vector creation techniques found in the literature and discovered that these can alter the classification accuracy between 5% and 10%. We are considering incorporating the weights from the basis functions of the dictionary learning method as a way to describe the features on which we are training.

We are additionally investigating a novel approach to handle the issue of correspondence between unordered features in images. Rather than find optimal correspondences between features, we are treating correspondence as a latent random variable that should be integrated out. A naïve approach to solving this requires summation over all possible permutations, which is intractable. We are exploring a method involving the importance sampling to reduce algorithm complexity, and our preliminary results show that this approach is feasible.

Finally, what have accomplished thus far is to develop an algorithm for learning models from "scrambled" data (i.e., the features come in arbitrary order). This is a common

PN13011/2492

property for data that comes as output of an image feature detector (e.g., SIFT). Another property of such computer vision feature detectors is that they return false positives, or clutter that is not part of the object of interest. Our method also handles this problem by probabilistically assigning features as foreground or clutter. The algorithm is scalable to large numbers of features by efficiently sampling high-probability correspondences, rather than solving the exact combinatorial matching problem. Our current tests on simulated data show that our method can produce promising results, and we are currently working to apply this to real image data.

For FY 2014, we will formulate a model for incorporating shape information into the image feature representation (i.e., relative position, orientation, and sizes of image features to each other). This is in contrast to the "bag-of-words" type models that do not take shape information into account. With this shape-based representation, we plan to use manifold learning in to capture the variations in our data, enabling more accurate identification of objects (or birds).

Mapping Molecular Dynamics Algorithmic Parallelism to Heterogeneous Architectures

T.P. Straatsma

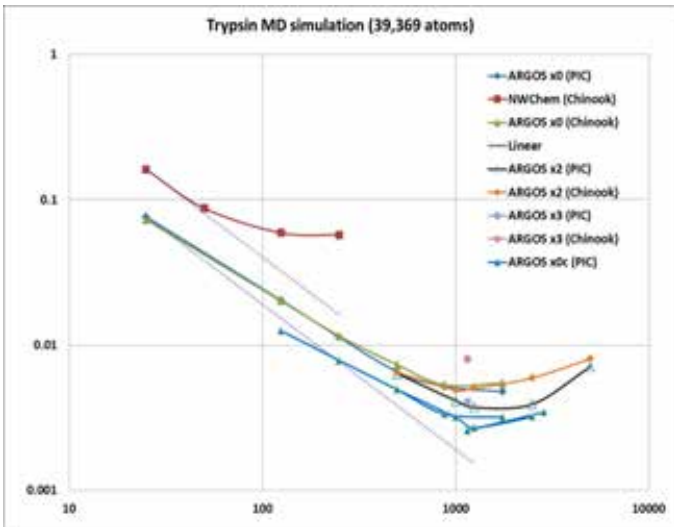
Novel computer science and software engineering approaches are applied to develop a high-performance molecular dynamics (MD) modeling and simulation implementation for biomolecular modeling research.

A principal and general challenge for biomolecular modeling and simulation in support of environmental molecular science involves the generation of extended molecular trajectories. These trajectories use advanced simulation methodologies to capture those features central to the research questions being addressed. This applies to understanding important biological processes relevant to human health and environmental issues such as carbon cycling and energy security by creating fuel from biomass. The molecular-level processes to be understood through the application of computer simulation include protein folding, protein docking, complex enzymatic reactions, and the association and function of large protein and protein-DNA complexes. The common situation in all these cases is the need for sufficiently large ensembles of conformations to capture the relevant events, which requires highly efficient computer codes to be designed and applied on the largest supercomputers. Achieving high scalability to remove current bottlenecks in applying MD simulation codes on very large processor count computer architectures is approached by formulating an implementation strategy that leads to significant performance improvements and through an extensive analysis of the algorithms used.

Based on a domain decomposition approach, we designed a new data structure to exploit locality and reduce communication and memory requirements. Distribution of the workload taking into account the workload based on relative orientation within the physical decomposition allows a number of processes to be used of up to 42 times the number of cells in the system. The most important challenge for scalability of classical MD is the need for synchronization after the evaluation of forces and after updating the coordinates. Large process counts make the use of global synchronizations, which is acceptable on simulations using less than 1,000 processes, quite inefficient. In the redesigned kernel, synchronization is made implicit by processes waiting for the number of expected neighbor contributions before continuing. This implicit local synchronization approach effectively "buffers" operating system jitter, resulting in much improved scalability. This feature is accomplished in part by a new Global Arrays feature put-notify, which provides a convenient means of using a push-data model for communication, instead of the traditional pull-data model.

The approach described above has been implemented in an MD kernel written in previous year and improved in FY 2013, extended to the application in heterogeneous systems and tested over a range of simulation conditions. Optimizations for intra-node parallelism and topology-aware assignment of cell-pairs results in a highly optimized simulation framework capable of achieving simulations of close to 70 ns/day. This is accomplished by communicating data needed by multiple processes on a node only once to that node, using a polling mechanism for each of the processes to determine the availability of the data. This approach is possible as one of the design features of Global Arrays. In the simulations, one core per node is reserved for the Global Arrays/Aggregate Remote Memory Copy Interface server thread, which is necessary to deal effectively with the communication requirements of these simulations. In addition, Multiple processes handle control tasks, such as global accumulation of kinetic energy for constant temperature and pressure simulations as well as with I/O requirements such as trajectory recording. Compared with the NWChem runs, the new kernel scales to an order of magnitude more cores and results in a wall clock time per step an order of magnitude smaller.

This research culminated in mid-FY2013 and included the development of multi-ensemble implementations, energy optimization using *Energy Templates*, and the implementation of free energy methodologies.



Wall clock timings for a 39k atom solvated trypsin simulation using the cell-distributed NWChem MD module and the newly designed cell-cell pair distributed kernel.

PN11049/2371

Multimedia Analysis of Cyber Data

Deborah A. Payne

Our goal is to create a multimedia analytic environment for cyber intelligence that identifies multimedia similarities from cyber data retrieved from web sites to infer relationships or common interests between entities.

Traditional cyber intelligence is performed by looking for indicators of threats in log data. This analysis has been applied to large structured or semi-structured text data summarizing the activity of a machine displayed in complex spreadsheets without inclusion of the multimedia content accessed. Contemporary analysis tools work with either log data or multimedia content in isolation, which do not provide the analyst with the ability to derive patterns of interest across information silos. The creation of a system that allows analysts to review multimedia data will aid in uncovering patterns that are overlooked when analyzing just the textual data. For example, an analyst might be interested in baking a cake and is searching for the text words flour, egg, and oil. With a multimedia system, the analyst can also search for images of cakes, flour, eggs, and oil.

This work is resulting in the development of approaches for applying multimedia analysis to cyber intelligence. The technical research and implementation are of use to analysts at PNNL and could be applied to a broader audience. This research leads toward a visual analytic system for multimedia analysis that can combine relationships in multimedia data with relationships in text. We have been identifying the features and entities necessary to exploit, extract, and learn the relationship patterns in multimedia content.

In FY 2012, we made advances in two broad areas: the creation of a framework for synthesizing multimedia data, and the development of a visual approach for linking image relationships with textual relationships. This set the stage for our progress in FY 2013, where we improved the results provided by our framework and developed a second visual approach to meet user need. The second visual approach, Topic Bubble, provides users with the ability to graph actors and locations while adding supplementary dimensions based on similarity of multimedia content. The Topic Bubble visualization aids in the detection of patterns in multi-modal content in context of complex Boolean queries like entity, key term, domain, and content. It provides at-a-glance understanding of multimedia content across multiple analyst and system created pivots.

The multimedia relationship framework created during the last year identifies patterns within image and video data by applying classifiers that run during data processing. Classifiers are created by first detecting features and then training the system to recognize images with similar features. Each media type has a variety of features associated with the type. For those images, we calculated several features, including dense scale invariant feature transform descriptors, pyramid histogram of gradient descriptors, and red-green-blue joint histograms. In FY 2012, we trained the system using a series of support vector machines – one for each feature type – using a pre-collected set of exemplar images producing a composite score.

The support vector machine approach that we applied in FY 2012 was adequate but not customizable across a range of image search and retrieval tasks. For FY 2013, we moved the training to an ensemble approach in order to achieve a more accurate, consistent predictive performance. We applied Bayesian Model Aggregation (BMA) to help our framework identify the features that are most important to use in each specific classification and to determine how to combine them. For example, the features that are best at classifying images of dogs might not be the same optimal features for classifying images of cars. BMA jointly utilizes all valid information and training inherent in each classifier to construct an aggregate. The aggregate represents a conditioning of "performance" based on an ensemble of assumptions and information rather than on any single constituent's "performance." The BMA approach to combining similarity metrics in an image classification task is a novel approach that has improved our classification results. We hope to apply this method in the future to allow custom-created classifiers on the fly and image search engines to become customized based on user feedback fed through BMA.

Our efforts additionally led us to a research topic applicable to the broad area of aggregation beyond just image classification. We discovered that a successful application of BMA is possible only when several underlying features provide a statistically unique contribution to the combined score. The method does not significantly improve results when any one feature dominates the score. The application of BMA diagnostics to a poorly-performing ensemble led to future research in ensemble design and verification.

We are working with subject matter experts to ensure that our advances will be applicable to cyber intelligence. In FY 2013, we added the relationship framework with the BMA classifiers and the new visual interface into the fully operational system that we deployed in FY 2012. They have a compelling success story from the application of our system to their mission.

PN12050/2451

Multimedia Information Fusion

Shawn J. Bohn

This project has the potential to allow researchers and non-researchers to increase the information content signal-to-noise ratio by viewing content in context.

We have seen an insurgence of multimedia content over the last 20 years specifically from the proliferation of digital cameras, smart phones, and video-capable tablets that can easily upload to and download from news and social media sites (e.g., YouTube™ with 100 hrs of video uploaded/min, Facebook™ with 300 million images uploaded/day, and Instagram™ with 40 million images/day). With this onslaught of multimedia information, our ability to analyze content holistically necessitates that content be fused as a means to monitor, track, and assess events. Our project focuses on identifying and exploiting contextual information in its various modes by addressing "what" defines context for various modalities; "how" context is extracted, represented, and synthesized; and "where" context can be applied as an informational aid.

The most important aspect of using context is dealing with the semantic gap. For media images, the gap is between low-level features and high-level semantic constructs and has been the target of research for years. We noticed that multimedia contextual theory has progressed to a point where it could be applied to the fusion of multimedia content. We aim to minimize the gap by associating a context with each piece of media, thereby providing a link between content, context, and the ability of ultimately fusing multiple media for holistic analysis. The gain by initiating this research is to lay a foundation that combines contextual theory with the fusion of multimedia data for analysis.

Our first task was performing an extensive literature survey within the communication field and various technical domains. We determined that only in the past 5 years has the communication community begun in earnest to quantify the elements of context in the various modalities. The catalyst for this movement was the popularity of the internet as form of communication. The technical communities are incredibly fragmented on what context means and how it applies to the problem they are solving (e.g., auto-image annotation). Several key papers exemplified what has been done in multimedia fusion and provided information about those who use context in the analysis of multimedia content.

Limiting our focus on multimedia, we concentrated on image-text fusion, having determined that there were alternative approaches to fusing content from the two modalities. The approaches could serve as a baseline as we designed, developed, and proved a context model. We used context theory to develop a hierarchy of contextual elements that was used in our model development. We realized that for every media element (e.g., image), there is content that specifies the context to various degrees that appear in different forms, whether in semantics or metadata about the media element, and needed to be weighted in our model.

To validate our approach, we engaged our subject matter expert in crafting a user study to understand what elements of text separately or independently help quantify the context for a collection of images and text articles. The user study would help us understand whether context played a role or whether the user's personal knowledge had any bearing on how the information was analyzed and grouped. To perform the user study, we obtained data to meet the study requirements. Specifically, we needed documents that included one or more embedded images and contained a variety of representations in how they referenced images, either implicitly or explicitly.

We also investigated several caption extraction tools that appeared appropriate for our needs but were too prototypical in their software state. Instead, we opted for a more simple solution of manually capturing the elements of context and storing these references as a sectioned document. Our approach was to determine whether the contextual elements were sufficient to describe images as envisioned by the author. Our results were promising, as was the initial model. Specifically, we used the hierarchy of contextual elements, weighted their contribution, and inserted the combined elements into the document space as baseline. The results told us how far our contextual representation was from the original document and whether it was better associated with other content of like semantics (i.e., whether an image was associated appropriately with the correct content).

To determine the weighting parameters of the model required, we performed the user study and used those results as training data to understand how the weighting would be performed. Unfortunately, we could only determine if the elements of context (which represented the image) had deviated from the original document. Because our research only covered three-quarters of the year, we were unable to execute our user study completely. However, we feel that many of these questions proposed by this research would address and expand our level of understanding and ability to fuse multimedia content more effectively for holistic analysis.

PN13026/2507

Multi-Resolution Data Model and Directed Data Reduction, Reconstruction and Aggregation

Kerstin Kleese van Dam

This project is delivering a collection of event and relationship descriptions and algorithms for identifying these phenomena across a multi-channel historical data archive.

The power grid requires bidirectional, real-time data flow to identify and respond to changes in demand, including quick reaction to extreme events, operational monitoring, daily and medium-term operational planning, and long-term facility planning. Unlike the current grid with a relatively small number of power sources, the future grid will have thousands of potential sources, including hybrids, wind sources, and solar panels in addition to traditional substations and power plants. Each of these sources will operate independently and thus respond differently to fluctuations in price, resources, and operational circumstances. Consumption devices will also be smarter, using real-time pricing to make decisions about their level and timing of consumption (e.g., an air conditioner will cycle on depending on not only current temperature compared to the desired but also the current price of electricity). Each of these production and consumption units as well as the connecting power infrastructure will be equipped with smart sensors that will report on status and deliver operational information and act as receptors of information and instructions (i.e., reduce demand/production).

The original focus of this project was developing a multi-resolution model as a data transport framework for the power grid that compressed identified features and prioritized data for transfer based on the importance of the data and the requirements of the receiver. The fundamental assumption underlying this work was that the amount of data generated by various sensors in the power grid (PMU data in particular) in the future would overwhelm communications systems, and this type of streaming re-prioritization of information would be required to ensure that the critical information was effectively communicated to the power engineers.

During the first year of our project, we quantified the amount of information that can realistically be expected from PMUs nationwide once the technology was fully deployed. Our best estimate was a maximum of 50k PMUs distributed across the nation when fully deployed, leading to an aggregate national bandwidth requirement of 144MB/sec. Given current communications protocols (e.g., 10GigE, 802.11n) and trends, our analysis implies that there will be no bandwidth issues in transporting the data generated by the PMUs. We published a conference paper outlining our conclusions.

During FY 2012, we focused on developing a better understanding of the generated data and translating the results into algorithms for detecting the events. We obtained access to ~1.5 years' worth of PMU data (2 TB) and identified several events of interest within this data set. Next, we performed an interactive analysis using the R statistical package running over a Hadoop cluster dynamically generated on the PNNL Institutional Computing (PIC) environment. With over 415,000 hours of PIC computing time, we characterized the underlying data by interactively testing the hypothesis against the entire data set. This was critical because many events of interest rarely happen, limiting the value of traditional techniques such as data sampling.

For the beginning of FY 2013, we spent significant time developing several algorithms to identify erroneous data elements, ranging from fairly straightforward data stream analysis (e.g., if the value remains constant for more than 10 min, there is a problem with the sensor) to more complex analysis (e.g., comparing frequency trends across PMUs to identify when a sensor is not reporting correctly). Coded in R, these algorithms were applied against our entire data set and were used to generate a "cleaned" data version in which all known PMU sensor errors were eliminated. We are using this new data to identify other events of interest. An initial version of our generator trip identification algorithm has been created, and the results of its application across the historical data set are being reviewed by domain experts.

For the rest of FY 2013, we demonstrated the feasibility of transitioning our models to a distributed platform with limited computational resources by working with the Volttron™ team to deploy our data cleaning and generator trip detection models. We selected Volttron™ as our demonstration deployment method because it is a particularly challenging platform for data analysis. In Volttron™, a model is deployed as an agent on a device (e.g., a PMU or PDC) that has access only to information that passes through that specific device. As a result, any assumptions about access

PN11054/2376

to external data or a global perspective must be removed from the model. Further, each device has limited resources; thus, models must be modified to work on streaming data (e.g., by employing windowing techniques) and has a limited amount of information in order to execute. Finally, the models must be enhanced to work in near-real time because information ineffectively processed information will be lost. This work has been well received and was published as a book chapter, presented and published in the latest IEEE conference, and was a part of a proposal for a separately funded project. In addition, the developed analytical infrastructure was adopted by PNNL's Platform for Regional Integrated Regional Modeling and Analysis (PRIMA) for the analysis of building energy simulation results for the northeastern interconnect. Finally, the results were also submitted to an international workshop on high performance computing.

During the course of this project, we developed a repository where we can record metadata about specific events for fast retrieval through either a SQL interface or a custom API. We designed an ontology capable of representing events in general and have used it as the basis for design of the relational database schema. This repository is populated with generator trip events identified by our algorithms. In addition, we created a custom Java API that allows applications to answer specific questions about generator trip events. While the standard SQL interface provides a generic query capability, our API provides a domain-specific query capability that eliminates the need for the calling application to understand our event ontology. This API optimizes query performance by ensuring queries use the underlying indices appropriately. The infrastructure component of this project was be transitioned and integrated into a new project in and thus will not be an ongoing task for our current work.

Our initial algorithms focused on basic statistical analysis of a single data stream – that of PMU frequency data. Our final amount of work was expanding on that foundation in two important ways. First, we incorporated additional data streams (or modalities or channels). In particular, we included other PMU channels in our analysis such as phase angle and external information such as weather. We established this multi-channel data set, which is used as the basis for significantly more advanced statistical analysis.

Our first step was to increase the number of data sets to which we have access. We obtained an update for our PMU data stream that brought in an additional 2 TB of PMU data covering the past year. This data set alone provided multiple channels for our analysis, including frequency, phase angle, and voltage. While this information is derived from PMU data, the resulting information provides a very different perspective of what is happening. We have been collecting weather data from several locations across the Pacific Northwest over the last year and incorporated this information into our analysis. Additionally, we worked with AEP to obtain smart meter data from their distribution system that will provide a distinct sensor stream based on the grid's distribution network, allowing us to demonstrate the generality of our data analysis framework.

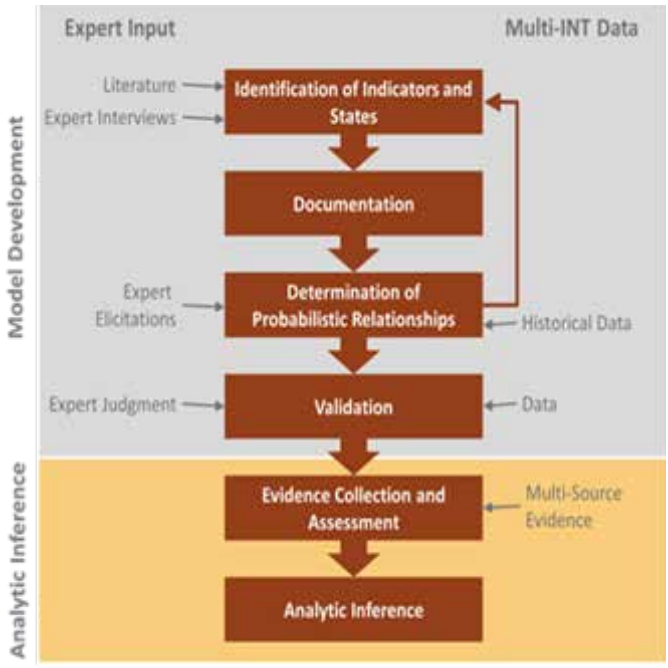
Once we developed and validated a statistical model capable of detecting a specific event (such as bad data or a generator trip), the algorithm was disseminated for use by other researchers. While the platforms for deploying analysis capabilities are being developed by other projects, we worked closely with these teams to ensure that our models have a broad reach. This close collaboration is required because adapting an existing model for deployment involves a deep understanding of both the constraints of the deployment platform and the model itself.

Multi-Source Signatures of Nuclear Programs

Paul D. Whitney

This project is developing, validating, and demonstrating contextual, model-based data signatures in a variety of nuclear-intent analytic challenges to identify nuclearization features and support analysis consistent with the informed intuition of nuclear proliferation subject matter experts.

Detecting and identifying the intent to develop or further a nuclear weapons program requires input from diverse sets of expertise, and requires understanding diverse information and data. Currently, analytic activity is executed by teams of analysts. While this typically crosses multiple disciplines, the entire set of required capabilities does not reside within any single analyst. This project addresses the heterogeneities in data and expertise to develop normalized, context-informed data summaries used to assess nuclear intent. These issues are addressed by representing analytic expertise and known relationships as influence diagrams or other modeling formalisms. The model-based approach codifies expertise and makes explicit various process models. Through this project, we are demonstrating the signatures in the context of nuclear non-proliferation models and data. Our goals for FY 2013 were to develop and demonstrate models of the political and physical processes associated with nuclear proliferation. We leveraged expertise available from the Hanford Site to support aspects of this modeling.



Model lifecycle, connections to expert input, and multi-source data

Modeling framework. In FY 2013, this project developed models related both with facilities and in the political context of facilities. The project also updated the methodology for eliciting the parameters for the models from subject matter experts. The models formed the basis of signature developments for facilities and assessing the political status of countries within which the facilities are located. The figure shows the model lifecycle from model construction to model use. Given a collection of expert-based mathematical models for evaluating features of nuclear facilities and a collection of possibly relevant data objects, the expert-based models are evaluated against the content of each data object. The model outputs are comparable across all data objects, thus normalizing different data types for comparison. Model outputs that are unusually distributed compared with the typical distribution across a row of features are selected as "topical" and become the basis for features. Because the models carry domain context, the resulting signatures based on those features can be mapped. These numeric signatures can then be used for clustering, outlier detection, or comparing facilities.

Analytic examples. The goal during FY 2013 was to facilitate the development of expert-based models by assembling an example with data. Given the unique co-location of PNNL with respect to the Hanford Site and its associated expertise, we developed a detailed model related with nuclear materials production. The model was calibrated via elicitation from Hanford Site veterans and used to motivate the construction of new signatures.

Expert-based models of facility context. The final goal for FY 2013 was developing expert-based mathematical (specifically Bayesian network) models for facility context related with the safeguard status of the host countries and facility security. The former focused on political and social factors relevant to a state's decision to seek nuclear technology, and the latter on aspects of physical security.

Elicitation methodology. During the elicitation phase of the project, we learned that our subject matter experts would prefer a different form of question to answer. Previously, we had been asking for a relative likelihood assessment of a pair of scenarios generated from the model. We altered the methodology by increasing the specificity of the generated scenarios and asked which scenario was more likely to lead to an outcome of interest.

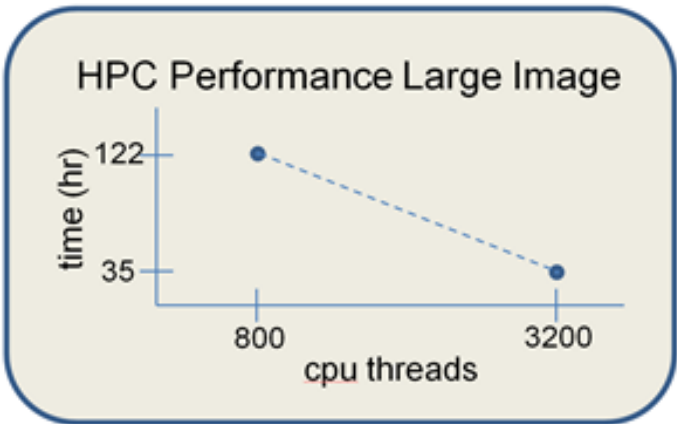
PN12053/2454

Nanoscale-Macroscale Three-Dimensional Integration Using High Performance Computing

Andrew P. Kuprat

This work establishes the capability to generate 3D datasets of macroscale data at nanoscale resolution by collecting multiple serial 2D images at nanoscale resolution and developing the high performance computing software to reconstruct the data. The capability to associate nanoscale events rapidly into the context of the macroscale is essential to achieving greater innovation in multiple disciplines.

Systems biology assumes that a life form is specified by interactions between genes, macromolecules, and cells, which do not act alone. In fact, systems biology presupposes that the spatial organization of specialized cells working closely with their neighbors determines the functional and structural attributes of an organ. Technological advances in the tools used in systems biology have led to an explosion of information that has changed our approach to the study of disease. This includes the complete sequencing of genomes for many different organisms; the development of sophisticated tools for generating genomic, proteomic, and metabolomic data; and the development of interactome network maps. A significant limitation of these techniques is that they present a static snapshot of cellular events with no spatial information. The capability to associate nanoscale events rapidly into the context of the macroscale is essential to achieving greater innovation in research.



Processor count scalability on the Cray XMT using parallelized image registration code.

This research brings genes, macromolecules, and cells imaged at sub-micron spatial resolution into a 3D system context. Using high resolution histological techniques, a small organ, organism, or bacterial population can be fixed, serial sectioned, treated with specific cellular and/or sub-cellular identifiers, and digitally imaged at 230 nm. These images can be computationally stacked into a volume representation at subcellular resolution.

Development of this capability is part of the emerging field of bioimage informatics. The primary impediment to recreating a volumetric image of histological data at subcellular resolution is that the images are extremely large. Each image can be up to 100 GB in size, and an entire volumetric dataset can occupy over a petabyte of data. Fortunately, 3D reconstruction of deformed tissue sections is a highly parallelizable task that can make efficient use of high performance computing resources.

In collaboration with the University of Washington, we commenced production of serial sectioned multicellular pulmonary datasets. Each image in these test sets is approximately 80000 × 80000 pixels, which is equivalent to the image size produced by a 6400 megapixel camera. These datasets represent different embedding conditions, various sectioning distances, and two different organisms: a mouse and a rat. By acquiring datasets under these different variables, the effects of these variables on computational performance can be examined. Then, optimal experimental conditions can be determined for future data collection efforts.

An image registration algorithm based on optical flow calculations was compiled and tested on the Cray XMT supercomputer for a pair of small test images to ensure proper functionality of the registration. During this time, the source image successfully deformed to match the target image. Next, the registration function code was re-written to be suitable for parallelization and Cray-specific parallel command-calls were added. Good scalability on a large number of threads was observed; increasing the number of central processor unit threads to 3200 from 800 (4:1) decreased calculation time by 71% (3.5:1) for registration of large images. Additional optimization efforts continued to improve performance.

Warp-smoothing is an essential step during 3D reconstruction to retain the proper overall shape of the sample. To begin implementation of this process, an initial 3D

PN11056/2378

reconstruction of a low-resolution version of a test dataset was created by applying an initial iteration of the warp smoothing approach. In addition, we tested the application of calculated elastic registration warp grids to annotated locations within the test dataset. This arrangement will allow features detected in the unaltered data to be properly translocated into the reconstructed dataset.

In FY 2012, we applied the registration and warp-smoothing techniques developed in FY 2011 to the high resolution mouse pulmonary dataset (also acquired in FY 2011). Results suggest that the pulmonary data are more challenging than the hepatic test dataset for image registration performed at the highest resolutions. This is likely due to the sparse nature of pulmonary tissue and the significant variation in shapes in sections 100 µm apart. This additionally suggests improved quality can be achieved via decreasing the spacing between acquired tissue sections.

In FY 2013, this project tested a highly parallelized approach for calculating and applying 2D rigid and elastic registrations via graphical processing unit (GPU) hardware through the use of the reg_f3d function contained within NiftyReg software package, developed at the University College London. Over previous approaches, the reg_f3d implementation allowed higher resolution input images, greatly improving quality of registration, and significantly reduced computation time.

Single image resolution increased from 1024 × 987 to 2000 × 1712, effectively doubling dataset fidelity. The CPU-bound equivalent of reg_f3d consumed approximately 300 sec per warp, with the number of simultaneous registrations equaling the number of cores on the CPU. On the other hand, the GPU version processed each registration in approximately 15 sec. The number of simultaneous registrations allowed was proportional to GPU memory, where each task occupied 340 MB of memory. With our test setup and dataset of 239 images, CPU-bound registration was carried out on 12 logical cores, resulting in a run time of 298 min. The GPU configuration involved 6 GB of GPU memory (17 parallel tasks), resulting in a run time of 10 min.

Further, efficient and accurate warp smoothing algorithms that minimize reliance on iterative calculation of warps from images were implemented. The norm was calculated for all registrations within the warp kernel for an image n, and these norms were subsequently averaged. Outliers were defined as any transformation with a norm greater than product of the threshold multiplier and the average of the norms within the kernel. Those in excess of this value were removed. The remaining images were weighted based on a discrete Gaussian distribution with controllable sigma, allowing for fine-tuning of distal vs. proximal influence.

Collectively, the tasks completed in this project successfully established the capability for 3D reconstruction of macroscale specimens at nanoscale resolution using high performance computing resources. This establishes the capability for a variety of impactful applications which require linking very large multidimensional data across scales.

PN11056/2378

Network Analysis and Modeling of Illicit Nuclear Trafficking

George Muller

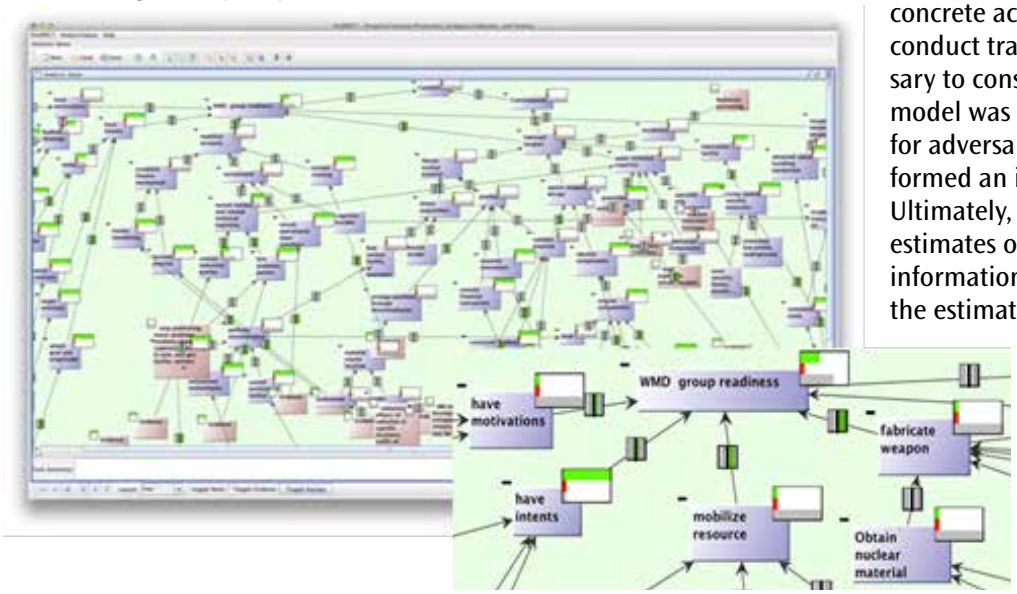
This work decomposes the problem of understanding how illicit nuclear trafficking can occur and identifies ways in which the interagency can ascertain the degree to which non-state groups are engaged in these activities.

The threat of illicit nuclear trafficking is an issue faced by every nation, and the mission to counter this threat spans international and domestic agencies. Current international and domestic programs exist to secure nuclear material at its source as well as detect and support interdiction of illicit materials in transit. Illicit nuclear trafficking networks can lead to nuclear proliferation as state or non-state actors expand their ability to identify and acquire expertise, technologies, components, and materials related to nuclear weapons. The ability to characterize and anticipate the key nodes, transit routes, and exchange mechanisms associated with and utilized by these networks is essential to influence, interdict, and disrupt the functions and business processes of these networks.

Complex paths hide the networks or individual perpetrator's identity and make detecting an instance of an illicit trafficking activity very difficult. The class of instances of this kind of activity forms a pathology of transactions that individually look normal but when linked together indicate evidence of malicious activity. In this way, illicit trafficking activity pathologies are easily hidden amid many millions of normal transaction records, and it may take months just to assemble the data into huge centralized databases where analysts can search for them.

This project builds on ongoing work that explores how knowledge about the structure and function of illicit trafficking routes and networks in analogous domains provides a framework for analyzing and understanding network operations in the more elusive and less documented nuclear smuggling domain. We will advance the science behind graph pathology detection and trace-back. Biologically inspired, agent-based approaches such have not been extended to be used in data analytics problems such as graph pathology mining and detection. Thus, one of the ways this project will advance the state-of-the-art is by using a decentralized agent-driven framework to solve the graph pathology detection problem. This approach will also be more impactful in cases where there is lack of cooperation or incomplete data. The novelty is also in the application of such an approach to the illicit nuclear trafficking domain.

During FY 2013, we developed a conceptual model for illicit nuclear trafficking to decompose the problem into concrete activities that could be performed to conduct trafficking and to enable an adversary to construct an improvised weapon. The model was used to construct three scenarios for adversary activities, which collectively formed an initial set of task sequences. Ultimately, the model would support estimates of future activities and direct information-gathering resources to improve the estimates and support the disruption of adversary activities. In addition, an inference model was developed based upon the activities defined in the conceptual model. This inference model uses evidence-based reasoning (e.g., the Dempster-Shafer decision theory) to connect hypotheses with evidence to assess group "readiness." This information is then fused through hypotheses weights



INT group readiness is evaluated using a belief network formulated from an evidence-based reasoning model. Group readiness is influenced by motivation, intent, resource mobilization and weapon fabrication. Each of these hypotheses is influenced by lower-level hypotheses, driven by evidence. The result represents the degree of progress that a group has made toward an end use goal.

PN13012/2493

and confidence of existing evidence. In addition to the conceptual model, an object model was developed to support future analysis efforts of illicit nuclear trafficking. This object model specifies the individual entities and relationships that are meaningful for the illicit nuclear trafficking problem.

Existing work on non-state illicit nuclear trafficking is limited. Historical incidents are scarce, and the information surrounding these incidents is equally uncommon. This problem requires a systems approach, which frames our entire approach to the problem. We have embarked on this method to integrate direct and indirect information about the problem to estimate the degree of activity and identify sources of additional information that can be used to strengthen these activity estimates in the future.

We will prioritize our efforts for FY 2014 among the following tasks. First, we will formalize system requirements and develop detailed scenario activity networks. After validating and verify conceptual and object models, we will integrate inference model with data sources. Finally, we will develop a COTS software plugin using the object model. These tasks enable our project to provide useful insight into the problem and allow us to demonstrate the approach for future clients.

Next Generation Network Simulations for Power System Applications

Jeff Daily

We are developing a next-generation modeling tool that captures the interaction of communication technologies with the power grid. This enables power systems engineers to understand and address the challenges of communication quality of service on the control, monitoring, and operations of future power systems.

New smart grid technologies and concepts such as dynamic pricing, demand response, dynamic state estimation, or wide area monitoring, protection, and control are expected to require considerable communication resources. As the cost of retrofit is too high, future power grids will require the integration of high-speed, secure connections with legacy communication systems while still providing adequate system control and security. Considerable work has been performed previously to co-simulate the power domain with load models and market operations; however, limited work has been performed in integrating communications directly into a power domain solver.

We are developing a scalable, high performance simulation infrastructure for data communication networks for power transmission and distribution power grids. Our goal is to develop a unique infrastructure that will position PNNL as the leader in modeling and planning power grid communication networks. This project will yield a general scalable high performance computing (HPC) communication network simulation, modeling, and planning infrastructure. We will investigate bottlenecks and limitations in state-of-the-art simulation infrastructures, proposing new approaches, parallelization methods, and high level algorithms to enhance performance. We will map and optimize existing simulations to state-of-the-art HPC computers, proposing new engines for integration with current simulation infrastructures. We will also interface existing power grid simulation infrastructures with the new proposed network simulation, identifying network properties and requirements representative of the power grid network domain.

For FY 2013, we evaluated the feasibility of continuing the development of our two approaches: rewriting the entire software stack targeting HPC systems or creating a middleware for connecting robust stand-alone simulators (e.g., GridLAB-D, ns-3). We focused our efforts on the latter because it would provide the greatest impact in terms of time to solution. The two main requirements for integrating GridLAB-D and ns-3 (and other simulators) are reuse and separation. The goal is to make the modules of both simulators available for integrated simulations and, where necessary, allow the different simulators to pass information into or out of one another. In this way, users can implement and experiment with different power and communication hardware/protocol combinations.

It is possible to integrate the simulators by adapting the modules of one simulator for use with the other one. Unfortunately, such adaptations can complicate the code base of the simulators, introduce errors, and often recreates work than has already been done. Hence, it is important to keep the code bases of both simulators separate, allowing for independent but parallel development. For example, modules of ns-3 could be adapted to work with GridLAB-D by directly linking the ns-3 functionality through the core operations of GridLAB-D. However, this complicates the code base of GridLAB-D, as developers also have to maintain these modules. Moreover, the adaptations can introduce errors to the existing code leading to crashes or incorrect results. Considering the requirements, integration middleware was implemented, within which both simulators run independently and communicate only when they need to exchange messages. The middleware facilitates the message exchange and time synchronization.

The time synchronization module keeps the internal times of both simulators in sync. This is a critical operation in order to coordinate the message exchanges. For example, a message sent by a GridLAB-D module at time *t* should also be forwarded to the communication network at time *t*. If

Integrating communications into a domain solver for the simulation of communication and power systems to form an interrelationship. Communication system latency, cyber security, and topology all will affect the power system itself and the controls to operate the power system.

PN12054/2455

ns-3's internal time is less than t, then the message can be forwarded back to GridLAB-D at $t_1 < t$. This may result in errors during calculations.

It is possible for the communication network to "lose" packets (e.g., by simulating network errors). In such a case, the counters for sent and receive messages would never be equal. Therefore, the synchronization algorithm described above would force the simulators to progress one step until completion, severely impacting the performance. To prevent this, the time synchronization module was programmed to declare packets as lost when the sent and received counters do not stabilize after n time steps after it is sent. The user can set the value for n.

Having started with the Message Passing Interface (MPI) glue code in FY 2012, we continued development using the highly flexible and scalable ZeroMQ messaging library. The final result is the Framework for Co-Simulation (Fenix).

Fenix is an interface, library, and middleware that facilitates the time synchronization and exchange of messages between multiple simulation software instances. Fenix was developed to integrate GridLAB-D, ns-3, and the MCA transmission solver, but it is generic enough to use with other simulation software. It follows many of the practices developed by High Level Architecture but is specifically tuned for the power domain.

The Fenix software has since become a flagship product. Both internal and external clients are asking for information on how to integrate their software systems with GridLAB-D using Fenix. Fenix was presented at a cyber systems workshop this year with great response from the audience. The Fenix software fills the niche of getting many different power simulators to work together to solve the next generation of power system problems.

Real-time High-Performance Computing Infrastructure for Next-Generation Power Grid Analysis

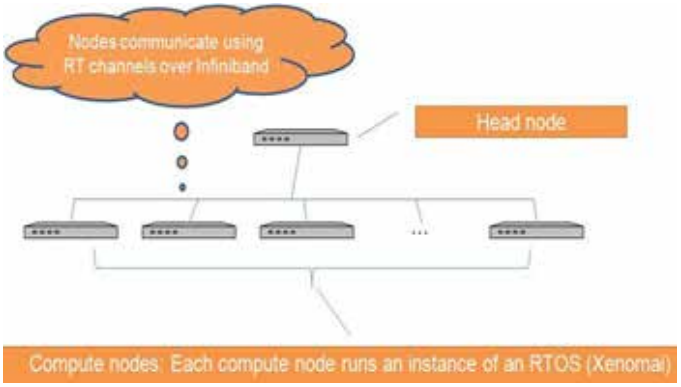
Peter S.Y. Hui

We are developing a platform for real-time computation in a high-performance computing (HPC) environment with the context of supporting real-time future power grid analysis. Specifically, we are creating a real-time operating system for HPC machines along with real-time libraries to support key power grid analysis kernels.

Traditionally, real-time computing has been considered in the context of single-processor and embedded systems. Indeed, the terms real-time computing, embedded systems, and control systems are mentioned in closely related contexts. However, real-time computing in the context of multi-node systems (specifically high-performance, cluster-computing systems) remains relatively unexplored, largely because until recently, a need has not existed for such an environment. With the future power grid growing significantly, HPC platforms are integral to analyzing the grid. This project investigates infrastructure necessary to support cluster computation over large datasets in real-time. Our motivating example is an analytical framework to support the next generation North American power grid, which is growing in size and complexity. With streaming sensor data in the grid potentially reaching rates on the order of terabytes per day, analyzing this data subject to real-time guarantees is daunting and will require an HPC capable of functioning under real-time constraints.

We spent FY 2011 building the infrastructure foundation. We scoped the requirements for a small-cluster computer and purchased one Infiniband switch and one Ethernet switch. We developed a traffic prioritization for the Infiniband hardware interconnect that would enable higher priority traffic to be transmitted with higher precedence than low-priority traffic. For a real-time operating system (RTOS) to be deployed on each of the compute nodes, we evaluated several candidate operating systems, which led us to choose Xenomai with its support for hard real-time deadlines, its simple and clean user friendly API, and its wide and active user community. Finally, we prepared to integrate the software stack and sample power grid kernels into the RTOS real-time kernel.

For FY 2012, our focus was developing a real-time version of a power grid analysis routine, specifically an application for dynamically estimating the electromechanical states of the power grid. Briefly, the chosen application targets one of the goals in simulation-based modeling of a power grid network



High-level overview of the targeted infrastructure. Compute nodes run a local copy of an existing real-time operating system (e.g., Xenomai), while data are moved between nodes via a real-time Infiniband layer that was developed as part of this project.

to assess conditions, such as network stability at a given time. The application models the network using a system of mathematical equations for network states and then solves these equations using an ensemble Kalmann filter. To maintain the rate of data generated by the Phasor Measurement Units deployed on the power grid, the goal of the system is to compute results within 30ms.

In FY 2013, we expanded on concepts developed in the previous years, tying them together to create a prototype real-time, HPC-based power grid analysis kernel. Primarily, this effort entailed integrating the RT modules developed in FY 2012 into a dynamic state estimation application, and using formal, automata-theoretic approaches to show that the crucial execution paths operate within the required time bounds. We developed a method of formally verifying the code to run within the stated time constraints, something that has not previously been done in a parallel environment such as ours. One of the challenges was capturing complex timing interactions that occur in a parallel computing environment and guaranteeing that with all of the various possible control flow paths among the processes and with complex fork/join interactions between them, all timing requirements could still be met. We took the approach of first modeling the control flow of each of the parallel processes using timed finite automata, specifically including points at which child processes were spawned and joined. Next, we analyzed control flow automata to determine the worst case execution time of all component processes. We also modeled timing interactions between the processes.

We presented two research papers summarizing our work, one at an international workshop and the other at a regional power grid conference.

PN11065/2387

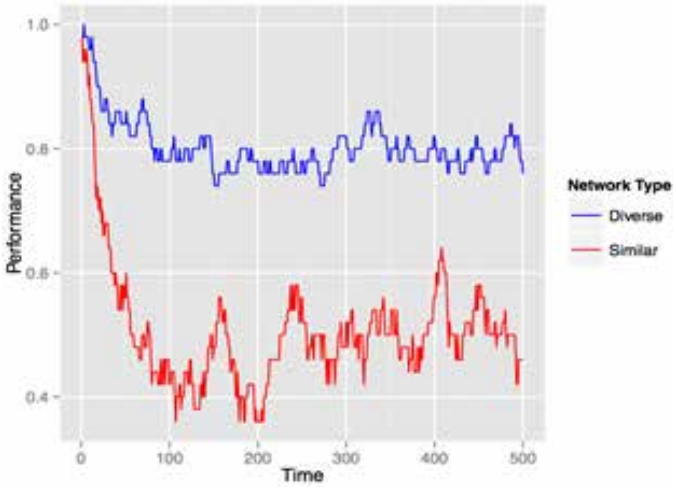
Robustness

George Muller

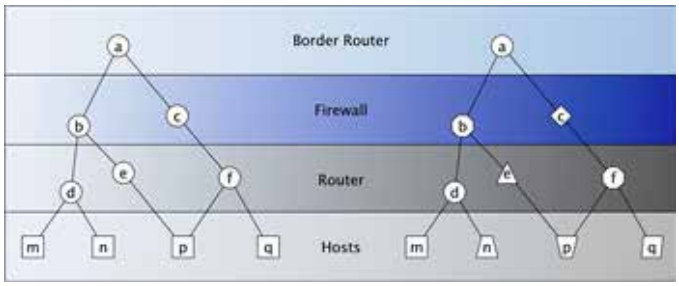
Robustness provides inherent protection against significant consequences from successful attacks through improved information system architectures.

In today's cyber security paradigm, attackers hold an asymmetric advantage for attacking networks due to a large set of vulnerable opportunities, the low cost of deploying an attack, and minimal consequences for executing an attack. This forces networks to operate in a state where compromise becomes a certainty, requiring resiliency to continue providing the functions necessary to perform the business mission. Robust networks, which reduce the impacts of disruptive events associated with cyber attacks, are a key aspect of resilient networks, and are critical to establish resilience as the default operating mode for vulnerable networks.

Currently, the approaches to comprehensive resilient cyber security strategy are not well known. Broad methods such as deception and diversity have been proposed; others, including changing a system's configuration using genetic algorithms and varying the configuration of virtual servers to create dynamic, random attack surfaces, have been explored and demonstrated some success. The expectation is that these methods afford defenders rather than attackers an asymmetric advantage to security. This is achieved in large part by adding complexity to the security layers of a network that arise for both attacker and defender, and the degree to which interfaces with current methods and approaches will be affected by these new



Simulated network performance is increased as diverse components continue to operate under heterogeneous attack types.



Network resiliency is influenced by the diversity of the equipment deployed on the network, even in the same network layout.

methodologies is not well understood. In addition, security approaches focus specifically on the response to a disruption or attack. There is limited research and even less operational implementation on proactive defense strategies, which are largely limited to moving target defense, the feasibility of which remains uncertain.

Robustness aims to modify network construction and behavior to mitigate the consequences of successful attacks. Our work explored underlying theoretical foundations upon which resilient cyber systems can be developed and operated through extensive background research. Building on this research and integrating it with the development of cyber security metrics, these theoretical foundations can be developed, documented, and bridged to tangible, operational solutions that can be executed by the cyber security community. The result will be a functionally defined architecture that describes methods available, provides a basis for each resilience attribute, and offers the set of conditions for which each resilience attribute is well suited to guard against.

The robustness project conducted basic research into resilience theory, existing approaches to resilient network systems, and metaphors that exist in the biological and infrastructure domains. The project aimed to identify network architectures –configuration and behavior of its components – that yield more robust system level behaviors. If the assumptions are that networks will continue to increase in complexity, that vulnerabilities will continue to be exploited, and that network performance is partially or wholly affected by exploits, then a key means to reduce exploit impact is to change to the structure of networked systems in such a way that system performance is not dependent on the functioning of individual elements within the system. This work lays the foundation for constructing information networks in a way that enables system-level performance in the face of disruptions to lower-level system components or sub-systems.

PN13065/2546

We developed a general approach to robustness through the application of network characteristics that contributed to robustness. Presented at an IEEE conference, our position paper highlighted existing approaches, gaps, and hurdles in implementing the robustness approach. In addition, existing approaches to resilience were researched. We established the essential properties of robust networks, including redundancy, hardness, flexibility, and diversity.

We developed an initial simulation to demonstrate network performance using configuration-dependent failures on a network. The model demonstrated an initial result of varying robustness properties of a network on overall performance based upon the configuration of the network. Because our approach was strikingly similar to another project, this project and the work conducted under it were rolled into the other. This enables a more straightforward, integrated approach to resilient architectures through the complementary methods of robustness and reconstitution. However, our initial results demonstrated that system performance is influenced through the robustness properties of redundancy and diversity.

Scalable Knowledge Extraction on Extreme Scale Scientific Data

Abhinav Vishnu

We proposed a framework for data analytics to provide tools for data analytics by scalable design and implementations of algorithms in classification, correlation analysis, clustering, and association rule mining.

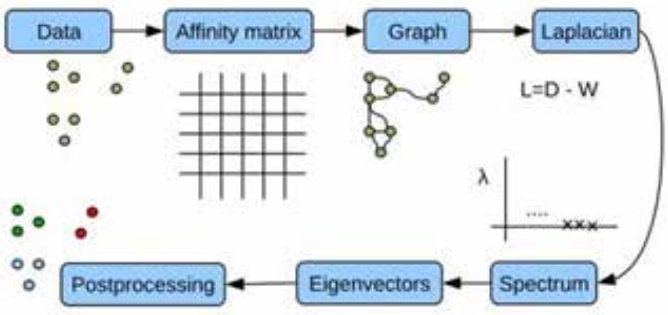
Today, scientific simulations and instruments produce exorbitant amounts of data. The primary objective of data collection is to provide modeling for predictive analysis. While there has been significant research in applying domain knowledge to data modeling, non-parametric modeling is becoming equally important to support domain modeling, especially in the cases where it is too difficult to isolate the effects of variables under consideration. In this context, data mining/machine learning – mechanisms to extract knowledge from data – is becoming extremely important. While several sequential algorithms exist to perform data analysis, they cease to be useful for the scale of data expected to be generated in the near future.

In this project, we worked to identify useful data analysis algorithms, conducting research on several of these well-known algorithms for parallelizing in distributed memory machines and providing a path forward to publishing them in the near future. Previous attempts were made to provide sequential algorithms for data analysis. Similarly, open source software is being used for data analysis. An important algorithm in data analysis is classification in which support vector machines are popular due to their high accuracy. While much of our focus is on structured data, we do not focus on graph datasets. To complete the literature study, we would add that there are several software suites that are becoming well used for graph algorithms. There have also been a few attempts to provide parallel algorithms for data analysis. The closest attempt is NU-Minebench, which provides several algorithms for clustering and classification.

The primary objective of our project was providing a library of parallel algorithms that can be used by scientists to perform data analysis at large scale. We believe that our work combines the traditional high performance computing (HPC) research and data analysis research effectively. In essence, the objective is to perform co-design of data analysis algorithms on HPC machines that have been traditionally used for applications in various science domains. As a result, we intend to use as much of the infrastructure in hardware and software, which has been used in traditional HPC infrastructure, such as Message Passing Interface and Global Arrays. The expected outcome of the project is potentially an alpha software release of the research conducted on algorithms in clustering, classification, and association rule mining algorithms.

Simplicity of implementation, efficiency through standard linear algebra packages, and improved performance over traditional algorithms in the real-world perceptual grouping problems have resulted in a broad adoption of the spectral clustering algorithm. Though this algorithm is widely used, we find its presence predominantly in sequential DM packages and only a few in parallel, making its way into our package. It is worth noting that in addition to being used as a data clustering algorithm, spectral clustering is used for segmenting images, an important step in image processing. To the best of our knowledge, this project is the first effort to study the performance of high performance spectral clustering on the chemical imaging domain.

We provided a scalable implementation of K-means algorithm, which we have tested on the genome wide assembly dataset. Our algorithm provides almost linear scalability with weak scaling. This step is important, as K-means is used for early analysis of algorithms. In addition, we designed a scalable algorithm using sequential minimal optimization (SMO), frequently used for classification. While the traditional SMO algorithm is sequential and a parallel algorithm was proposed, there was redundant work. We designed and implemented the algorithm to provide smart shrinking of the dataset and load balancing so that all processes have an equal amount of work. We are currently evaluating the scalability of this approach, and the initial results look encouraging. Finally, we designed a parallel FP-Growth version and proposed a series of optimizations in memory utilization, merging, and load balancing, which provide much improvement in performance compared with the state-of-the-art. The importance of mining-extracting useful information from data is of utmost importance for domains such as national security, high energy physics, behavioral modeling, and medicine.



Steps in parallel spectral clustering

PN13066/2547

Scalable Sensor Data Management Middleware

Jian Yin

This project is motivated by the explosive growth of power grid system sensors and an increasing number of innovative analysis applications that can use sensor data to improve operation efficiency and increase power system reliability.

Smart grid technology promises to enhance and automate the monitoring and control of electrical distribution and is essential to integrating variable renewable energy sources that will reduce carbon emissions and enhance energy independence. To fulfill the lofty potential of smart grid, an unprecedented amount of data must be collected to understand the human and device behaviors that enable efficient distribution and consumption of electricity. However, these data are not useful unless they can be made available for analysis to understand control decisions and turn them into information and knowledge. An integrated data layer to mediate the massive amount of data and the large number of applications is the only viable option to provide efficient access to a large number of applications.

In this project, we built scalable, flexible data management middleware to collect, select, store, replicate, preprocess, and deliver the data generated from a large number of sensors deployed in smart grid systems. Our middleware provides a general set of abstractions to meet the data access needs of a large, diverse set of analysis and control applications. We gathered design requirements of real-time data middleware for future power grid systems to handle the explosive growth of sensors and control and analysis applications that use these sensors to improve the efficiency and reliability of power grids.

Our system is designed for scalability and predictability to meet real-time requirements, and storage is optimized for rapid insertion and deletion of high-speed data from a large number of sensors. Specifically, we designed a specialized log structure-inspired schema to store power grid sensor data written directly to the file system block device layer using interfaces, eliminating high overheads and unpredictability associated with higher level interfaces such as file systems, in which one access can lead to multiple accesses for reading the inode block, varied indirect blocks, or the data block itself. We also designed specialized data structures to monitor data and free space in the storage devices. By leveraging the fact that data are often inserted in temporal order, we minimized metadata so that it could be kept easily in main memory.

In our system, a data insertion is translated into only one disk access that allows customized indexes for data retrieval. Our design can aggregate local and remote main memory and SSD to enable most metadata lookups to happen in the types of memory that support random access. We can also exploit remote direct memory access to utilize remote RAM and SSD efficiently. Next, we prototyped major features of our real-time middleware system. Note that our system leverages Linux buffer caches on computer disk devices to achieve performance. In preliminary tests, our system achieved several orders of magnitude performance improvements over traditional systems. Performance is predictable and consistent, which is crucial for real-time data ingestion. Our results produced several papers, presentations describing the design and initial implementation of our real-time data middleware system, and client funding for managing power grid data.

In FY 2013, we built on our previous work to continue improving performance, reliability, and usability of our system. We built the mechanism to allow our system to scale over multiple machines. Our system is implemented as close as possible to hardware to obtain the maximum possible efficiency. Preliminary results showed that our system significantly outperformed similar commercial systems and provided consistent, predictable performance, which is important to meet real-time guarantees. Experimental results show that our system performs close to the hardware limit.

We also developed mechanisms to improve the reliability of our systems. This capability allows our system to continue to provide service in the face of partial failures such as computer crashes and includes mechanisms to detect failures and reconfigure the system to use other machines for services. For this reason, we engineered our system to be deployed easily over a wide variety of commodity hardware. An added feature is that we built in the capability to allow multiple language support, as our efficient implementation minimizes the overhead for multiple languages. We demonstrated that other parts of larger systems could interface with our system in program languages including C and java. In our system, we also allow computation to overlap with communication, hence minimizing latency.

In FY 2014, we will find opportunities to apply our technologies as we explore internal and external funding opportunities. Our capability can be applied to research and development as well as operational environments. The amount of data in power grid systems continues to grow explosively, and the future power grid systems are becoming increasingly dynamic to improve efficiency and reliability. As the need to collect, select, store, replicate, preprocess, deliver, and use data continues to grow, we believe that there could be many opportunities. We anticipate work that will allow us to apply the capability developed in this project to different parts of power grid systems.

PN11066/2388

Scalable Solvers for Uncertainty Quantification of Large-scale Stochastic Partial Differential Systems

Guang Lin

This research is focused on developing fast iterative solvers for the numerical solutions of stochastic partial differential equations (SPDEs).

SPDEs provide mathematical models for the uncertainty quantification in numerous physical and engineering applications, including flows in heterogeneous porous media, thermo-fluid processes, flow-structure interactions, and similar situations. These models are proposed to quantify the propagation of uncertainties in the input data such as diffusion coefficients, forcing terms, boundary conditions, initial conditions, and geometry of the domain to the response quantities of interests. The main challenge for the numerical solutions of SPDEs is the phenomenon known as the "curse of dimensionality" for the discretizations in the stochastic domain.

The objective of this research is to develop fast, scalable iterative solvers for the large linear systems resulting from the stochastic Galerkin discretizations of SPDEs. The size of such linear systems increases dramatically when increasing both stochastic and spatial accuracy, hence the demand for efficient solvers. On the other hand, the matrices associated with the Galerkin discretizations have rich structures such as block sparsity pattern, hierarchical structure, and tensor product structure. We are designing new solvers that utilize these special matrix structures with the goal of combining multilevel techniques in stochastic and spatial domain to achieve solver efficiency. More precisely, we are using successively finer grids in space and hierarchical polynomial basis in the stochastic domain to build efficient preconditioners for the Krylov subspace methods such as the conjugate gradient (CG), biconjugate gradient stabilized (BiCGSTAB), generalized preconditioned conjugate gradient (GPCG), and generalized minimal residual (GMR). The performance of the new solvers is tested and compared with traditional solvers based on block diagonal preconditioners.

At the commencement of our work in FY 2013, we used the finite element method to discretize spatial dimension of the underline SPDEs and generalized polynomial chaos (gPC) expansion for the discretization in stochastic dimension. In particular for elliptic type SPDEs, we used the standard linear finite element; for stochastic Stokes equations, we used the H(div) finite element. When the random variables were uniformly distributed, we employed Legendre polynomials to span the stochastic approximation space. Additionally, hermite polynomials were chosen in case random variables satisfied normal distribution. The global approximation space was the tensor product of the finite element as well as the gPC spaces. The resulting linear systems of equations inherited this tensor structure, which is important for the design of scalable solvers.

Continuing our efforts, we combined the optimality of the multigrid method in physical space with the hierarchical structure in probability space and obtained new preconditioners for the Krylov type iterative methods. Specifically, we proposed a block triangular preconditioner for methods, including GMR, GPCG, and BiCGSTAB. We also considered a symmetric block preconditioner for the standard CG method. Our numerical results indicated that the block triangular preconditioner was more efficient than the traditional block-diagonal preconditioner for SPDE problems with large random fluctuations. We also analyzed the theoretical bounds for the spectrum of the preconditioned system.

The proposed preconditioners were tested on elliptic equations and Stokes equations with random coefficients, the latter of which were assumed to satisfy either uniform distributions or log-normal distributions. For stochastic Stokes equations, we proved theoretically the existence and uniqueness of the solution of the continuous problem as well as the corresponding discrete problem. Further, we proved that the solution of the discrete problem could achieve the optimal first-order convergence in the spatial approximation space and the optimal order in the stochastic approximation space on mildly structured meshes. Our numerical experiments verified these theoretical results.

We also developed multigrid algorithms for stochastic Stokes equations using block Jacobi and block-Gauss-Seidel smoothers and obtained optimal convergence rate.

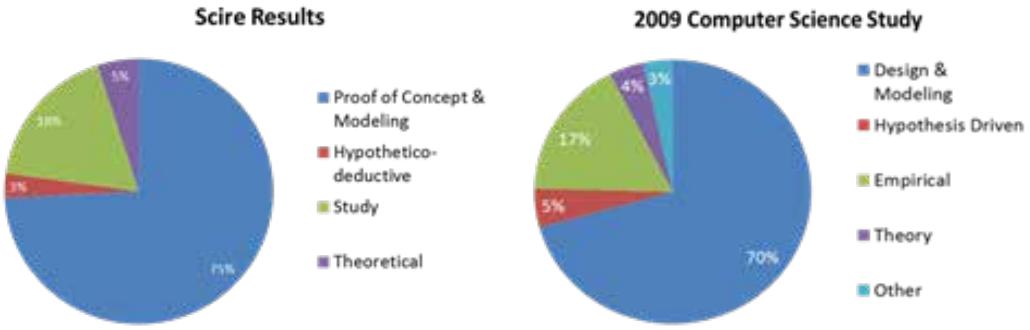
PN13003/2484

Scire: Scientific Process for Validation and Verification

Thomas W. Edgar

The aim of this project is to develop and exercise a methodology for verification and validation when performing modeling and simulation, experimentation and studies, and theoretical research.

Older, more established fields have generated foundational laws and theories upon which to base their engineered systems. Cyber security still lacks the necessary fundamental understanding of cyber space to engineer effective solutions. Cyber security needs to move beyond tradecraft and building interesting systems to include scientifically rigorous knowledge gathering and engineering based upon science. For example, before chemistry became scientific, people believed alchemy was possible. Without instituting a scientific approach, the status quo will remain; we will continue to be unable to differentiate fear, uncertainty, and doubt derived mysticism from real solutions; and we will slide further down the asymmetric slope of the attacker advantage.



Research empirical comparison with computer science

To address the scientific rigor of cyber security, we are developing a scientific guidance document that researchers will leverage to develop their scientific questions and design a rigorous research approach. The document will address all of the scientific issues necessary to have valid and impactful results. An extensive literature survey from the field was performed to capture and integrate the strong methods and measurements that were already derived. In addition, we are defining scientific approaches to fill the necessary gaps. In addition, a multidisciplinary scientific committee is being formed to leverage this guide document to work with other projects to define and design scientific validation plans and to collect feedback and improve the guide document.

The first task performed under this project was an extensive literature study to understand challenges of the field and methodologies used to perform research. This process collected, sorted, and analyzed ~ 5700 documents. Empirical analysis of the large corpus of data was performed to understand what types of research is being performed and by which publishers. From this effort, we learned that the cyber security field lacks a strong fundamental science effort and has not coalesced around standard rigorous approaches to practice research.

The second step was performing an empirical comparison to research in computer science. Similar studies have been performed in the broader general computer science research. Because cyber security in effect grew out of computing science, we wanted to discover if we were better or worse than our parent field. The studies we compared criticized the limited amount of fundamental research being performed, and we found the cyber security field to perform even less.

From this understanding, we decided that the most impactful method for improving the validity of research would be to develop a cyber security scientific method guide document to walk researchers through a rigorous scientific process and to explain the importance and dependence of fundamental research results to applied research efforts. An initial draft of this document was created during this fiscal year with the intent to edit and revise it with future efforts.

Finally, we leveraged the capabilities of a summer intern to trial ideas in the guide document. We do not assume to have all the answers, so we felt that it was important to have an independent person attempt to use our ideas to discover gaps in our methodology and limitations in our approach. Following our process, the intern drafted an experimental design for a pilot study evaluating the determinism with SCADA communication, a widely held and yet untested assumption. The feedback from the intern throughout the process helped us to refine the message in our first draft of the document.

The plan for next year's efforts is to create a scientific committee that will work with other projects to evaluate the effectiveness of our document while also providing the service of helping improve the rigor of several other similar projects. In addition, we will evaluate how to improve the integration and control of user and adversarial behavior within experimentation – two scientific challenges that were discovered during this year's research – for inclusion in the methodologies of the guide document.

PN13057/2538

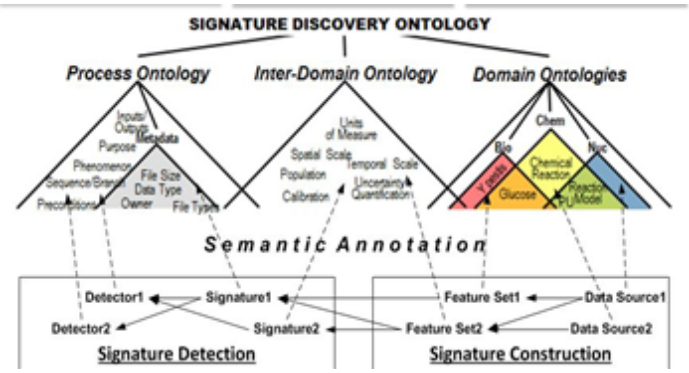
Semantic Workflows for Signature Discovery

Cliff A. Joslyn

We seek to generalize signature methods across domains by providing analytic measures of semantic similarity of ontologically annotated workflow components within the Analytic Framework (AF).

Equipping workflow components and data items with semantic specifications representing the type and content of information flows allows measurement of similarities and recommendations about interoperable methods and algorithms. To this end, this project develops methods, measures, and algorithms to guide scientists and researchers to automate construction of signature discovery workflows and identify new composite signatures. A workflow to serve one domain (bioforensics or biomarker discovery) may have individual analytic components and full or portions of analytic workflows useful for another domain (e.g., chemical forensics or insider threat discovery). An annotation process can register workflow components (both data objects and software capabilities represented as I/O ports of workflow components) with their semantic content. Given a library of such annotated components and data objects, we have developed, deployed, and validated novel mathematical measures which can accurately recommend matchings appropriate on a semantic basis.

Focus selection. Developing semantic technologies targeting a broad area of applicability like signature science requires a phased approach. To support tractable development, we identified a focused, domain-specific application sufficient to demonstrate applicability and feasibility. We sought to instantiate and demonstrate the value of semantic technologies by selecting the domain of cyber-sequence analysis, interacting closely with the MLSTONES project. In addition to integrating with other AF demos, MLSTONES provides coherence by focusing on discrete sequential data originally from the cyber domain but with generalizability to biological analysis.



Workflow and ontology relationships and processes

Semantic services for AF middleware. We provided the ability to equip software services as AF workflow components with semantic metadata, including semantic annotations of component ports and data items in the AF library through SAWSDL protocol. User roles (component developer, component annotator, and workflow creator) interact with the AF library and a knowledgebase (semantic triplestore) to register the semantics of component ports and measure similarity. This allows Boolean and quantitative type-matching at workflow construction and execution times.

Semantic content development for signature discovery. Identifying, integrating, and developing needed semantic resources requires lexical (terminological) and ontological (conceptual) development. We developed the sequence analysis ontology (SAI) and knowledgebase (SA-KB) by combining novel ontological development with integration with existing ontologies, including our own signature discovery ontology (SDO).

Comparing semantic annotations. Quantitative annotation comparison requires development, implementation, and testing of analytical methods to measure similarities of workflows, components, and data items in the context of user cases for the recommendation of data and component generalization and selection. We advanced our mathematical methods for metrics in ordered sets for measuring distances between concepts and sizes of regions circumscribed by groups of concepts in semantic hierarchies. These stand in contrast with industry standard measures based on so-called semantic similarities (SSes) but are rarely true metrics.

Both our new poset and traditional SS-based distances can be deployed within an aggregating Hausdorff measure to measuring overlaps between multiple groups of concepts. We developed algorithms to identify all the semantic classes implicated by individuals in the SA-KB, representing annotated data objects and workflow components. We then manually identified rank-ordered recommendations to provide ground truth and verified a high level of recommendation quality (Spearman rank-ordered correlations of $> 80\%$) for poset-based distances but a rather low (ranging from -20% to 40%) correlations for SS-based distances.

Future work includes deployment of these methods not already in workflow recommendation but managing novel protein annotations for another project.

PN12070/2471

Signature Discovery Analytic Framework

Adam S. Wynne

This project creates a software architecture that integrates novel and existing data analysis algorithms into a single, enterprise-strength system that allows users to take advantage of PNNL's methodology of systematic signature discovery across multiple domains.

Next-generation signature detection capabilities require a signature development platform that provides advanced tools and services to discover, verify, and exploit complex threat signatures in multiple heterogeneous data sources. Software tools and infrastructures for the discovery of domain-specific signatures are in use and well known, but none has the flexibility and reusable methodology required to be extended to multiple domains. These tools are typically applied ad-hoc for the task at hand or are integrated using workflow software, which provides a graphical interface to integrate disparate software tools. In either case, methodology is not made explicit in that it is tied to a particular scientific domain or not documented systematically. This makes it difficult for signature development tasks to be repeated and applied in other contexts. Although the workflow tools help to automate and integrate distributed computing resources, they are often challenging to apply in practice and cannot individually take advantage of the high-performance computing resources required to solve many of these problems.

To build a system that enables users to employ a fundamentally new methodology, we helped develop signature discovery process (SDP) methodology. To support the execution of commonly used signature discovery methods and algorithms, we created a flexible, lightweight integration platform that can incorporate existing and legacy tools into a single system. These tools are made available by creating web services, which are accessible from a wide range of client tools. The objective is creating a novel software architecture that combines a rigorous signature development methodology (developed under another research task) with highly flexible, robust software architecture and associated tools. Thus, this project will result in a lasting software toolset that can be used to develop signatures in a wide variety of domains. This knowledge will help achieve several goals by reducing trial and error, allowing users to reuse existing tools and approaches, and providing a reproducible mechanism for constructing and evaluating signatures.

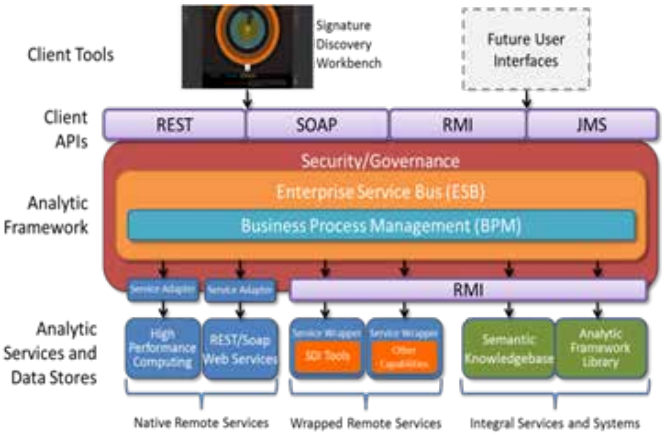


Diagram of full analytic framework

In FY 2012, we worked toward our project goals by creating a service-oriented design and comprehensive vision for the analytic framework consisting of the following high-level elements: signature discovery workbench, analytic components, integration framework, analytic framework library, and semantic knowledgebase. This year, we implemented the design and focused on making the system production-ready to support signature challenge projects in FY 2014. The resulting implementation comprises an enterprise-quality integration platform that exhibits the following qualities:

- **Modularity** – The integration framework allows analytic services to be readily added and removed.
- **Deployability** – The system can be easily deployed and fielded through the use of flexible profiles that abstract away host-specific information.
- **Interoperability** – Web services are used to make analytic tools available to a wide range of client tools and programming languages.

This system was used to integrate the analytic capabilities of five other projects and led to demonstrations for each.

By the conclusion of the year, we delivered the Analytic Framework software and comprehensive architecture documentation to be used in developing signatures for signature challenge projects. The maturity of the system is exhibited by the fact that it has been adopted as the integration framework for PNNL's power grid work as well as other projects sponsored by the Defense Threat Reduction Agency and the U.S. Department of Homeland Security. Additionally, PNNL Institutional Computing has recognized its value and is considering supporting the framework for use with future research.

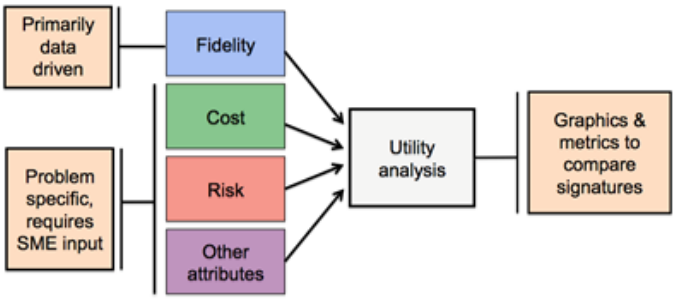
PN11067/2389

Signature Quality Metrics

Landon H. Sego

This project provides a holistic assessment of signature quality in terms of fidelity, cost, risk, and utility with the ultimate goal of identifying optimal signatures for a particular application.

Signature systems that detect, predict, or characterize a phenomenon of interest are used by scientists in virtually every field of study. Every signature discovery effort needs a transparent approach for evaluating the quality of the resultant signature system. Consequently, a domain-agnostic methodology for assessing the quality of signatures will help researchers ensure that signatures achieve their intended purposes. The need and significance of signature quality metrics (SQM) can be summed up by this well-known statement attributed to Lord Kelvin: "If you cannot measure it, you cannot improve it."



Components of the SQM methodology

SQM provides a holistic approach for assessing the quality of signature systems as they move through the stages of research and development, construction and evaluation of prototypes, and eventual deployment in an operational environment. SQM enables researchers to address the following types of questions:

- Is one signature better than another?
- Does it make sense to attempt to improve upon an existing signature by modifying that signature or even creating a new one?
- What can we do to improve the performance of a poor signature?

We assembled a team of experts with diverse backgrounds, including statistics, economics, psychology, operations research, risk analysis, computer science, and decision science. Drawing on this diversity of perspectives, we developed a systematic approach for measuring signature quality. Finally, we demonstrated the methodology in the context of various signature discovery efforts.

Specifically, SQM consists of the following four components:

Fidelity: This term examines how well a signature system detects, predicts, or characterizes the phenomenon of interest. Fidelity includes metrics such as sensitivity, specificity, and accuracy.

Cost: Cost is the resources expended to develop, deploy, and/or utilize the signature system. Specific examples include the cost of detection systems, consumable reagents, and labor.

Risk: Risk is the likelihood and consequences associated with decision errors that may result by employing the signature system. This idea helps investigators assess the tradeoff between false positives and false negatives.

Other attributes: Any other factors that may distinguish one signature system from another that are not already accounted for by fidelity, cost, or risk are included in this category. Examples cover concepts such as time-to-decision, system portability, human safety, ease of use, and the like.

SQM also includes techniques to aggregate the component metrics into a single decision framework that can be used to compare various alternatives using decision and utility theory. Ultimately, SQM provides a transparent process for assessing and visualizing the quality of nearly any signature system.

We successfully demonstrated the application of SQM in four primary areas: 1) comparing algorithms that process gamma-ray PVT sensor scans for radiation portal monitoring; 2) various assays for bioforensic analyses of *Bacillus anthracis*; 3) signature systems that use bioforensic analyses of *Yersinia pestis* and scientific literature; and 4) data quality predictions for proteomic datasets. Each demonstration included a conference or journal article as well as documented software used to implement the analysis.

We also documented the entire SQM process in a journal article. The methodology and software developed from this project will be useful to future signature discovery efforts in any domain.

PN11068/2390

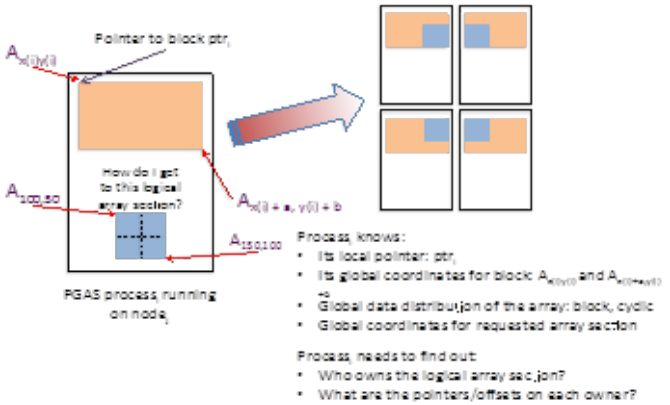
Single Node Optimizations for Extreme Scale Systems

Daniel G. Chavarría-Miranda

This project investigates the necessary design and implementation techniques to support efficient utilization of node resources for extreme scale scientific applications.

Extreme scale systems provide an unprecedented opportunity for scientists to attack larger, more complex problems and reach solutions dramatically faster. These new architectures promise to be qualitatively different from existing computers, and extracting their full capabilities requires significant innovations in programming models and the underlying primitives used to implement them. The next generation of computers will be characterized by larger numbers of nodes, each with hundreds to thousands of cores. Although the collective capabilities of each node will be large, the resources available to individual cores will be much smaller. Therefore, resources will need to be shared to a much greater extent between processes than before, much higher levels of parallelism will need to be exploited by applications, and sophisticated scheduling will be required to avoid significantly limited bandwidth.

We believe that the efficient use of node memory and processing elements are principal factors to achieving the levels of scaling and performance required for exascale applications. Project outcomes include development of new optimization techniques for computation on emerging extreme scale nodes; the integration of such capabilities with Global Arrays (GA) for system-wide scalability; and the demonstration at scale using a molecular dynamics driver application. Applications with communication-intensive execution patterns such as molecular dynamics can benefit even more when they can effectively exploit intra-node resources (shared memory, multicore processors, accelerators) to reduce their communication requirements. Careful assignment of logical tasks to multicore nodes to exploit data access locality is essential to reducing inter-node communication requirements for communication-intensive applications.



GA uses pointer metadata to compute offsets into other processes' address spaces and issue precise get() or put() operations for composite array slices.

In FY 2013, we focused primarily on two activities: performance evaluation of schemes to reduce metadata overhead and development of support for NUMA-aware distributed data allocation. We designed and implemented a framework to evaluate different metadata management strategies from full replication of metadata on every process to fully distributed metadata, which must be "remotely" queried on every communication operation.

Our analysis on metadata management strategies uses the GA's library-based PGAS model as a vehicle. GA needs a number of different metadata items to support operations on global-view dense arrays, which are its hallmark features. Among those items are the number of array dimensions, the primitive type of its elements (integer, single or double precision, etc.), sizes of array partitions (Cartesian distributions), and pointers to remote array sections. Currently, GA uses a fully replicated metadata strategy in which all "p" processes participating in an array instance creation maintain copies of the metadata for that instance ("p" could be all processes in the parallel job or a subset created via GA's process groups mechanism). This strategy enables direct, constant time (O(1)) access to any required metadata element by any GA process; however, it has a space cost of O(p) in all cases.

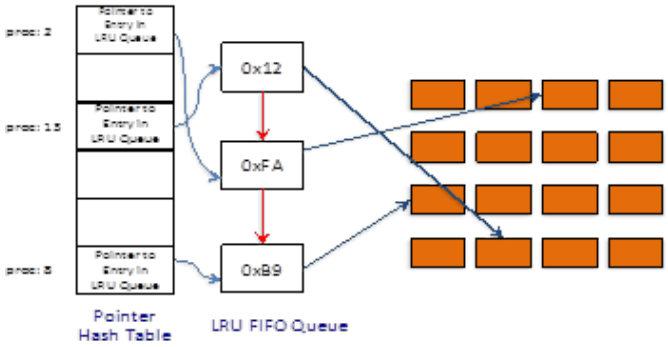
In typical GA applications, the number of active array instances has tended to be small (no more than 10 active instances at any point); however, the number of GA processes has been steadily increasing as computational scientists have gained access to ever larger HPC systems. For example, some of the largest GA-based simulations using the NWChem computational chemistry package have used up to 180,000 GA processes. If we consider only the memory space to store 64-bit remote pointers to array sections owned by all these 180K processes, we are using 1.38 MB per GA process for a single array instance. If the application is using 10 of these array instances, then we are using

PN11069/2391

13.80 MB of memory for data that is only incidental to the application. Given the small increases in (or even flat) memory capacity per core on trans-petascale systems, this becomes a non-trivial amount of space to use for metadata.

Given the abundance of cores on petascale and transpetascale systems, we propose the use of dedicated processes distributed across the system for the purpose of managing metadata. These metadata server processes are in charge of storing full replicas of all required GA metadata and responding to queries from compute processes. The processes are not intended to participate in the overall execution of the application except in their capacity to store and serve metadata. We consider multiple strategies for the placement and number of metadata server processes for GA applications, which include all processes being metadata servers to themselves and a logarithmic number of metadata servers across the GA process ranks. We also consider using k $\log_2 p$ servers (where k is > 1) to understand the tradeoffs between using more server processes and potential performance improvements.

Because compute processes no longer have direct access to all metadata, we use a fixed-sized LRU software cache to store recently used metadata pointers. The LRU cache size is kept small to guarantee a constant bound on the memory used for GA metadata. Our experiments evaluate several factors with respect to GA metadata accesses: number of processes, LRU cache size, and number of metadata servers. For each application, we present a principal result, which is the average metadata access time.



The LRU cache is implemented using an O(1) hash table lookup keyed by GA instance and remote process.

A key metric for the effectiveness of the metadata access mechanisms is the average metadata access time. A successful scheme should have an average access time close to the corresponding to a memory lookup when metadata is fully replicated. Average access times that are closer to the time required for a network transfer indicate that the metadata management mechanism is not effective. We average access times across the different GA process instances and also across different accesses present in the communication traces for each application.

Synergistic Integration of Feature Recognition and Analysis for Chemical Imaging Data

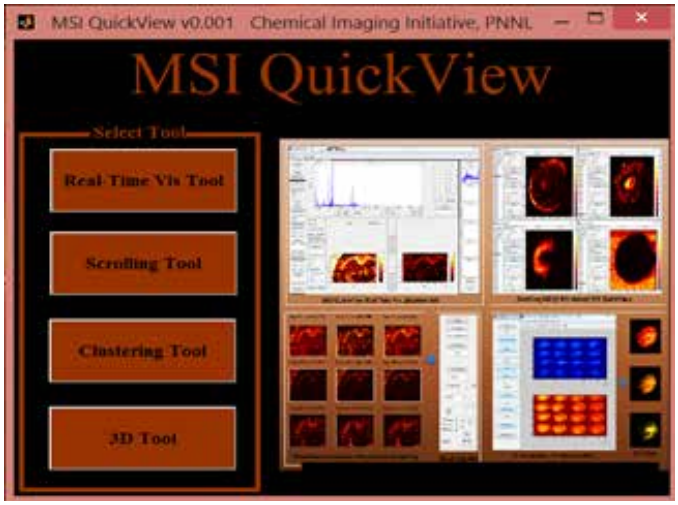
Mathew Thomas

This project is building unique new critical research capabilities to perform image analysis and feature recognition of large spatial time-varying datasets. The immediate impact includes facilitating knowledge-discovery in the fields of soil remediation and renewable energy production.

The ability to see the molecular details of chemical processes in both time and space is critical to future research in multiple areas of national interest, including energy production and utilization. Developing new instrumentation and integrating information from multiple data sources is an important step in achieving such new capabilities. Generally, there is a need to make existing feature recognition algorithms accessible to scientists and to customize such approaches to the particular needs of the specific data collection instruments and feature types. Currently, tools and techniques are being developed to image chemicals, materials, and biological molecules at resolutions ranging down to nanometer scale.

This project was initiated to establish new capabilities that allow scientist-driven interactive feature recognition and analysis for their chemical imaging-based data. The goal was to develop a toolkit with feature recognition algorithms integrated and customizable to specific chemical imaging acquisition protocols to make new analysis possible and provide more direct feedback to researchers during the image acquisition process. The toolkit was to be integrated into an image handling framework spanning from data acquisition to final results interpretation, which allow tools to augment positively data processing steps such as segmentation, clustering, and data visualization/interpretation.

In FY 2011, we initiated a discovery process to identify specific needs and skills in feature recognition and analysis for chemical imaging research. We created a prototype tool with graphical user interfaces to allow interactive feedback to chemical imaging scientists, and we applied this capability to two other projects. Our feature recognition tools are modular and expansible, enabling their application to a variety of other chemical imaging projects. For FY 2012, the toolkit was refined to increase both the utility of the toolkit and the capability for chemical imaging research. We additionally incorporated new objectives in the field of heterogeneous catalysis analysis. This work developed a test based on features stored in a quality database that determined if catalysis image data were unsuitable for analysis.



The main interactive interface of MSI QuickView for visualization and analysis of high spatial resolution mass spectrometry data (Nano-DESI).

PNNL High Resolution Mass Spectrometry. During FY 2013, in addition to extending the real-time capability of MSI QuickView, new tools with graphical user interfaces (GUIs) were developed, as described below.

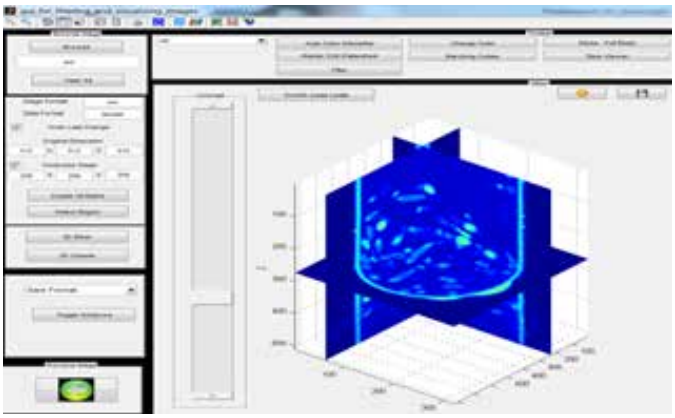
- **Scrolling GUI**. This tool allows the user to scroll through quickly and/or click through and perform image processing on the ion images. By default, only single ion images at a time are loaded into computer memory. This efficient approach enables multiple instances of Scrolling GUI to work simultaneously to compare different data set. The GUI provides the option to load the entire data matrix into memory as well.
- **Clustering GUI**. This tool uses the agglomerative hierarchical clustering function in MATLAB. The clustered images are saved to a PDF file, with each class starting at a different line. The tool can be used to generate ion images for a full data set by specifying a coefficient value such that all ion images belong to one class.
- **3D GUI**. This tool uses a semi-automated approach to separate and register 2D slices imaged side-by-side to generate 3D images. The optical image is first warped to change it to the coordinate system of the 2D ion images created using the Real-Time Visualization GUI by picking identical pair of points on the optical image

PN11074/2396

and an ion image and then using the spatial transformation function in MATLAB. The warped optical image is then cropped out into individual tissue slices. Starting with the first slice as the reference image, adjacent slices are registered to the coordinate system of the reference image using the 2D spatial transformation function in MATLAB by picking identical pair of points on both images. The registration, warping, and cropping parameters are saved and used to automatically repeat the process for other 2D ion images from the data set. The registered slices are then loaded into an open-source visualization and analysis toolkit for 3D visualization.

Image processing capabilities include normalization, time-alignment, intensity scaling, smoothing, and overlay of ion images. The user can create lines scans or regions of interest (ROI) over the ion image to display the intensity spectrum for pixels within the ROI as a function of location. All pixel values within the ROI are displayed in a table and can be saved to an excel spreadsheet. Only single lines of data are loaded into memory at a time. The ion images can be normalized to the TIC (total ion current) or a user-specified m/z range. To handle datasets containing MS/MS imaging data for 92 precursor m/z values, new functions were added to first split the raw files into 92 different files before analyzing them in MSI QuickView.

ALS X-Ray Tomography. Additionally in FY 2013, BiofilmViewer was extended to enable better segmentation, analysis and visualization of biological samples imaged using X-ray tomography techniques. Several existing open-source tools like Chimera, Icy, ImageJ, Vaa3D, Drishti, Cell Profiler, Paraview and BioImageXD, each with different capabilities,



The main interactive interface for Biofilm Viewer for visualizing and segmenting biological samples imaged using X-ray Tomography techniques.

were embedded into BiofilmViewer. This allows for automatic data transfer between the different tools. These tools combined with Velo (collaborative data management and analysis infrastructure) and REXAN (component library for Rapid Experimental Analysis) and GlobusOnline supported the remote near-real-time analysis needs, allowing scientists to receive critical feedback while at the experimental facility, increasing the quality of the resulting experimental data.

The toolkits have been integrated into the framework for managing chemical imaging data, which provides faster and more accurate feedback to the researchers conducting imaging experiments. The toolkits developed in this project for this year have contributed to four papers and one online journal article.

Targeting Extreme Scale Computational Challenges with Heterogeneous Systems

Oreste Villa

We developed approaches, techniques, and methodologies (in the form of software libraries or prototypes) to facilitate scientific application mapping on large-scale heterogeneous machines.

Novel high performance computing (HPC) systems are heterogeneous: they integrate both multicore x86 processor together with Graphic Processing Units (GPUs) for general purpose computation. Developing and implementing approaches that efficiently exploit all the processing elements at the same time is fundamental to reach the highest performance enabled by these platforms. Under this project, we studied and implemented techniques and methodologies to support scientific applications on large-scale heterogeneous system. Challenges associated with the use of this system involves refactoring of commodity applications to support new hardware features, which are dictated by evolution in architecture design necessary to achieve higher performance (shorter simulation times or larger simulations) or reduce power consumption.

This project is an important cornerstone in understanding benefits and applicability of new computer architectures for applications of strategic importance such as molecular science and classical and quantum chemistry. The methodologies and techniques developed will allow the next generation of molecular science, classic, and quantum chemistry applications to run efficiently on next generation large-scale heterogeneous systems. For instance, we implemented novel code generators to allow for that enable easier application porting on new architectures (from high level description manageable by domain scientists) and software layers that enable transparent load balancing of batched solver for small linear systems across heterogeneous processing elements, while also considering power consumption.

Specific code generator. We developed and implemented a domain specific code generator and related auto-tuners for tensor contractions in NWChem that target a variety of architectures. The code generator can produce code automatically for GPUs with CUDA, OpenCL with and without multicores (AMD, NVIDIA), standard C/Fortran, and DGEMM + Sort. The code generator supports many options and different optimization strategies depending on the target device. Different code versions and varying processors present different performance tradeoffs depending on input type. We implemented runtime auto-tuning strategies that exploit heuristic training methods to select the most efficient (for performance and power) options. We also demonstrated up to 8 times performance improvement over conventional approaches and applied the new tensor contraction engine on the multi-reference coupled cluster problem, obtaining significant speedups with respect to previous solutions.

Quantum chemistry. We implemented a novel chemistry reaction module for subsurface flow, climate modeling, and carbon sequestration codes. The chemistry reaction module simultaneously solves a large batch of small linear systems (millions with up to 100 variables each), exploiting the Newton-Raphson method and LU decomposition. We tested three different versions: first, we verified that the CUBLAS version has limited flexibility, only working with small systems (32 variables) and only supporting partial pivoting. The NVIDIA solution can manage larger systems (up to 76 variables) and is faster but only supports partial pivoting. Finally, our custom implementation is slightly slower but more flexible, supporting systems over 120 variables plus partial and full pivoting. We also integrated the module a heterogeneous load balancer able to distribute systems transparently to solve CPU cores and GPUs and divide the workload so that different components complete their own part simultaneously, obtaining an overall speedup. When these approaches are integrated in STOMP, the full application provides 6 to 7 times speedups on large problems.

Subsurface flow, climate modeling, and carbon sequestration. Besides working on the reaction module, we evaluated opportunities to accelerate the coupled flow and transport component for subsurface flow, climate modeling, and carbon sequestration simulation. We looked at two different strategies. The first consists in substituting the de facto standard solver for these problems (PETSC) with a CUDA multigrid linear system solver under development at NVIDIA (NVAMG), helping to identify the challenges in providing an easy, efficient application programming interface for substituting standard solvers with GPU-accelerated solutions. The second strategy was developing our own custom code for GPU solvers, in which we implemented Gaussian elimination with partial pivoting. Preliminary tests show that our code provides significant speedups with respect to the CPU solution.

This year, our accomplishments have been widely disseminated in the community. Highlights included two journal articles, presentations at two well-known HPC conferences, and a technical presentation at another technology conference.

PN12074/2475

TASCEL: An Execution Model for Task-Based Optimizations

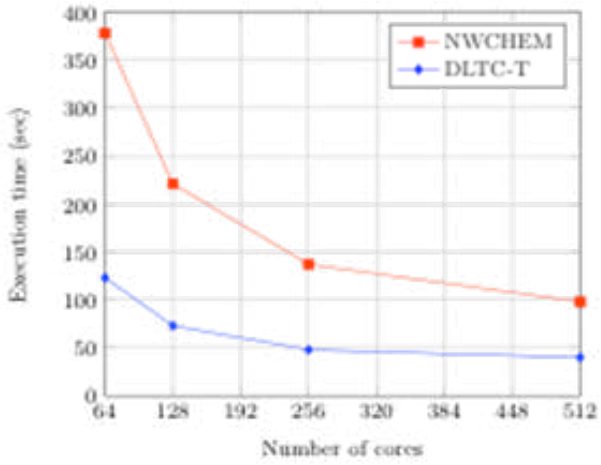
Sriram Krishnamoorthy

We are developing a task-based execution model that exposes a higher level interface to programmers with a runtime that manages automatically the scheduling of computation and communication to optimize various metrics.

Performance portable implementations of applications for next generation supercomputers requires a careful application design to adapt execution schedules to match architectural and input-specific application characteristics as the problems and scales at which they are executed evolve. Optimization strategies vary not just across architectures but also with the type of problems (strong or weak scaling, capacity or capability runs, stand-alone or a larger calculation). Increased scale and variability in extreme scale computing systems makes it challenging to manage them effectively. The scale-out in high-end systems has resulted in small load imbalances, causing large performance interruptions. The churn in the hardware architecture necessitates a programming paradigm abstract enough to be programmed easily to key hardware features while detailed enough to provide performance portability. Challenges of load balancing and fault tolerance are exacerbated when a process-centric programming model is employed by the application.

This project focuses on developing task-based algorithms for distributed memory systems. While several specific instances of this model have been studied previously, we will design a runtime that incorporates key elements of prior approaches with focus on the design of scalable algorithms suitable for extreme scale platforms. The runtime manages scheduling of the computation and communication to optimize various metrics (power, performance). Developed algorithms will handle various use cases for data access characteristics, computation granularity, communication requirements, and anticipated regularity. Algorithms also optimize key computational chemistry methods and will be released as part of the TASCEL library for use by the wider scientific community, enabling development of performance portable applications for adaptation to underlying architecture with selective localized changes to the application through careful choice of algorithms and optimizations. The result will be sustainable software for effective porting to future supercomputing platforms with reasonable effort.

We developed a thread-safe, active message-based framework for the design and development of task applications. Thread-safety enables natural sharing of data in a many-core system. Given the anticipated increase in core counts on future systems, thread-safety prepares applications for effective execution on future extreme scale systems. The active message framework enables efficient execution of complex transactions that could involve multiple round-trips using remote memory access protocols. The multi-threaded framework supports flexible progress for active messages, enabling development of concurrent observer computations that act as sub-programs running concurrent with the application, observing and reacting to hardware actions.



Comparison of NWChem and DLTC performance in evaluating CCSD on green fluorescent protein.

A popular approach to scheduling task-parallel programs, work stealing is flexibility inherent in load imbalance and can result in seemingly irregular computation structures, complicating runtime behavior. We designed an approach to trace async-finish parallel programs efficiently using work stealing by identifying key properties to trace task execution with low time and space overhead. We demonstrated that the perturbation from tracing is within the variation in execution time with 99% confidence and the traces are concise, amounting to a few tens of KB per thread in most cases.

We developed Dynamic Load-balanced Tensor Contractions (DLTC), a domain-specific library for efficient task parallel execution of tensor contraction expressions, a class of computation encountered in quantum chemistry and physics. Our framework decomposes each contraction into smaller unit of tasks, represented by an abstraction referred to as iterators. We exploit an extra level of parallelism by concurrently executing tasks across independent contractions through the functionality implemented in TASCEL. By executing tasks across independent contractions concurrently, DLTC eliminates many synchronization points enforced by conventional approaches, leading to higher processor utilization. We demonstrated the improved performance, scalability, and flexibility for the computation of tensor contraction expressions on parallel computers using examples from coupled cluster methods.

PN12075/2476

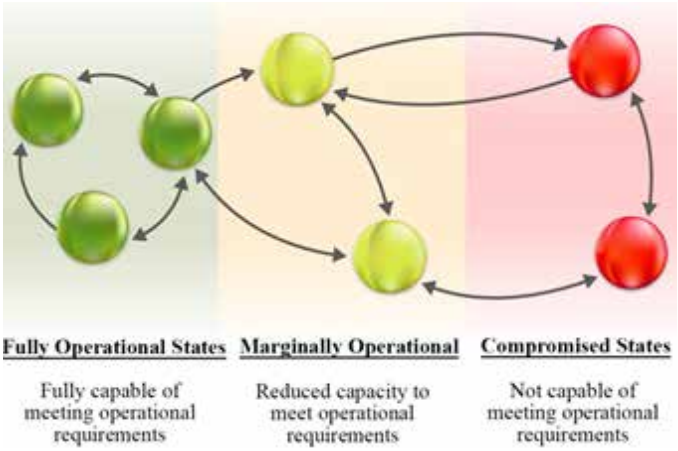
Theory of Resilience

Pradeep Ramuhalli

We are developing and evaluating a theoretical framework for resilience to provide the technical basis for evaluating cyber defense and reconstitution approaches.

Effective approaches for resilience in cyber systems are needed to keep critical infrastructure operational in the face of adversarial attacks. Natural and intentional manipulation of data, computing, or coordination are the most impactful ways an attacker can prevent an infrastructure from realizing its mission goals. Thus, the ability to maintain mission-critical cyber-system operations in the face of disruptions is critical. The quest for such resilient systems has seen a number of technical approaches (such as moving target defense) that attempt to add specific properties such as diversity, redundancy, deception, segmentation, and unpredictability to make the system more resilient to one or more attack vectors. However, the approaches cannot be applied readily after a system has been compromised. It is also difficult to ascertain whether such methods are applicable under any attack scenario or if limitations exist.

The objective of this research is to develop and evaluate such a theoretical framework that incorporates key concepts of resilience in cyber systems and integrates decision making so that decisions in a compromised environment may occur with the goal of operation continuity in a fast, cost-effective manner. The resulting framework will enable 1) identification of key components of resilient cyber systems; 2) decision making with uncertain information (i.e., the design and recovery decision-making methodologies may need to function with limited information); and 3) asymmetric advantage (increased cost to the adversary should be a factor in the decision making). The results of this research will also provide insights into measurement needs and success criteria for asymmetric resilient cyber infrastructure.



Conceptual representation of reconstitution in cyber systems, moving the system from one of several compromised states to a fully operational state.

Current frameworks for resilience encompass four basic concepts that collectively enable resilient design (anticipate, withstand) and dynamic reconstitution (recover, evolve). In FY 2013, the focus of the technical developments was on dynamic reconstitution. Including recovery and adaptation, the reconstitution response may require significant reconfiguration of the system at all levels to render the cyber system resilient to ongoing and future attacks while attempting to maintain the continuity of operations.

We developed a state-space based formulation for use in dynamic reconstitution of cyber systems. Specifically, the state C_t of a cyber system at time t is defined as a representation of its key properties, including network connectivity, critical service configurations, and cost. When a system is compromised, it is seen as moving from a fully operational state to several compromised states as it loses the ability to maintain continuous operations. Reconstitution is then defined as moving the system back to one of a few fully operational states, possibly through several intermediate states. In general, determining secure configurations in a reasonable time frame is difficult; however, computing "good enough" solutions may be possible. The proposed formulation enables the application of optimization approaches with metrics for resilience and continuity of operations to determine at each time step subsequent to the disruption a set of states (such as network connectivity, services, and hosts) that secure the continuity of operations while improving the resilience and minimizing the cost of cyber defense.

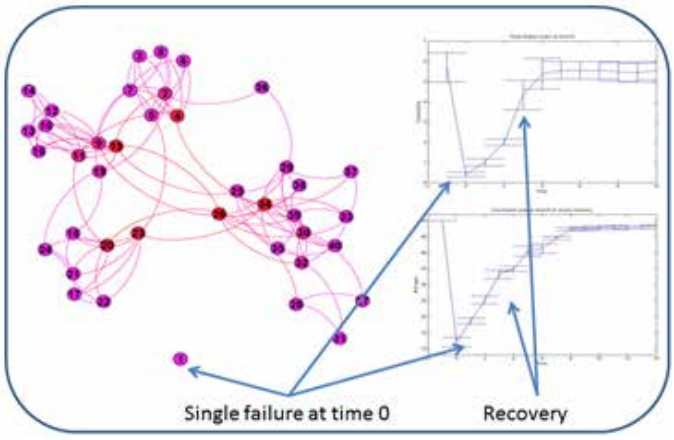
We tested this proposed formulation using simulations that abstracted small (30–40 node) cyber systems as graphs. An n-node network with m critical services is described by a graph G that captures the connectivity structure of the network. Here, a node represents some computational resource, such as a computer or a server that can run any subset of the m services. The matrix ϕ describes the configuration of each node with respect to the critical services. A set of random networks are simulated with random connectivity patterns and configuration matrices. The simulation assumed that events can occur that result in hardware and services becoming unavailable, thus disrupting continuity of operations. Using the state-space formulation for

PN13063/2544

these systems, we applied a genetic algorithm to determine solutions for connectivity and service deployment that maintained continuity of operations.

Results of simulation studies indicated that the proposed approach may be viable for reconstitution of compromised cyber systems. Applying the proposed approach to reconstitution, network services were able to be reconstituted quickly after a fault at time step zero. At the same time, the reconstituted system in the simulations was deemed to be more resilient based on the selected metric (network connectivity). Unlike existing techniques, the proposed approach was seen to be viable even when disruptions occurred during the reconstitution process. Based on simulation results, the approach was capable of reconstituting systems and improving overall resilience in the presence of serial disruptions (i.e., as the system is recovering, a subsequent disruption occurs); however, the recovery may be somewhat slower as the approach needs to account for severe reductions in capacity. The formulation also provides a mechanism for evaluating performance of other approaches to reconstitution of compromised cyber systems. This formulation and results were presented at an IEEE international conference.

To date, the simulation studies included several assumptions that will need to be relaxed if the results are to be more generally applicable. This research is planned for



Example of simulation results. The network connectivity map (left) at time 0 shows a failure of a node due to some event that negatively impacts metrics measuring connectivity (top right) and continuity of operations (bottom right). When applied after the failure, reconstitution methodology helps the network recover while improving resilience (measured using connectivity) and continuity of operations.

FY 2014. In addition, next year's research will examine expanding the theoretical formulation for reconstitution to address concepts of overall resilience of the cyber system. Theoretical development and limited experimental evaluation will be performed to address decision making with uncertain information, and asymmetric advantage.

Visualizing Uncertainty in Conceptual and Numerical Models for Geological Sequestration

Luke J. Gosink

This research will help scientists characterize and quantify uncertainty in the conceptual and numerical ensembles that are commonly used to assess the viability and long-term impacts of sequestering carbon dioxide at select geological regions.

Despite being actively researched for several decades, the field of uncertainty quantification (UQ) is still rapidly evolving and developing. Existing UQ research attempts to quantify the uncertainty in a given system based on the system's reliance on instrumentation (used to obtain ground truth), conceptual modeling structures (the mathematical and domain-specific models used to create a system that can predict and simulate a real-life event), and numerical approximations (the methods used to approximate the operation of the system). A fundamental challenge to assessing uncertainty and accuracy in sequestration models arises from the fact that ground truth data – essential for most UQ and risk analysis tasks – is largely unavailable; in many cases reliable, sufficient subsurface ground-truth is too costly or infeasible to obtain.

Our objective is to address the above challenge by developing novel methods for characterizing uncertainty that can be used in the presence of minimal ground truth measurements. We combine the analytical strengths of statistics with the interactive power of scientific visualization and visual analytic methods. This approach addresses uncertainty through the comparative analysis of varying scenarios, considering all likely events predicted by numerous systems (i.e., varying conceptual and numerical models) for a fixed sequestration site. Our integrated approach is novel in the geological sequestration, visualization, and uncertainty quantification communities. As an integrated component of the Geological Sequestration Software Suite (GS^3), this project supports risk analysis, management, and mitigation tasks for sequestration activities through a unique combination of techniques derived from the fields of statistical analysis, scientific visualization, and visual analytics. Our work will ultimately allow scientists to advance the understanding of sequestration modeling and more accurately assess the performance of and long-term impacts to sequestration activities at potential reservoirs.

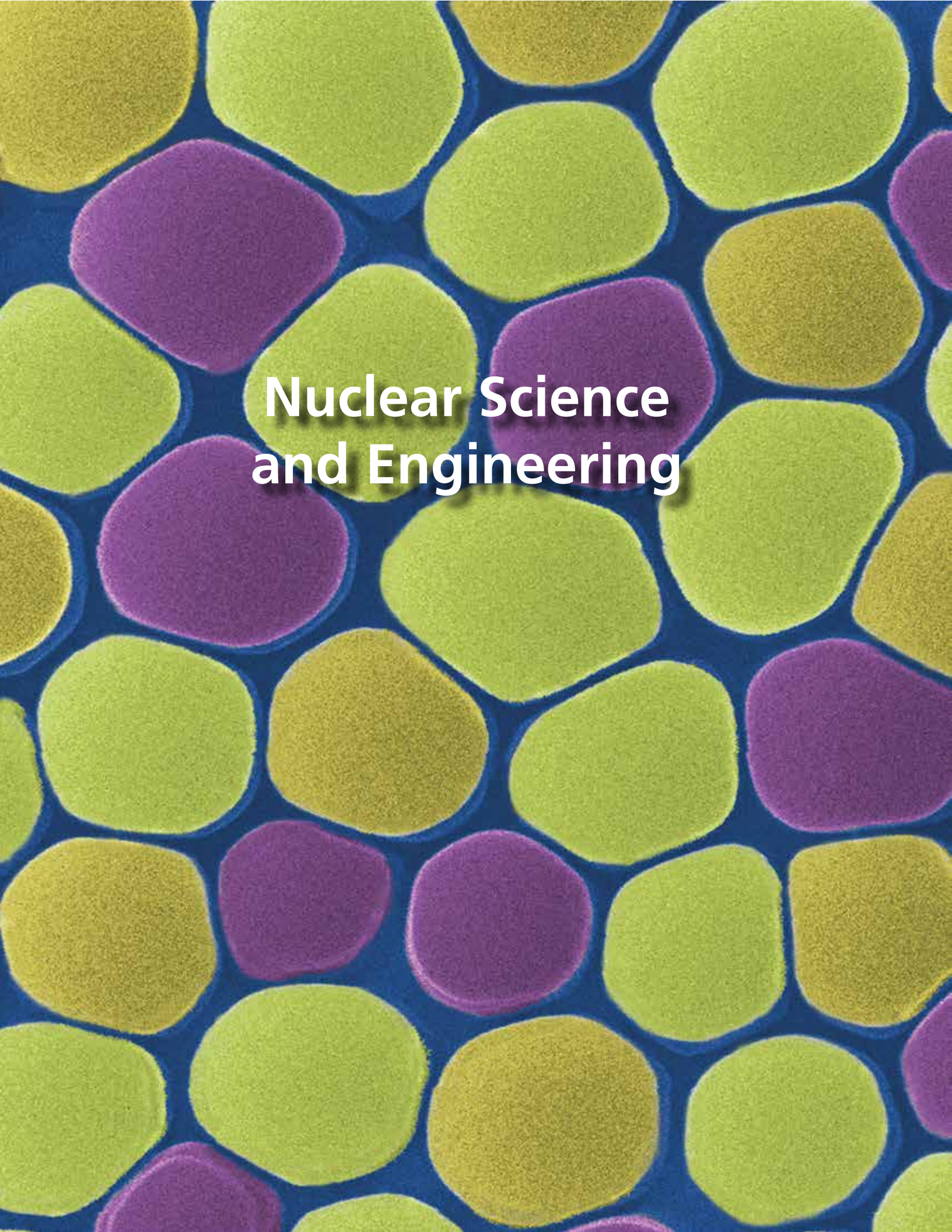
Our work during FY 2010 focused on three stages: identifying aspects of uncertainty to address within GS^3, identifying viable strategies to address these uncertainties, and acquiring data for testing and validating our strategies. Initially, we identified the varying sources of uncertainty present in each of the existing projects supporting GS^3. We collaborated and consulted with other projects and identified both elements of uncertainty and potential strategies for detecting and characterizing these uncertainties. Ultimately, we identified uncertainty to address in two stages of GS^3's modeling framework: its conceptual and numerical model development stages.

For FY 2011, we completed construction of our statistical analysis framework for performing UQ. This framework is based on a Bayesian approach called Bayesian Model Aggregation (BMA) that combines ground truth observations (i.e., monitoring data) with the predictions obtained through an ensemble of conceptual and numerical models. The framework facilitates UQ analysis for both the ensemble and its constituents by providing a mechanism for identifying and quantifying the spatial, parametric, and temporal regions where uncertainty and model errors are greatest. We extended this statistical framework with a preliminary suit of scientific visualization software tools. These tools, which are working prototypes for this project's eventual application deliverable, help to convey the information generated from our statistical framework in a more meaningful way by visually conveying (through isocontours and surfaces, scatterplots, and time series displays) where model uncertainties are greatest.

During FY 2012, we completed validation studies on our statistical framework. Our approach and the validation results were presented at several conferences. We also completed the development of our software application called Dynamic Real-time Ensemble Analysis and Monitoring (DREAM). This software allows modelers to launch and monitor an ensemble of numerical simulations. Additionally, while simulations are running, DREAM enables modelers to visualize and analyze the simulations in-flight. The integration of our statistical framework with DREAM effectively forms an uncertainty quantification pipeline. This tool is a core component in the GS^3 environment and will therefore be widely accessible for all GS^3 clients.

In the first half of FY 2013, we completed the DREAM software tool and integrated this software into GS^3. We also published our methodology for characterizing and visualizing predictive uncertainty.

PN10079/2321


Nuclear Science
and Engineering

Argon-39 Measurement

Jill M. Brandenberger

This project developed and tested an ultra-sensitive argon nuclear measurement capability that has relevance to groundwater residence times for climate change, environmental remediation and national security applications.

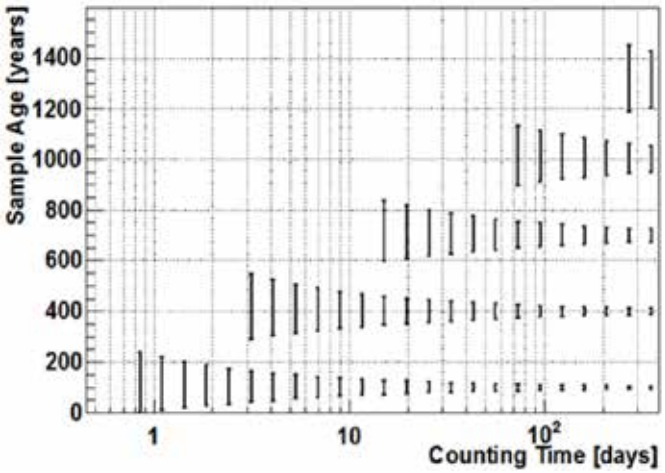
The scientific challenge of this project was to establish a U.S.-based capability for ultra-low-background proportional counter measurements of argon to support treaty verification and water reserve (e.g. groundwater) measurements using radioisotopes. Argon-39 (^{39}Ar) is a cosmogenic isotope found at a constant, but low, concentration in the atmosphere. It is thus present as a dissolved gas in rainwater and snow which recharges aquifers. Once in an aquifer and thus isolated from exchange with the atmosphere, ^{39}Ar radioactive decay decreases its abundance. The 269-year half-life of ^{39}Ar compared to the 12.32-year half-life of tritium can allow determination of the age of groundwater sources over an important time range spanning 50 to 1000 years. This elucidates the time-scale of impacts expected due to modern changing climate patterns.

This project provided the first precise measurements at PNNL of atmospheric vs. geologic (ancient) ^{39}Ar activity. This confirms PNNL's low-background proportional counting system, constructed with ultra-clean materials in the new shallow underground laboratory, has the capability to make argon isotopic measurements with the precision necessary to address water resource questions. Along with others, this project established an ultra-sensitive ^{39}Ar nuclear measurement capability and demonstrated it though measurement of ancient argon. Ancient argon is depleted in ^{39}Ar and possible sources include dissolved gas in deep ancient aquifers and argon trapped in natural gas reservoirs. The ancient argon was provided by collaborators from the Fermi National Accelerator Laboratory (Fermilab) via a National Nuclear Security Administration (NNSA) research project aimed at identifying a reliable source of ancient argon for use in radiometric gas-proportional counters.

The use of the ancient argon in PNNL detectors effectively eliminated ^{39}Ar activity and allowed detector background signals to be elucidated, thus establishing the magnitude of a modern ^{39}Ar signal vs. limiting backgrounds. To support the ^{39}Ar detection, the project developed a degasification instrument for removing argon from groundwater and leveraged an argon separation and purification process for argon gas separations from soil gas and air samples. The argon separation system was developed for ^{37}Ar measurements in support of treaty monitoring. Specifically, this project provided the optimization and experience on the system to support future argon work.

The ultra-sensitive ^{39}Ar measurement capability leveraged the shallow underground laboratory and other research to develop low background materials and proportional counters. We specifically provided operating parameters and analysis packages for ultra-sensitive measurements of ^{39}Ar and ^{37}Ar activity. A significant number of argon samples (around 24) derived from environmental soil-gas collections were measured on this project in FY 2013. These measurements supported project research by establishing calibration, analysis, and measurement optimization for ^{37}Ar. Additionally, this work directly supports current proposals for development of ^{39}Ar measurements for groundwater chronology and ^{37}Ar field systems.

Ancient or geological argon measurements from a gas well samples were used to demonstrate the performance of these proportional counters and the sensitivity of ^{39}Ar measurements on samples as old as 1,000 years. These published results demonstrated PNNL's capability to measure ^{39}Ar at a sensitivity required for chronology of groundwater with recharge rates of anywhere from 50 up to 1,000 years. This lays the foundation for a new project to provide the age of select groundwater aquifers in support of climate change science and groundwater remediation. This capability also supported efforts to establish argon calibration sources to provide the international community with a tool for detector inter-calibration.



Expected water age in years vs. counting time with 1σ error bars.

PN12009/2410

Atomic Mass Separation for Enhanced Radiation Detection Measurements

Gregory C. Eiden

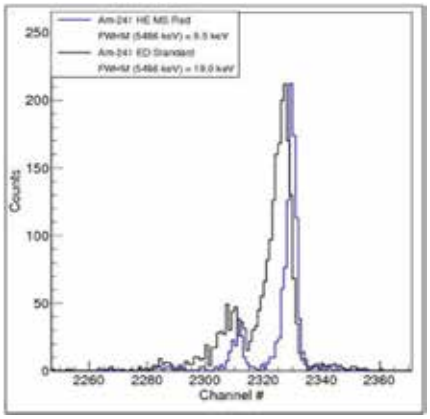
This project seeks to improve the measurement and detection of radionuclides of interest in nuclear security.

Sensitive measurement of radionuclides requires high fidelity chemical separation of the isotopes of interest from the sample matrix and state-of-the-art radiometric counting of the nuclide decay events in the separated sample. Previously, the only option for preparing samples for counting was chemical separation. While most elements have distinct chemical behavior that can be leveraged for separations, there are (with few exceptions) enough similarities that separations are never completed. This research seeks to develop separations that are based on radionuclide atomic mass. With mass spectrometers (MSes), this type of separation can be many orders of magnitude more effective than the chemical method.

Our objective is to explore analytical performance of the prototype instrument, develop modifications, and obtain proof-of-concept data for key beta and gamma emitters. The prototype system can be used for almost any radionuclide and sample matrix. Our focus is developing instrumentation and methods for use in fixed laboratory analysis of radioactive environmental samples associated with fission and activation products from reactors (reactor accidents or routine releases) and terrorist use of a nuclear device. We anticipate that our novel approach will create a paradigm shift in the way radiochemical separations are thought about by experts in the field. The technique will lead to improved methods in both fixed and field deployed instrumentation, especially for nuclear accident and terrorist incident response.

We developed a prototype instrument with which individual atomic masses can be implanted into a solid target, thereby enabling isotopes of a given atomic mass to be characterized radiometrically. The instrument is a modified quadrupole mass filter-based inductively coupled plasma MS. The detector was replaced by a simple metal foil to capture mass separated ion beams. Future instruments will include complex target arrangements: multiple targets, variable target regions, and ion detectors.

Initial experiments sought to address several analytical or method issues: efficiency of ion beam capture, the spatial distribution of captured ions on the target, the effect of molecular ions and ion scattering on mass separation efficacy. We are also considering means for calibrating the counting of MS-RAD separated species on solid targets: it is not enough to obtain a well separated isotope to count, as the measured activity must be related to the original sample amount to obtain its specific activity. Calibration will be addressed in FY 2014. The instrument was first tested using stable isotopes to characterize how effectively one isotope could be separated from an adjacent mass isotope. We obtained more than four orders of magnitude separation of Europium isotopes, limited by the analytical dynamic range. Based on our knowledge of the quadrupole mass filter performance, the actual separation should be ~6 orders of magnitude. In FY 2014, we will obtain higher dynamic range data to demonstrate six orders of magnitude separation or more.



Comparison of ^{241}Am alpha spectrum for conventional sample preparation (electrodepositionED) vs. MS-RAD mass-based separation and capture on a metal foil target. The full-width at half maximum of the MS-RAD spectrum is about half of the electrodeposited sample.

To aid our analysis and experiment design, we developed a computer program that calculates decay rates as a function of time for mixtures of fission products from different fissile isotopes undergoing fission induced by neutrons of various energies. The program also figures the beta, gamma, and alpha radiation as a function of time for each fission product. This program solves the Bateman equation for in-growth and decay of isotopes at a given mass and thus provides a detailed look at what is expected at each mass that might be captured on a target using MS-RAD.

The first radioactive isotope implanted was ^{241}Am. The spectra obtained demonstrated a key advantage of an MS Rad separated sample for alpha spectrometry: the atoms of interest are captured at very low kinetic energy on the metal foil and are thus at the surface of the foil or implanted at most a few atomic layers deep and the sample is virtually free of any carrier solution. The thin, clean MS-RAD sample results in alpha peaks that are both less shifted in energy (less straggle of the alphas escaping from the sample) and narrower. Modeling of the high resolution ^{241}Am MS-RAD spectra provided new insight into the processes and nuclear data parameters responsible for select features in the alpha spectra.

In FY 2014, the focus will shift to beta and gamma emitters because these are the main decay modes for the isotopes of interest in the analysis of fission products, either old fission products from cooled, spent nuclear fuel or fresh fission products from a reactor accident or a terrorist use of a nuclear explosive device. The prototype instrument will be modified to enable multiple implants to be accomplished during a single pump down of the instrument. Analytical issues, including methods for calibration of specific activity determinations, will be addressed such that analyses can be compared with conventional radiochemical methods.

PN13015/2496

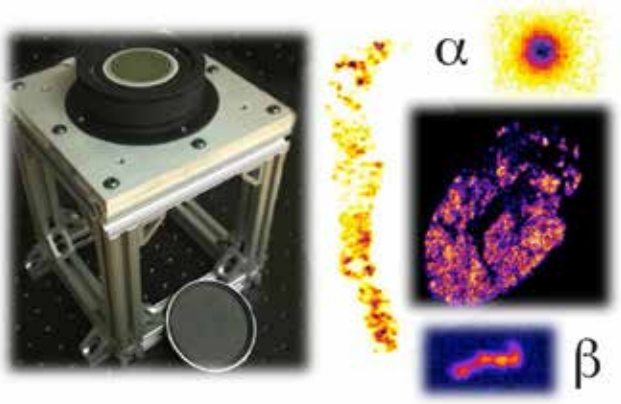
BazookaSPECT Neutron Imager

Brian W. Miller

This project is developing a high-resolution ionizing radiation imaging detector technology for specialized use in medical imaging, nuclear science, and national security applications.

During the last several years, BazookaSPECT, a new class of gamma-ray detector, has been developed that combines columnar scintillators, image intensifiers, and charge coupled device/complementary metal oxide semiconductor (CCD/CMOS) sensors for high-resolution imaging. This detector has been employed primarily in pre-clinical medical imaging applications such as single-photon emission computed tomography (SPECT), which monitors level of biological activity (instead of just taking a picture) in a 3D region of anatomical structures. We are interested in exploring and characterizing the use of BazookaSPECT for detecting additional forms of ionizing radiation (e.g., neutrons, alpha particles, fission fragments, betas, and conversion electrons) for use in national security and nuclear science applications.

Our objective of this project is to establish a detector capability at PNNL and investigate, demonstrate, characterize, and improve the detection response to ionizing radiation. We are building a high-resolution position-sensitive neutron detector (~100 µm resolution, 10–16 cm diameter) that will be utilized in a variety of imaging experiments such as neutron radiography, diffraction, and neutron coded aperture imaging. In addition, we will seek to develop applications that are enabled or improved upon by the technology. After establishing the capability, we aim to expand BazookaSPECT in terms of hardware, software, and applications.



The PNNL iQID detector, images of alpha and beta particle interactions, and selected alpha-particle autoradiographs of tissue sections.

As noted above, our focus was evaluating the response of the detector to other forms of ionizing radiation, including fission fragments, alpha, neutron, and beta particles. Results have been successful and, as a result, the detector has been renamed ionizing-radiation Quantum Imaging Detector (iQID) to represent its broad response to ionizing radiation.

During FY 2013, we focused on three major tasks:

Developing iQID hardware and software. We developed an iQID camera that has an active area 125 mm in diameter to accommodate applications that require a large imaging area. A goal of this project has been to make the detector portable so that it can easily be transported for use with collaborators outside of the laboratory. To this end, we developed an easily transportable iQID system that runs using a laptop computer.

Studying iQID's response to beta, alpha, fission fragment, and neutron particles. Early this year, we demonstrated that iQID has high sensitivity to variable forms of ionizing radiation. It quickly became apparent that the detector would work well as an auto-radiography camera. In particular, it provides a number of benefits over other methods, including higher spatial resolution, sensitivity, real-time imaging capability, portability, and activity quantification. We showed that the various particles can be discriminated from one another using signal information and/or different scintillation materials.

Developing new applications and collaborations. Shortly after discovering the potential of the iQID camera as a digital autoradiography detector, we established a collaboration with researchers at Fred Hutchinson Cancer Research Center in Seattle, WA. They focus on radioimmunotheraphy research, which uses cancer-targeting antibodies labeled with alpha-emitting isotopes. The alpha particles are highly potent at killing cancer cells and are a promising therapy. However, estimating dosimetry at the cellular level or within a small cluster of cells is key to their research. They were excited at the prospect of a detector that would enable digital autoradiography of alpha particles for microdosimetry. This collaboration continues, and iQID is becoming an integral part of their research studies. In addition to imaging alpha particles, we are now starting to explore the potential of the iQID for beta-particle autoradiography.

In FY 2014, we will continue our emphasis on these major tasks from the last year. In particular, we will explore methods for improving the detector spatial resolution with beta particles. We will also focus on internal work and collaborations for nuclear science- and national security-related applications.

PN12012/2413

Characterization of a 14 MeV Neutron Generator and Measurement of Fission Products Produced

Erin C. Finn

The installation and characterization of a 14 MeV neutron generator at PNNL has enabled nonproliferation research and development (R&D) by supporting the production of unique radionuclides for use in nuclear advancements, particularly the measurement and evaluation of possible new short-lived isotope signatures.

Historically, PNNL's efforts in nuclear nonproliferation R&D have been thwarted by a lack of readily available, relevant radioactive standards. In addition, realistic radioactive samples are required to test new measurement and data evaluation methods accurately and to improve nuclear data libraries. The facility housing the 14 MeV neutron generator at PNNL is unique and enables us to achieve a spectrum with minimal scatter. Before the 14 MeV generator can be used to produce the samples required for high accuracy nuclear measurements, the instrument must be carefully set up and characterized to ensure a consistent high flux of 14 MeV neutrons with minimal downscatter.

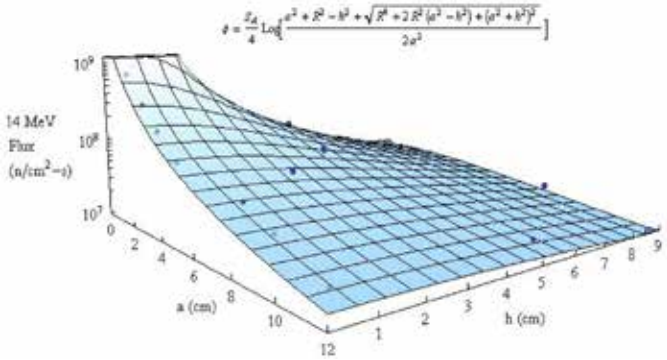
The objective of this work is to make the generator capability available to a diverse user group. Maintaining the capability requires that all appropriate procedures and permits is in place, along with an understanding of how to operate the system and what results to expect from various operating schemes. To accomplish this, MCNP modeling is used to predict results, which are then experimentally tested and validated.

Effective experiment planning requires a baseline series of quantitative experimental measurements of the neutron environment, flux, and spectrum available. Such measurements are made using standard dosimetry techniques, wherein high purity wires or foils with known neutron activation cross-sections are irradiated under known conditions. The activation of the materials is measured by gamma spectrometry and is used to calculate the number of neutrons incident on the sample. Additionally, the use of multiple materials concurrently allows measurement of the neutron spectrum in addition to the neutron fluence.

The Monte Carlo N-Particle (MCNP) model built last year was updated in FY 2013 to provide improved energy resolution of the neutron spectrum. The improved model more closely agrees with experimental data that were collected to enhance the preliminary characterization completed in FY 2012. The new experimental data expand our understanding of spatial effects on the neutron spectrum and flux. The comprehensive characterization has enabled the good prediction of irradiation conditions for diverse samples in multiple geometries under both this and other projects. The characterization of the neutron generator performance in our unique facility was submitted to *Nuclear Instruments and Methods Part A*.

Finally, a fission product sample was produced to evaluate the fraction of fissions induced by 14 MeV neutrons compared with lower energy neutrons. Greater than 92% of the fissions were induced by 14 MeV neutrons. The results demonstrated the ability to produce a sample primarily composed of products from 14 MeV neutron-induced fission and verified predicted rates of fission. Also, the outcome determined that it is feasible to produce fission product samples large enough to perform radiochemical separations for measurements. A paper on this work was submitted to the *Journal of Radioanalytical and Nuclear Chemistry*.

The performance of the neutron generator has been extensively characterized for both the neutron spectrum and the neutron flux at multiple positions relative to the neutron generator head. This data have enabled successful irradiation planning for multiple users, including the NNSA Office of Defense Nuclear Nonproliferation (DNN) and the National Center for Nuclear Security R&D-funded projects.



MeV flux with experimental data points superimposed on the surface ϕ, which describes the anticipated 14 MeV flux in space around the neutron generator target. The quantity "a" is the vertical distance from the center of the target disc, and "h" is the horizontal offset from the center of the target disc. Dark blue points lie above the surface, and light blue points lie below.

PN12014/2415

Exploiting Correlated Radiographic and Passive Signatures for Threat Detection in Cargo

Erin A. Miller

This project develops a method to improve radiological threat detection in cargo containers by systematically combining information from passive radiological sensors and radiography.

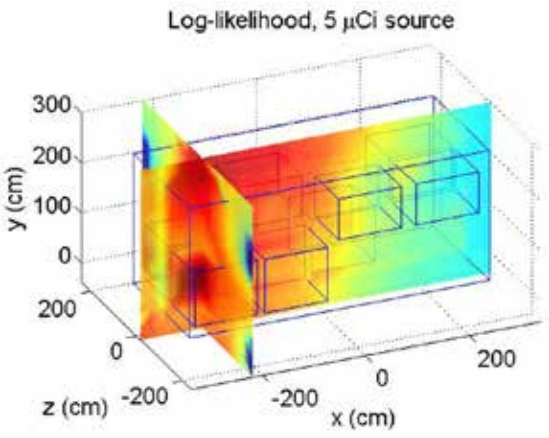
At present, cargo radiography is used to detect contraband but not radiological threats. Previous work that combined radiography and passive sensors for cargo has largely focused on high-level data fusion, where both pieces of information are evaluated separately. For this project, we are investigating the use of radiography for providing information about the problem geometry and focusing on the problem of cargo screening. Specifically, we are developing an inverse algorithm for describing a possible radioactive source in cargo using both radiographic and passive portal monitoring data. Both these technologies are currently deployed for cargo screening and are widely acknowledged to provide complimentary information: the method developed here may provide a quantitative, cost effective way to harness this information. We are also building off of a previous effort that developed a source localization algorithm for combining multiple passive sensors, given a particular estimate of attenuating or scattering materials.

Our research covers three key areas: developing a method for making a 3D geometry estimate from a radiograph and evaluating the impact of this approximation; investigating low-dose radiography with poor spatial resolution to provide this kind of shielding geometry estimate; and evaluating quantitatively algorithm performance compared to existing approaches. These approaches will allow us to evaluate feasibility and its sensitivity to the quality of the geometry estimate and lay the groundwork for future injection studies of the method.

In FY 2012, we established the basic approach for the likelihood calculation. A simulated test case was constructed consisting of boxes of mulch in a 20-ft cargo container, asymmetrically distributed. Initial calculations were performed, demonstrating the ability to localize a source that improves as source strength is increased (decreasing statistical noise). A method was developed for incorporating radiographic information by coarsening the image to ~20 cm pixels and backprojecting each pixel toward the source. With an estimated composition, the radiograph is used to assign a density to each region. Finally, a study was made of two low-dose radiography scenarios: dose sufficiently low for a driver passing through a portal 10 times/day for a year, and the ultra-low case where a worker stands across the lane from the beam each workday for a year. Radiographic images were simulated, with pixel sizes ranging from 3–20 cm for these dose scenarios.

Work in FY 2013 assessed the algorithm's ability to locate a source correctly and estimate its strength over multiple statistical realizations at a range of strengths for a source hidden randomly in the container. This test was performed assuming perfect knowledge of the container, a geometric estimate based on a radiograph, a homogeneous estimate of the container, and no estimate of the container. Other calculations were performed to account for radiation scatter from the hidden source, scatter in the radiographic image, and the effect of assumptions about cargo composition. Calculations of detection based on identified source strength were compared with those based on gross counts. Finally, a series of experimental measurements were performed using commercial systems that coupled radiography and passive measurement.

We have found that the methods developed led to enhanced localization, with increasing benefits as source strength is increased. The ability to localize the source is greatly improved by estimating the radiation transport environment, even in a coarse fashion: 20 cm pixels from a radiograph or even a homogenized view of the container reduce localization error by a factor of two over a calculation without transport correction in a low-noise case. The ability to estimate source strength is likewise improved by nearly any estimate of the transport environment. These results indicate the possibility for improved detection of radiological material in a complex shielding environment.



Log-likelihood map for a 5 µCi source placed in a cargo container containing boxes of mulch. Filled boxes are indicated with blue squares. The true source location is in the dark red region, shielded between two mulch boxes.

PN12030/2431

Three Dimensional Neutronics Analysis Capabilities for Nuclear Archeology Applications

Christopher J. Gesh

Nuclear archeology is a vital capability that the United States needs for many missions such as arms control, nonproliferation verification, and the commercial nuclear power industry. While PNNL is the leader in developing nuclear archeology methods, significant work is required in maintaining a fully operational and robust national laboratory capability.

Nuclear archeology is the measurement based study of nuclear facilities to determine historical information and past operating history to confirm a facility operator's declaration independently. One of the key nuclear archeology techniques used at PNNL is the isotope ratio method (IRM), which estimates the energy or plutonium production in a fission reactor by measuring stable isotope ratios in non-fuel reactor components and relating those ratios to neutron fluence at the sample locations. Based on the fluence and sample location, reactor numerical models are used to calculate the integrated energy or plutonium production of the reactor. Increasing the fidelity of the neutronics analysis allows for more accurate assessments of the reactor operation and use of samples from a wider variety of locations.

This work establishes baseline computational models, methods and capabilities for the current 3D transport codes required to provide timely response to program requests and to establish reasonable expectations from the process. Though the computational approaches are well established, applying them to complex, full-core, high-fidelity depletion problems is challenging and continues to advance with improved computing hardware. Maintaining a core staff with a competitive skill set in the rapidly advancing nuclear analysis field requires regular exercise and training of new staff. A rapid response is also important, which means that personnel need to have relatively recent experience with the tools.

During FY 2012 and FY 2013, multiple analytical approaches were studied to identify the strengths and limitations of each. The computer codes evaluated were 3DB (diffusion theory, current state-of-the-art), MCNP6 (Monte Carlo transport) and Attila (discrete ordinates transport). MCNP6 models of a Magnox reactor and the Hanford B-reactor were used to characterize sensitivities of the results to the parameters that would be useful in quantifying uncertainties in this type of calculation.

Although a MAGNOX core model was successfully imported into Attila, where the mesh was generated and materials assigned, developing a suitable burnup cross-section library was more problematic than originally expected. The difficulties in cross-sections and convergence were such that final Attila results were not obtained.

Measured fuel isotopics as a function of burnup for B-Reactor were used to validate MCNP6 modeling. Above 600 MWd/t, the difference between MCNP6 and empirical production rates drops under 5% for total plutonium production per ton of uranium fuel and the percentage of ^{240}Pu agree to less than 3% relative difference. The MCNP evaluations also characterized the differences in production rates based on core location.

MCNP and Attila models were created and used to analyze a specific PWR reactor benchmark problem. These models illustrate the complex modeling and acceleration techniques required to obtain reasonable results for this type of variable fixed source deep penetration transport problem. The average MCNP calculated neutron flux spectra at the capsule locations agreed within a few percent of the benchmark results.

The methodology was developed for using MCNP to support a nuclear archeology application for a small research reactor where only limited details were available on historical core configurations, dimensions, and materials. Detailed 3D MCNP models of a research reactor in Uzbekistan were developed to represent its major operational periods. These models were validated by comparison to limited published measurements and parameters. Techniques for characterizing the azimuthal and axial variations in the neutron environment outside the core as well as the sensitivity of these variations to changes in core configuration were developed for subsequent work in the selection of IRM sample locations in the reactor core barrel. Project samples were collected from the core barrel, and follow-on nuclear archeology work will interpret the sample data using the calculations.

A paper on IRM neutronics methods is in progress for the journal *Science and Global Security*. An internal report on MCNP techniques and results was also developed for future reference and benchmarking. It is expected that the IRT reactor modeling work for IRM will be presented at the 2014 TRTR Conference. In addition, the graphite reactor work and PWR fluence benchmark comparison are expected to be published in *ANS Transactions*. Finally, an invited presentation on the use of measured IRM data to support nuclear cross-section improvement will be made at the beginning of FY 2014.

PN12076/2477

Ultra-Low-Background Gas Measurement: Building an Advanced Capability

Allen Seifert

We are developing an ultra-low-background (ULB) gas measurement capability available to basic science and emerging application areas, leveraging our new facilities and achieving scientific competitiveness on an international level. This work builds capability for diverse new applications, resolves unique measurement challenges, and will result in high-impact demonstration measurements.

Extremely sensitive nuclear measurements are exploited in a number of scientific disciplines. Measurements of low levels of radionuclides in the environment allow their use as tracers and for age dating. Examples include studies of fate and transport of dust (^{7}Be) in the atmosphere, deep-ocean mixing (^{39}Ar), estuary and deep ocean sedimentation (^{32}Si), soil erosion (^{7}Be), and studies of the mechanisms and pathways for carbon fixation by soil biota and residence times for organic and inorganic carbon (^{3}H, ^{14}C). Low-level measurements are also useful in the verification of nuclear treaties; e.g., detecting ^{37}Ar produced by an underground nuclear test. Across all of these areas, the need for greater sensitivity is paramount.

ULB proportional counters have been used for decades for samples that can be prepared as gases; this method is particularly sensitive to low-energy beta and electron emitting radionuclides. Recently, PNNL developed a design with improved radiopurity materials, and this project moves forward from that foundation. High-sensitivity measurement methods will be developed for specific isotopes, and the detector design will be enhanced to allow greater sensitivity. Analogously, this project will develop the gas handling and quantification tools necessary to work with a variety of input gases derived from environmental samples. The suite of ULB proportional counters necessary for these high-sensitivity demonstration measurements will also be constructed, along with design improvements incorporated as they are developed. The ultimate goal is the development of an enduring ULB capability for the measurement of environmental samples that will position PNNL to compete with other recognized leaders around the world.

For this project, we placed focus in FY 2012 on developing hardware and software tools that would enable high sensitivity demonstration measurements. Design and construction began on a general-purpose gas-handling system intended to handle a broad range of sample types, providing quantitative gas mixing techniques for precise count gas formulation and pressure while maintaining sample integrity. Additionally, five ULB detectors were constructed during FY 2012 to enable the innovative new measurements, with an emphasis placed on performance at gas pressures and gas gains higher than previously employed. Finally, a modeling effort was initiated using Ansoft's Maxwell3D finite-element software and the Garfield gas detector simulation code to investigate pulse formation and detector end effects.

In FY 2013, the basic gas handling system developed in the previous year was expanded, tailoring the capabilities toward the newly identified analytes of interest. The carrier gas manifold capacity was greatly expanded to allow for the creation of more diverse count gases. Components for precision temperature and pressure measurements were added to improve gas quantification, and new independent load lines for both large- and small-volume samples were added. Coupled with significant sample purification and process improvements, all of these modifications, further enhance PNNL's capabilities.

With the hardware built in the first year, the ULB detector work focused this year on optimizing detector operating parameters to enable the quantitative measurement of the analytes of interest. Parametric studies of varying gas blends were used to identify ideal count gases, load pressures, and operating voltages. Improvements were made to detector calibration methods and analysis software tools. Further refinement of the modeling tools developed in FY 2012 led this year to significant insights into the internal electro-static and gas-physics performance, enabling improvements to the detector design and sensitivity.

New activities aimed at further enhancing PNNL's ULB proportional counter capabilities were incorporated into this project for FY 2013. A new low-noise, low-background proportional counter preamplifier was developed. This hardware is designed to operate at higher applied voltages than standard commercial units while meeting the necessary low-background requirements. Additionally, novel uses of the ULB proportional counter were explored, utilizing ^{3}He to enable low-background neutron detection.

The hardware and software improvements developed this past year have each contributed to a series of key radiotracer measurements of ^{39}Ar, ^{37}Ar, and paired ^{3}H/^{14}C that demonstrate proportional counter-based sensitivities that position PNNL to compete internationally with recognized world leaders in this field.

PN12078/2479

Uranium Enrichment Facility Signature Exploitation

Bruce D. Reid

This project leverages PNNL's surface science expertise to quantify material production at a uranium enrichment facility.

Currently, there are a number of countries with uranium enrichment technology that are not signatories to the Nuclear Non-Proliferation Treaty or have dubious nonproliferation credentials. Though a robust measurement technique for determining throughput and operational history of a uranium centrifuge enrichment facility does not currently exist, this technology could have valuable application for resolving international concerns associated with these facilities. The nonproliferation community has recognized that an ability to assess past enrichment facility operations is a "holy grail" of the nonproliferation verification challenge. It has been recognized previously that it may be possible to gain some understanding of the throughput and enrichment of the enrichment plants through examination of facility piping. To this end, our project will maximize signatures inside of a centrifuge and use PNNL's ultra-trace analysis capability to provide a unique method for estimating past enrichment plant operations. We will furnish a scientific basis for assessing the ability of various archeological techniques to predict accurately the throughput and enrichment level of uranium processed in the facility.

In FY 2012, significant time was dedicated to developing necessary collaborations with Oak Ridge National Laboratory and the University of Virginia to accomplish future objectives. Collaboration with the UK Department of Energy and Climate Change was also established to obtain the necessary materials. These collaborations have been encouraging, as the project received uranium enrichment facility hardware in late 2013. For FY 2014, sample preparation, analysis, and interpretation will be completed.

Initial modeling completed in FY 2012 estimated the depths for deposited decay products within the centrifuge walls. Additional modeling was completed in FY 2013 to estimate more accurately the growth rate of the UF_6 corrosion/passivation layer on the surface of centrifuge for all materials of interest. Having a good prediction of the corrosion/passivation layer growth rate will help with interpretation of the surface analysis. By estimating the growth rate, we can determine the length of the operation by measuring the thickness of the corrosion layer. For FY 2014, we will section a sample from the enrichment facility hardware to determine corrosion layer thickness for validation of the completed modeling. Validation of this analysis method will provide information pertaining to the length of time that the equipment was exposed to UF_6.

In FY 2012 and FY 2013, coupon samples of various materials of interest were fabricated to prototypic conditions and exposed to various enrichments of UF_6 for surface and chemical analysis. Initial measurements were completed with promising results. The current methodology for the combination of SEM, XPS, 4F-SIMS, and TIMS provided corrosion/passivation isotopic ratios, elemental ratios, and quantity of uranium. These samples offered the ability to develop a methodology that can be applied to actual hardware, which enables us to determine which measurement technique or combinations of techniques offer the most optimal characteristics of the uranium enrichment facility.

This year, additional testing was completed to determine the applicability of this method for samples exposed to multiple enrichments. Instead of TOF-SIMS, 4F-SIMS was used for this analysis. Based on coupon sample measurements, it was determined that the change in uranium enrichment measured as a function of depth in the corrosion layer is observable using 4F-SIMS. From these measurements, a method was developed for obtaining time-averaged enrichment of UF_6 and the temporal variation in enrichment campaigns. A journal article is underway to report the results of this work. Additional 4F-SIMS measurements are planned during FY 2014 to determine if this measurement technique can be used for formulating the elemental ratio of decay products in the rotor wall and corrosion layer.

TIMS analysis of Pa-231 and Th-230 content has been identified as a possible method for measuring a signature that can be used to estimate the cumulative mass throughput of a centrifuge. Procedures for completing the chemical etching of a surface and sample preparation required for TIMS analysis are currently underway. Efforts have been made to locate a UF_6 cylinder that has been exposed to UF_6 for a substantial period of time to validate the proposed analysis method. During FY 2014, initial TIMS measurements are planned to confirm the applicability of this measurement technique for determining the cumulative number of Pa-231 and Th-230 decay products.

Measurement	Confirm/refute consistency of
Averaged uranium isotopics in corrosion layer*	Time-averaged enrichment of UF_6 (nonlinear corrosion rate)
Depth profile of uranium isotopics in corrosion layer*	Temporal variation in enrichment campaigns
Corrosion layer thickness	Estimate exposure time to UF_6
Cumulative number of decay products (Pa & Th atoms)**	Cumulative mass throughput of centrifuge
Elemental ratio of decay products in rotor wall and corrosion layer**	Time-averaged enrichment of UF_6
Depth profile of implanted decay products in rotor wall**	Temporal variation in enrichment campaigns

** Proof-of-principle measurements of Pa & Th cannot be made without samples that have been exposed to uranium for a significant length of time
* Lab scale experiments have been focused on validating the accuracy and repeatability of these measurements

Uranium measurement type and corresponding time-averaged enrichment or temporal variation.

PN12080/2481

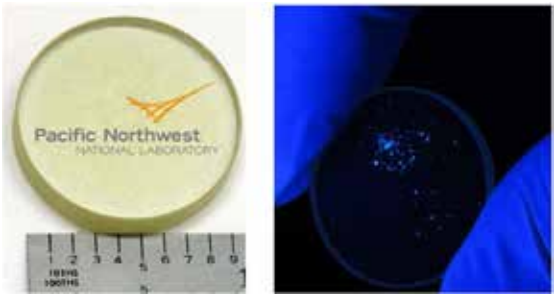
ZnS Scintillators

Mary Bliss

Our objective is to characterize optically active defects involved in zinc sulfide (ZnS) scintillation and to determine their spatial distribution and possibly trace chemistry. Optimization of this material for scintillation would have an immediate impact on radiation detection for medical and other applications.

Traditional ZnS powder scintillator is activated with Ag and Cl and does not transmit scintillation light through the bulk material. ZnS is one of the brightest known scintillators, the commercially available applications of which are limited because of the lack of transparency created by ionizing radiation. Normally, it is used for imaging screens or in conjunction with wavelength shifting plastics. Infrared window grade ZnS is created by chemical vapor deposition (CVD) and is subsequently made transparent to visible light by hot isostatic pressing (HIP). Not all samples scintillate, and several are non-uniform in their response; however, the scintillating portions have similar optical emissions as determined by photoluminescence.

We are investigating new optically active defects in commercially produced infrared window materials for potential use as a scintillating radiation detector. Because of ZnS's low transparency, we are investigating potential scintillation in commercially produced bulk ZnS. Several samples have emission spectra that are significantly different than commercial ZnS powder and indicate a different optically active defect. We recently discovered scintillation signals with alpha particles in commercially produced optically clear ZnS infrared windows.



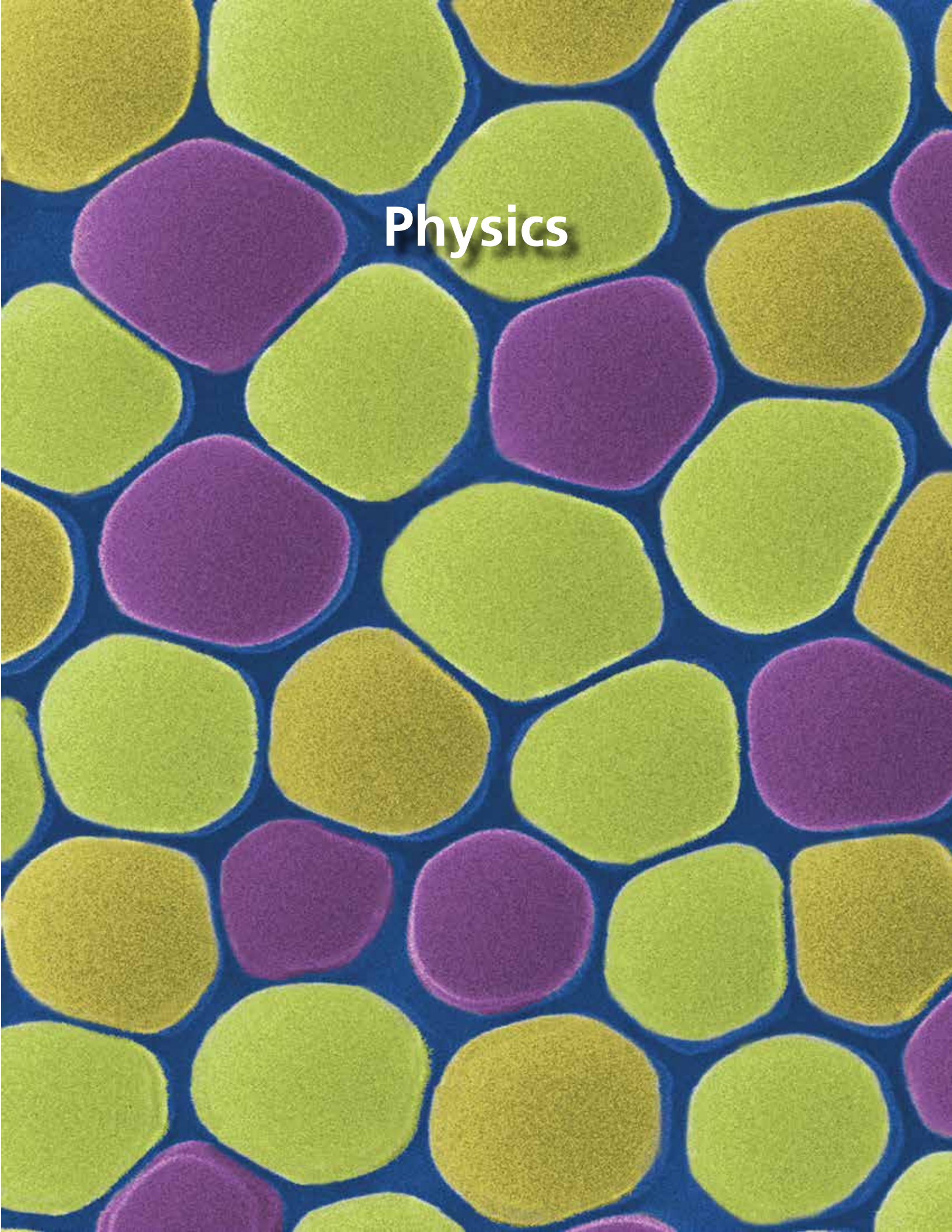
Left: A visible light transparent ZnS infrared window under ambient lighting; Right: The same disk under UV illumination, showing isolated fluorescent regions of unknown origin.

The response is very non-uniform, and we do not have samples thick enough to determine gamma ray response. Instead, we developed a model of the scintillating defect structure based on what we know of the sample processing history and optical emission behavior. The hypothesis is that oxygen replaces sulfur in the crystal lattice and then causes a distortion of the zinc atoms, and/or there is trace copper present from the raw materials. An optically active defect is formed by the association of the copper (or displaced zinc) and the oxygen. More data are needed to verify this hypothesis, which will generate enough material and process information for a patent application for the scintillating material and possible production methods.

In addition, we examined several single crystals and other commercial sources of visible to infrared transparent polycrystalline ZnS for optical emission, fluorescence lifetime, Raman spectra, and X-ray diffraction (XRD). The Raman spectra and XRD data show only subtle differences between the polycrystalline samples. Overall, the samples are cubic, but there are subtle indications of some hexagonal structure that is typical of ZnS. There are differences in texture that correlate with the optical emissions, which were detected by acoustic microscopy. It was expected that the XRD data would support the acoustic microscopy, but it did not. The acoustic microscopy ended up being more sensitive to grain boundaries than XRD. Likewise, optical emission data varied between the polycrystalline samples, indicating that some commercial HIPed ZnS infrared windows have optically active defects different from commercially produced scintillating powders. The ZnS was doped with silver and chlorine and has a strong emission at ~460 nm.

The optical emissions detected in infrared windows vary greatly between samples, and even within a single sample. Some samples have different properties on either large face and others varied within the face of the sample. This is surprising, given that the ZnS is first vapor deposited to form a monolith and then cut for HIPing in reactive atmospheres. The HIP process can take several days and fully recrystallizes the sample into a very fine-grained uniform visible light transparent monolith. A silver doped sample exhibited two emissions: one at the expected ~460 nm and the other at ~530 nm. Alpha particle response correlates to the optical emission data. Scintillation is observed only on samples that respond to UV stimulation. Some chemical variations were detected in the samples by X-ray fluorescence. Scintillation in these materials appears to be caused by trace components and were not intentionally added by the manufacturers.

PN13088/2569


Physics

Absolute Gas Counting Measurement Techniques

Richard M. Williams

This project aims to improve significantly the uncertainty in radiation measurements of low activity gas samples. With these enhancements, scientific endeavors ranging from radioisotope dating of environmental samples to nuclear explosion monitoring will benefit.

Low-level radiation measurements are extremely difficult due to ever-present background sources. In order to suppress this background, PNNL developed a unique method that produces ultra-pure copper, which is used to fabricate detectors that possess almost no radioactive elements. In addition, PNNL recently constructed a shallow underground counting facility that naturally shields measurements from a part of the cosmic ray background. With the further addition of a novel counting method, these existing capabilities combine to yield a truly unique measurement capability.

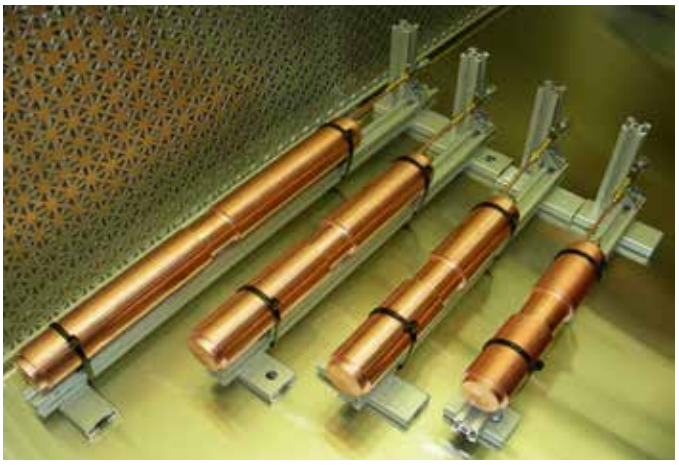
The objective of the project is to develop a measurement technique capable of determining the radioactivity of gas samples without the need for calibration, a process that requires reference standards often not available or difficult and expensive to produce. A handful of laboratories around the world employ the technique being developed, which is referred to as length-compensated proportional counting, but none of these facilities is focused on the same low-level measurements being attempted at PNNL.

During FY 2012, two different sets of proportional counters (detectors) were fabricated at PNNL and passed initial characterization tests. For FY 2013, background measurements were made in PNNL's shallow underground laboratory that showed excellent results, paving the way for future measurements. In addition, an above-ground counting apparatus was designed and construction initiated; this capability will be employed for the characterization of medium-level gases.

The method of length-compensated proportional counting relies on the use of a set of proportional counters that are identical except for their length. The final outcome of the measurement (i.e., the specific activity of the gas sample) is calculated by analyzing the differential count-rates (normalized by the differential gas volumes) from a set of unequal length counters. During FY 2012, two sets of counters were fabricated. The first is a set of detectors were fabricated using ultra-low background (ULB) materials that will be used to make the lowest level underground measurements. In addition, a second set of detectors was fabricated using high-purity, commercial materials; this set is intended for measurements in the above-ground counting system designed and was partially assembled during FY 2013.

In FY 2012, a key background test was performed with the ULB counter-set. The detectors were placed in the shielded counting tomb in the underground laboratory and background counting rates were measured to be roughly 100–200 counts per day over the full energy spectrum. These low count-rates enable measurements of low-level samples with counting times of days to weeks in PNNL's underground facility. In FY 2013, the project focused on developing an advanced counting system with the set of detectors fabricated from conventional materials at its heart. The radiation shield for this system is comprised of both passive and active components. Lead and copper passively stop environmental sources of gamma radiation, while plastic scintillation panels actively detect higher energy cosmic ray events. Finally, a method was developed that will allow the radioactive gas samples to be blended with the special detector counting gas and uniformly transferred into each detector, which completes the measurement suite.

With this phase of the project complete, the newly developed capability will be leveraged by other projects at PNNL. The international radiation metrology community will be engaged, and collaborative measurement opportunities will be explored to validate the capability that PNNL has developed for basic science applications and nuclear non-proliferation treaty monitoring research and development.



PNNL-fabricated set of unequal length proportional counters for use in making absolute specific activity measurements of low-level radioactive gas samples.

PN12002/2403

Anthropogenic Uranium Detection with X-ray Microscopy

Andrew M. Duffin

To support nuclear materials analysis, this project is developing new tools and techniques for analyzing uranium chemically and isotopically on the nanometer spatial scale.

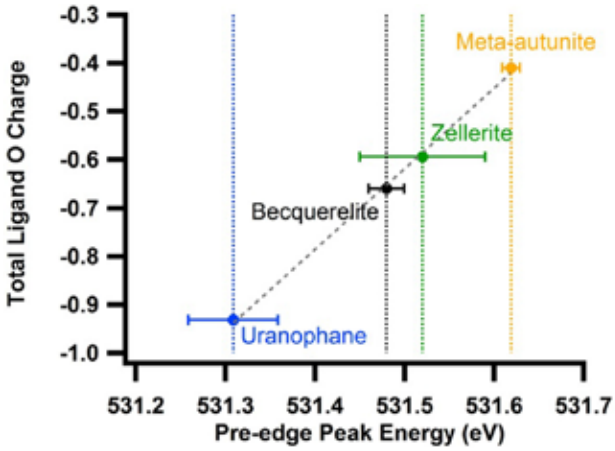
Each year, a considerable amount of research is conducted to detect and analyze anthropogenic uranium, often in a natural uranium background. Most of these studies ignore the chemical form of the sample and rely on mass spectroscopy to gain isotopic information. Unfortunately, this oversight often leads to wasted efforts analyzing natural uranium minerals. Recently, researchers at synchrotron facilities have developed and employed sophisticated techniques for analyzing samples with X-rays. These experiments yield valuable information about electronic structure and chemical bonding.

The objective of this research is to determine the utility of X-ray microscopy for the enhanced analysis of anthropogenic uranium. We are employing advanced X-ray tools developed at synchrotron light sources for identification of the different anthropogenic and natural chemical forms of uranium. We will analyze relevant natural and industrial uranium compounds for fingerprinting purposes. The fact that synchrotron light can be focused down to give nanometer spatial resolution allows for uranium analysis on diminutive samples (those with low detection limits). Given the proper signatures, spatially resolved X-ray absorption spectroscopy could efficiently and non-destructively screen samples for anthropogenic uranium, allowing for subsequent isotopic determination. The results of our research will be of great interest to DOE and could fundamentally improve anthropogenic uranium analysis.

The first year of this project focused on obtaining preliminary X-ray absorption data on a test set of uranium minerals. In addition, considerable effort was spent on developing methods for preparing the anthropogenic uranium samples. Specifically, we collected the oxygen K-edge X-ray absorption spectra for four uranium minerals: uranophane, meta-autunite, zellerite, and becquerelite. We purposefully restricted our initial efforts to collecting spectra at the oxygen K-edge, as this spectral region promises to reveal the most chemical information. Each of the minerals listed above as well as most anthropogenic uranium compounds contain a uranyl species (UO_2^+) surrounded by various ligands. Transition of electrons from the oxygen 1s orbital to the π* uranyl molecular orbital results in a pre-edge feature on the oxygen X-ray absorption spectrum. The specific energy of this transition is modified by the chemical environment (i.e., the coordinating ligands). We found that the pre-edge peak position does indeed correlate with the chemical environment or the uranyl species. That is, oxygen K-edge spectra exhibit promise to differentiate various chemical forms of uranium. Experiments are currently underway that will expand the spectral data library from the four initial minerals to more than 20 key uranium compounds.

The current experiments are developing a spectral library for future reference, and these experiments have been conducted on bulk samples. Future experiments will take uranium compound identification down to the nanometer spatial scale. Scanning transmission X-ray microscopy (STXM) has nanometer spatial resolution with the ability to obtain complete spectral information at each pixel. As a result, it is an excellent technique for uranium compound identification. With this goal in mind, a significant effort was made to develop techniques and expertise for anthropogenic nanoparticle production. These efforts have involved collaboration with Oak Ridge National Laboratory has and resulted in nanoparticle anthropogenic samples that are queued for STXM analysis.

During FY 2014, we will initiate the production of surrogate nanomaterials for analytical performance measurements. In doing so, we will commence measurements to determine detection limits and general analytical performance using these surrogate materials. Finally, we will transition to vacuum detection techniques for a comparison with the analytical performance measurements.



The nearly perfect correlation between the charge of the coordinating ligands surrounding a uranyl ion and the corresponding uranyl oxygen X-ray absorption pre-edge peak position.

PN13027/2508

Highly Efficient and Cost Effective Gamma Detection Arrays for FRIB

Luke E. Erikson

We participated in an accelerator-based experiment (which included the Multi-sensor Airborne Radiation Survey [MARS] array) with the overall goal of developing capabilities in the area of nuclear astrophysics.

The scientific motivation for studying the $^{34}S(\alpha,\gamma)^{38}Ar$ reaction is well documented, as nuclear reactions on ^{34}S play a role in a wide variety of astrophysical scenarios. For example, classical novae powered by thermonuclear runaway driven by accretion onto white dwarfs likely play a role in galactic nucleosynthesis because they inject material into the interstellar medium. In previous research about the study of intermediate-mass nucleosynthesis in ONe novae, the $^{34}S(p, \gamma)$ reaction is noted as having an impact on the final isotopic abundances. The focus was to determine the most important reactions in the Si-Ca mass region. Proton capture on ^{34}S has an impact on the nucleosynthesis of heavier masses in this region, most notably ^{35}Cl and ^{36}Ar. In some astrophysical scenarios (such as a 1.3 solar mass white dwarf experiencing an ONe novae), the main path actually runs directly through the $^{34}S(p, \gamma)$ channel so that the masses in the S-Ca region are directly dependent on this reaction. The typical Gamow window for ONe novae is in the range of 150 to 450 keV, and there exists little experimental data for the reaction in question. This requires models to use a statistical approach in estimating the reaction rate, as was done for the $^{34}S(p, \gamma)$ reaction.

Further, it has been shown in a recent sensitivity study done for type Ia supernovae that the final abundances of some nuclei in this mass region are somewhat sensitive to the (α, γ) reaction rates on ^{34}S. These reaction rates are important for a similar temperature range as the explosive oxygen burning stage of a massive star's life. In addition, the $^{34}S(p, \gamma)$ reaction is of great importance for calculations of final abundances in ONe novae, which occur around temperatures 0.3 GK corresponding to a Gamow peak at 350 keV.

Astrophysical reaction rates are primarily dominated by the presence of resonances with the Gamow windows. Often from a lack of experimental data, statistical modeling is used instead for calculations of reaction reaction rates. However, statistical modeling can provide a poor prediction of the true cross sections in regions where isolated resonances exist. A recent study showed that there is often an order of magnitude discrepancy or more between statistical calculations and experimental determinations of reaction rates. Clearly, experimental determinations of reaction rates are preferable to yield accurate models.

Originally, this project designed a new detector array that would be used in the next generation of nuclear astrophysics experiments to increase understanding of the formation and evolution of the universe. However, the project changed emphasis, with the primary goal during FY 2013 shifting to the successful participation in an experiment to measure resonance strengths of the $^{34}S(\alpha, \gamma)^{38}Ar$ radiative capture reaction. This measurement was performed at the Detector of Recoils and Gammas Of Nuclear Reactions (DRAGON) located at the Tri-University Meson Facility (TRIUMF) in Vancouver, British Columbia, Canada. The work was done in collaboration with a number of other institutions, including the Colorado School of Mines, McMaster University, and the University of York.

The intention was to include MARS array with the experiment to demonstrate the value of PNNL's approach to integrate HPGe detectors into beamline experiments. However, MARS has low efficiency compared with the existing BGO array at DRAGON, which exists in a close packed arrangement on each side of gas jet target. With pumps located both above and below the beamline, this would have left MARS with a significantly distant position, further decreasing the detection efficiency. For these reasons, the MARS array was not included.

The experiment was conducted during the middle two weeks of August, during which one of our team members participated in experimental setup, data acquisition, and preliminary analysis. Preliminary results include the direct measurement of previously unmeasured resonances as well as the accurate measurement of known resonances. An eventual publication – the second for this project – on this reaction is expected and will appear in *Physical Review Letters* or equivalent.

PN12038/2439

Low Energy Threshold Germanium Detectors and Science

John L. Orrell

We developed a novel maximum likelihood analysis method for testing experimental data for the presence of a time varying dark matter signal hypothesis.

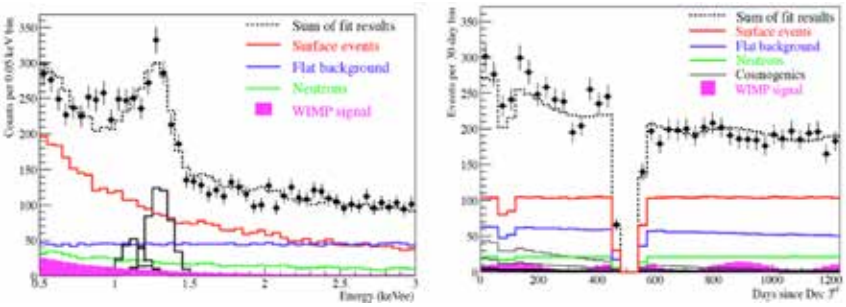
Unambiguous direct detection of cosmological dark matter particles is perhaps the single most pressing particle astrophysics measurement sought by cosmologists and theoretical particle physicists alike. Dark matter is well known by its gravitational effects observed throughout the universe, despite being entirely invisible to our telescopes measuring the electromagnetic spectrum from radio frequencies, through the visible spectrum, and up to gamma rays. The gravitational effects appear in deviations from uniform intensity of the cosmic microwave background radiation, galactic clustering, gravitational lensing, and the rotation of galaxies. The leading theoretical candidate for these gravitational effects is the presence of a weakly interacting massive particle (WIMP), whose gravitational influence is ubiquitous but otherwise does not participate in the electromagnetic or strong interactions of the Standard Model of particle physics.

We have been developing and fabricating a low-capacitance, low-background cryostat for germanium detectors to be used in ultra-sensitive radiation measurements made in ultra low-background environments at underground locations. Although not alone in the search for WIMPs, PNNL has pursued two unique avenues of development in methods for directly detecting dark matter particle using deep underground radiation detection systems. For hardware development, we completed a prototype low-background, low-capacitance cryostat for operating high purity gamma-ray spectrometers. This effort will allow an increase in size by a factor of three of the germanium crystal used in comparison with previously operated detectors of this same kind. Increasing the target mass directly increases the potential rate of dark matter interactions in the detector.

In addition to the low energy signals, a dark matter signal is expected to vary annually due to the earth's orbit varying with respect to the Milky Way's center. We developed a maximum likelihood analysis technique to search simultaneously for the low-energy and time varying signature of dark matter interactions in deep underground experiments. The software analysis method development is a sufficiently generic enough analysis method that will be applied to a future dark matter experiment.

Left: C4 cryostat design and cut-away view of the interior, showing the low-capacitance mount; Right: Completed prototype cryostat in final assembly in the PNNL underground lab cleanroom.

Next, we will team with the University of Chicago and CANBERRA Industries to install a 1.2 kg germanium detector in the cryostat that has demonstrated a factor of two for improvement energy threshold performance from prior detectors of this kind. The signature of dark matter interactions in a deep underground radiation detector is a low-energy nuclear recoil signal. Signals of these kinds are difficult to measure, as the energy deposited by the dark matter particle in comparable to the magnitude of the electronic noise of the system. Thus, a low energy threshold improves the detector's capability to measure dark matter induced nuclear recoil interactions.



Simultaneous energy and time maximum likelihood analysis for dark matter detection.

PN12045/2446

Resolving the Reactor Neutrino Anomaly by Precision Beta Spectrometry

Kim A. Burns

We are developing an experimental technique that results in precision measurements of the beta energy spectrum from neutron-induced fission using the Project X Injector Experiment (PXIE) 30 MeV proton linear accelerator as a neutron source.

Neutrino experiments at nuclear reactors are vital to the study of neutrino oscillations. The observed antineutrino rates at reactors are typically lower than model expectations. This observed deficit is called the "reactor neutrino anomaly." A new understanding of neutrino physics may be required to explain this deficit, though model estimation uncertainties may also play a role in the apparent discrepancy. Each fission event produces fission products that decay and emit electrons (beta particles) and antineutrinos, and precise measurement of the beta energy spectrum can be used to generate the associated neutrino spectrum.

During the course of this project, the team will research, design and test components for a system that promises reduced uncertainties for measured data to support resolution of the "reactor neutrino anomaly." During the first quarter of FY 2013, a thorough literature search was completed. Previous measurements and analyses were reviewed to identify primary and possibly secondary contributors to the uncertainty in interpreting reactor neutrino measurements. An initial statistical analysis of the previous experimental setup and spectral unfolding methodology was completed. Based on this analysis, it was determined that the sources of uncertainty that have the highest impact on the data resulted from the deconvolution of the beta spectrum from the spectrometer and the unfolding of the antineutrino spectra from the beta spectrum. This analysis has focused a majority of the work during the first fiscal year on identifying an appropriate beta spectrometer and completing extensive testing of the proposed method for completing the antineutrino unfolding algorithm.

A major advantage of an accelerator neutron source over a neutron beam from a thermal reactor is that the fast neutrons can be slowed down or tailored to approximate various power reactor spectra. To demonstrate feasibility of completing a fission based experiment at PXIE, an accelerator beam target will be designed. The initial stage of target design was competed using Monte Carlo N Particle (MCNP) modeling. Initial results indicate that it is possible to generate an appropriate neutron spectrum and neutron flux to support the experiment. A variety of possible target materials have been identified based on spectrum and flux requirements. The selection of the final target material will be based on fabricability, mechanical requirements, thermal analysis and facility requirements. The target material selection will be completed during FY 2014.

Fission foil targets containing a small (mg) isotopic fission target of the primary reactor isotopes ^{235}U, ^{238}U, ^{239}Pu, or ^{241}Pu will be developed that optimize the neutron field and contains the fission products but allows the delayed fission betas to be released so they can be collected in the beta spectrometer. Analysis is underway to determine the feasibility of using fission foils produced at PNNL based on the methodology developed at Oregon State University. Current analysis indicates a high probability of success using these fission foils, but a more detailed analysis is required to determine the fission product retention rate versus the fission beta release rate from the foils. Completion of the fission beta transport model through the fission foil and cover is required as input to the analysis of the extraction tube. This analysis will determine the maximum length of the extraction tube based on the number of betas that will make it from the fission foil to the detector entrance through the extraction tube, which will determine the beta spectrometer efficiency as a function of beta energy.

To reduce the uncertainty associated with the beta spectrum measurement, a beta spectrometer with well characterized uncertainties as a function of energy will be used in the experiment. Sets of wire chamber detectors at the entrance and exit of bending magnets is currently being evaluated for this application. If the wire chamber detector is determined to suitable, two magnets will likely be required, one optimized for lower energies and another for higher energies. A methodology is also being developed for track finding and track fitting, using Hough filter techniques, to identify the beta energy from the detected track. The track finding and fitting methodology has been focused on basic feasibility of the technique for this application, not detector design.

During FY 2013, a white paper on predicting reactor antineutrino emissions was produced, and the work was presented at a workshop and a conference. Additionally, two journal articles are in progress.

PN13008/2489

Search for New Physics at the Intensity Frontier

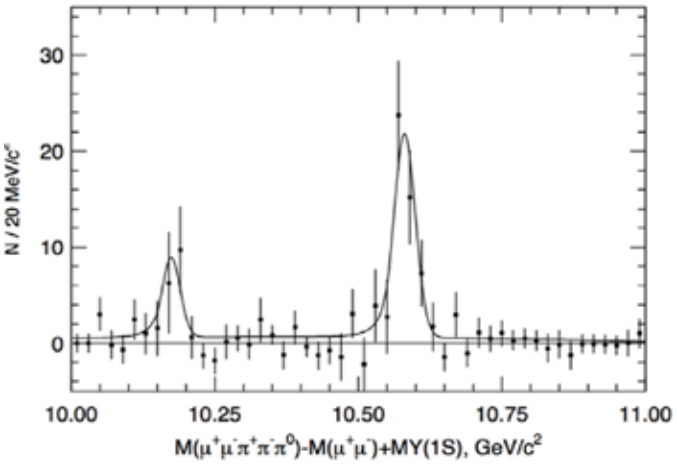
Gocha Tatishvili

The specific aim of this project and research is to search for new physics in the existing Belle data.

Operated at the asymmetric electron-positron collider KEKB from 1999 to 2010, the Belle detector accumulated the world's largest data sample for the study of B mesons (subatomic particles containing one heavy bottom quark and a light anti-matter quark) and Bottomonia (bound states between a heavy bottom quark and a heavy anti-matter bottom quark). Bottomonia states were first observed about 30 years ago. However, there remain areas in which knowledge of these systems needs improvement, including anticipated but unobserved states of matter. Specifically, Bottomonia spectroscopy is generally regarded as a key source of information necessary for the development of the understanding of the quantum theory describing the strong interaction that holds atomic nuclei together.

This project analyzes the world's largest sample of heavy quark decays recorded by the Belle detector at the KEK Laboratory in Tsukuba, Japan. The goals of this work are to discover and characterize the properties of exotic bound states of matter and anti-matter. This project will also lay the foundation for leadership in the physics program at the upgraded facility Belle II/SuperKEKB.

The discovery and characterization of the properties of exotic states of matter recorded by the Belle detector are the objectives of this project. One specific goal is to refine our understanding by making new and more precise measurements of the parameters of the Bottomonia states, their decays, and their transitions. The expected results of this endeavor provide important benchmarks for a variety of theoretical approaches. This project is seeding the initial collaboration between PNNL and the international Belle research team which, since the start of the 2013 calendar year, consisted of a collaboration of more than 450 physicists from 70 institutions of 20 countries.



Analysis of the Belle data indicated the first-time observation of hadronic transitions $\Upsilon(1D) \rightarrow \eta\Upsilon(1S)$ (left signal) and the former transition $\Upsilon(4S) \rightarrow \eta\Upsilon(1S)$ (right signal).

At the conclusion of FY 2013, this project has had several significant accomplishments. First, we reported the results of the rare hadronic transitions in which the final states $\eta\Upsilon(1S)$ were observed. The results were obtained using a data sample of 382×10^6 $\Upsilon(4S)$ decays collected with the Belle detector at the KEKB asymmetric-energy e+e- collider. We observed two hadronic transitions:

$$\Upsilon(4S) \rightarrow \eta\Upsilon(1S)$$

and

$$\Upsilon(1D) \rightarrow \eta\Upsilon(1S).$$

The latter transition $\Upsilon(1D) \rightarrow \eta\Upsilon(1S)$ was observed for the first time. For the former transition, $\Upsilon(4S) \rightarrow \eta\Upsilon(1S)$, we determined a branching fraction of the following:

$$B = (2.48 \pm 0.38(\text{stat}) \pm 0.10(\text{syst})) \times 10^{-4}$$

The obtained results provided important tests for theoretical approaches to understand quark–antiquark interaction dynamics and the formation of light hadrons. The result of $\Upsilon(4S) \rightarrow \eta\Upsilon(1S)$ branching fraction is in good agreement with the Experiment BaBar result and also indicates that above B/anti-B threshold hadronic transitions are unable to be supported by the conventional hadronic transition theory.

Throughout FY 2013, we were invited to speak at two international conferences and submitted four journal articles, including a publication in the European physics journal Il *Nuovo Cimento*. This project has enabled PNNL to build a reputation for scientific excellence through the discovery and elucidation of Bottomonium states in FY 2013. Next year, the PNNL HEP group focus will shift to the search for new phenomena in B meson decays.

PN12069/2470

Super CDMS Initial Engagement

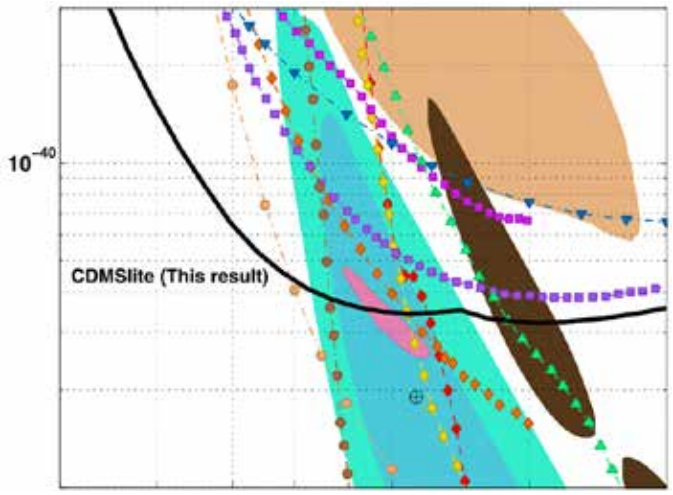
Jeter C. Hall

This work engages PNNL with cutting edge experiments probing the nature of dark matter.

A strophysical observations show that 84% of the matter in the universe is of an unknown, non-baryonic nature. The dark matter problem is a fundamental mystery in particle physics. To this end, the Super Cryogenic Dark Matter Search (SuperCDMS) is an experimental program to search for dark matter using germanium detectors operated at 10s of millikelvin. The SuperCDMS Collaboration has pioneered the direct detection of dark matter field. The basic idea is to shield sensitive nuclear detectors from all known forms of radiation to look for a new form of radiation: weakly interacting massive particles (WIMPs) dark matter. Because they are expected to be weakly interacting and neutral, WIMPs should scatter off nuclei in these detectors, much like neutrons, and can be distinguished from most known radiation, such as gamma rays.

This work is focused on light dark matter or WIMPs below the mass scale of ~10 GeV/c^2. Limits on new charged particles from particle colliders coupled with simplistic assumptions in theoretical phenomenology have led experimentalists to neglect this low mass range. This range is particularly difficult to search with traditional dark matter detectors because the detectors are based on searches for unexplained nuclear recoils. As the WIMP mass becomes much lighter than the detector nuclei, the energy imparted to a detector nucleus by a WIMP scatter is so small it is undetectable for many traditional dark matter experiments. However, a number of recent experimental results and theoretical explorations have rapidly increased interest in light dark matter. This work completes the world's most sensitive search for light dark matter in the mass range of 2–8 GeV/c^2.

We are leading a particular technique within SuperCDMS, the CDMS low ionization threshold experiment (CDMSlite). This experiment involves operating the SuperCDMS detectors with lower thresholds than previously possible with new custom electronics. We proposed this operation mode for SuperCDMS, designed the custom electronics, participated in the commissioning and data taking for this experiment, and finally led the analysis group to a result. Our results were submitted to *Physical Review Letters* and presented at a major conference in late FY 2013. The SuperCDMS Collaboration is pleased with the progress in FY 2013 and is anticipating an extension of this experiment with a longer exposure to investigate these results.



The final limits from the CDMSlite experiment are compared to various regions of interest and to limits from other dark matter experiments. Note that the circles represent experiments based on xenon nuclear targets where the experimental and theoretical systematic errors are significant.

In addition to leading this particular experiment, the SuperCDMS Collaboration is continuing to analyze legacy data and take new data. By engaging the Collaboration, PNNL is supporting these other scientific opportunities through data-taking shifts, collaboration meetings, analysis meetings, and scientific presentation. For example, the Collaboration published two papers in FY 2013 on light dark matter using legacy data taken with the previous CDMS-II experiment. These results were well received by the particle physics community as evidenced by the 73 citations (16 published) as of mid-August 2013.

By engaging scientifically with SuperCDMS, PNNL capabilities have been showcased to the broader dark matter community. The four major WIMP search experiments proposed to DOE have all approached PNNL regarding technical and engineering expertise to assist with their research and development efforts. Additionally, DOE has taken notice and is expecting proposals for additional scientific engagement, and we expect that the Joint Center for Low Background Materials will support the next generation of dark matter searches.

PN13058/2539

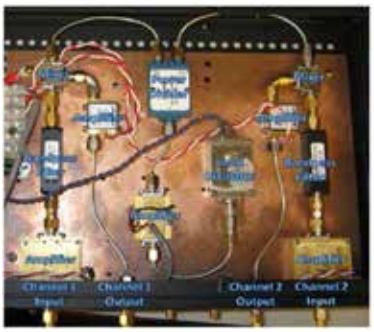
Ultra-Precise Electron Spectroscopy to Measure the Neutrino Mass

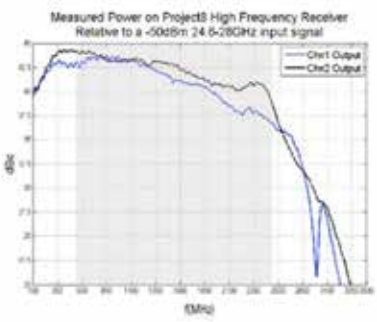
Brent A. VanDevender

This project will develop a fundamentally new spectroscopy technique to be applied in the search for the absolute mass of the neutrino.

Neutrino mass has far-reaching implications from theories beyond the Standard Model of Particle Physics to the evolution of large-scale structures in the universe. As the DOE Office of Science rates measurement of neutrino mass as one of its top priorities in basic nuclear science, we are developing a new technique in this field. Specifically, we are working with the spectroscopy of medium-energy (1–100 keV) electrons based on the detection of cyclotron radiation emitted by magnetically trapped electrons. The energy of a trapped electron is encoded in the radiation frequency, which can be measured with extreme precision. Our method promises energy resolution superior to existing techniques. The goal is to measure the neutrino mass m involved in tritium beta decay by observing a small reduction in electron energy E attributed to neutrino mass by Einstein's relation $E = mc^2$. Experiments based on existing techniques have reached their ultimate neutrino mass sensitivity and may still not be sufficient to weigh the neutrino, which could be up to 100 times lighter still.

For this project, we formed the Project 8 collaboration with the University of Washington (UW) Center for Experimental Nuclear Physics and Astrophysics, the Massachusetts Institute of Technology, the University of California–Santa Barbara, and the California Institute of Technology to develop the technique and use it to measure neutrino mass. Our immediate goal is to demonstrate the ability to detect the cyclotron radiation from individual trapped electrons and reconstruct an energy spectrum with a prototype experiment at UW. The prototype demonstration will be made with ^{83m}Kr, a radioactive gas that emits electrons with similar energies as tritium, but in a much simpler spectrum that will allow us to discern instrumental effects from true spectral features.





The Project 8 High-Frequency Receiver Chain Electronics unit (left) contains a reference oscillator that serves as a master clock for correlating signals on the two channels. Expected signals of ~25 GHz are "mixed" down to 0.5–2.5 GHz (right) in preparation for digitization and storage to disk. Significant redesign by PNNL engineers improved the performance of this critical component of the apparatus.

PNNL plays leading roles in three major aspects of the experiment. Strong radioactive ^{83m}Kr sources manufactured at PNNL are required for the prototype demonstration and will continue to be needed for calibration for the entire life of the Project 8 experiment. The radio- and microwave frequency electronics and receiver structures are designed and tested here. PNNL is now also host to high-performance computing capabilities required both to simulate and to analyze data from the experiment.

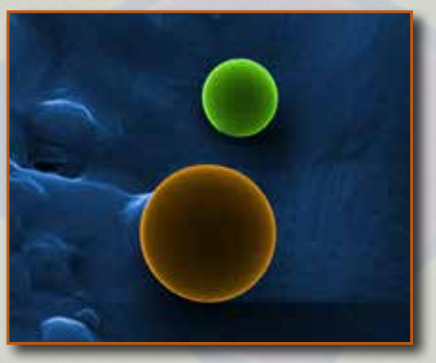
A new 1mCi ^{83m}Kr gas source was manufactured at PNNL this (and last) year and installed into the prototype apparatus at UW-Seattle. A significant new step was the completion of the first cycle by which these strong sources are exchanged and returned to PNNL for safe storage. After a full year, these sources are no longer sufficient for Project 8 due to natural radioactive decay. However, they are still quite hot by normal radiological standards and require appropriate handling and storage. A compact lead shield was also fabricated and installed so that people working in the laboratory around the prototype are not exposed to any significant radiation dose.

A redesign of the high-frequency electronic receiver chain by PNNL engineers has greatly improved the signal path that converts raw electron cyclotron radiation to digitized signal values on disk. Most importantly, a single stable reference oscillator is now used as the master clock for the two-channel system. Previously, each channel had its own oscillator and the two were not phase coherent. Further, they were found to be unstable with frequencies drifting over ranges very large compared to our targeted frequency precision. The new oscillator also has a frequency better matched to specifications of other components in the system. A phase coherent system will improve our overall signal-to-noise ratio. Preparations are under way to characterize the first-stage low-noise amplifiers and demonstrate the feasibility to detect the expected signal power above noise.

PNNL took on new scope in support of Project 8 in FY 2013. Our institutional high-performance computing capabilities are now host for Project 8 data analysis and instrument simulations, and these capabilities are not available at any of our collaborating institutions.

PN12079/2480

ON THE COVER:



Understanding the Origin of the Universe

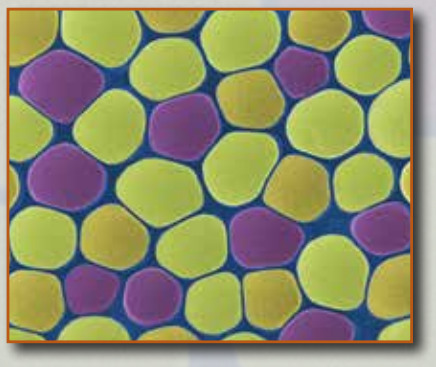
Great discoveries often need great tools. Powerful scientific accelerators, lasers, and electron microscopes rely on photocathode materials. This colorized image, obtained using a helium-ion microscope, shows a used photocathode electrode from the Thomas Jefferson National Accelerator Facility. The photocathodes direct different beams, such as x-rays, at samples. How the beam interacts with the sample tells scientists a lot. Expensive photocathodes degrade with repeated use. Understanding how and why the photocathode material degrades could lead to longer-lasting photocathodes, reduce the down time, and reduce the cost of the instruments. Further, improving photocathode technology will allow scientists to get deeper insight into the particles and forces that build our universe. Research was funded by PNNL's initiative in Chemical Imaging.

More Efficient Air Cooling in Buildings

The cooling of air in buildings accounts for approximately 40 percent of energy use in the United States. A novel and more efficient cooling method is adsorptive chilling utilizing new super-hydrophilic metal-organic frameworks (HMOFs), as seen in this image. PNNL researchers have developed HMOFs with porous frameworks and very high surface areas and other special characteristics. These HMOFs trap high numbers of refrigerant molecules, allowing for efficient adsorptive chilling applications that require little or no electricity. Research was funded by the U.S. Department of Energy Advanced Research Projects Agency-Energy.

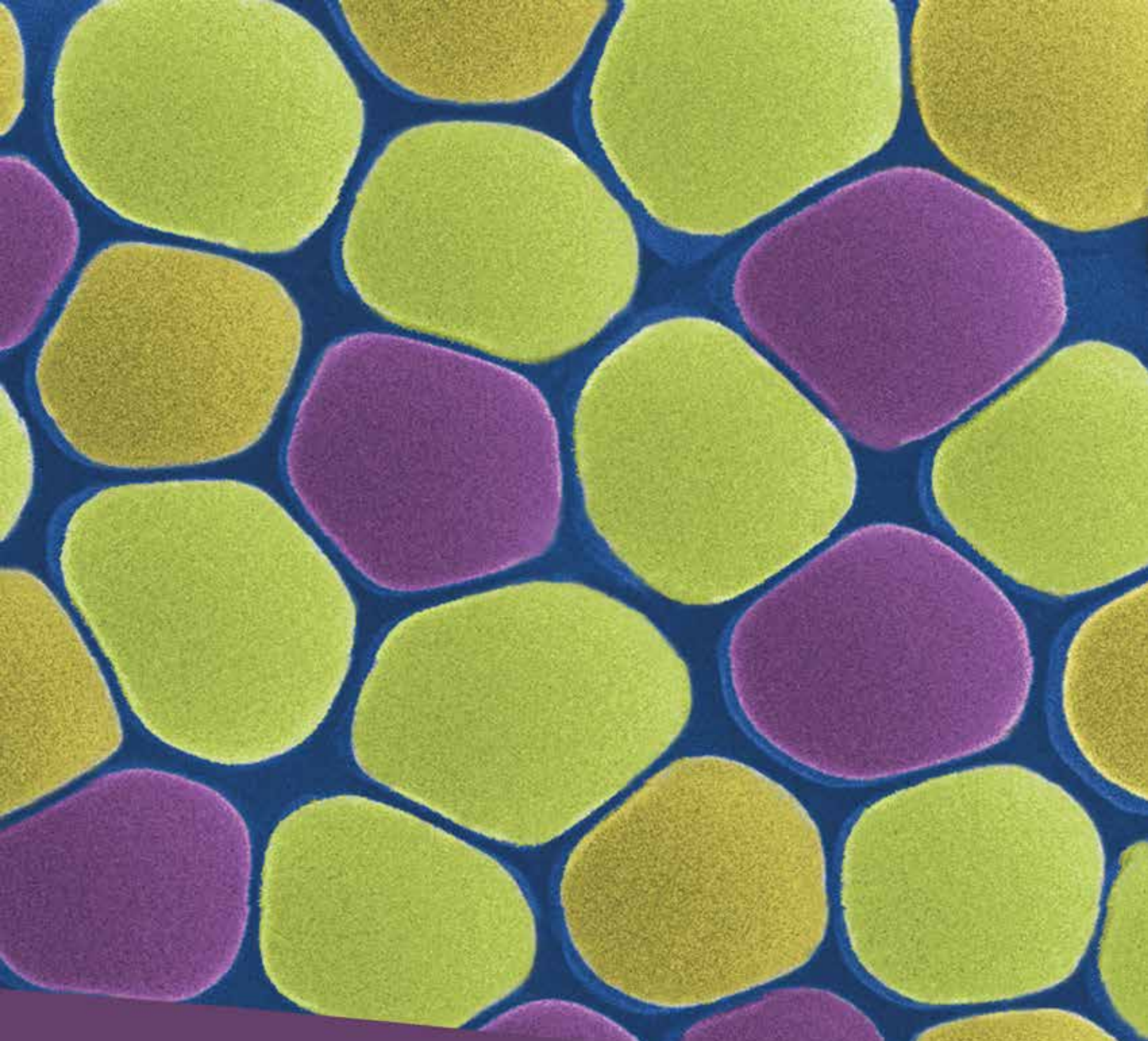
Alamogordo Crystals and Glass Aid in Describing Type and Magnitude of Nuc lear Events

This digital image shows soil particles altered by the world's first nuclear explosion in Alamagordo, New Mexico, in 1945. Using microscopic imaging tools, scientists rapidly identified the microstructure of crystalline and glassy particles altered by the intense heat of an explosion. This work conducted by researchers at PNNL and the University of Notre Dame examines what type of forensic information may be extracted from a post nuclear blast site. Research was funded by the National Security Agency.



Background Photo
Understanding Titanium Dioxide Nanotubes to Create Energy from Renewable Sources

This helium ion microscopic image shows the back side of an array of titanium dioxide (titania) nanotubes grown in an anodization process. Titania is a versatile transition metal oxide that is used in gas sensors, photocatalysis and dye-sensitized solar-cell applications. Helium-ion images of titania nanotubes arrays are being used by Pacific Northwest National Laboratory (PNNL) researchers and collaborators to improve understanding of the performance of this material in dye-sensitized solar cells for converting solar energy into electricity and to help solve the energy crisis. Research was funded by the National Science Foundation.



2013 Annual Report

Laboratory Directed Research & Development

www.pnnl.gov

Pacific Northwest
NATIONAL LABORATORY

Proudly Operated by **Battelle** *Since 1965*

U.S. DEPARTMENT OF
ENERGY

PNNL-23265