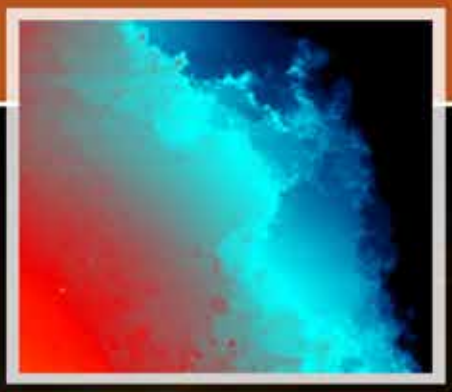
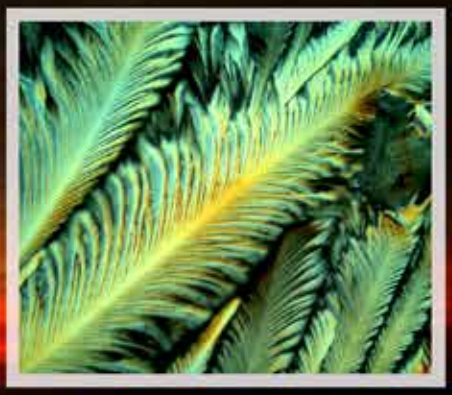
2017 Annual Report

Laboratory Directed Research & Development
at Pacific Northwest National Laboratory





DISCOVERY *in Action*

Pacific Northwest
NATIONAL LABORATORY

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-27379

Laboratory Directed Research and Development Annual Report

Fiscal Year 2017

March 2018

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

Pacific Northwest National Laboratory
Richland, Washington 99352

Contents

Chemistry

Earth and Space Sciences

Energy Supply and Use

Engineering and Manufacturing Processes

Materials Science and Technology

Mathematics and Computing Sciences

Nuclear Science and Engineering

Physics

Laboratory Director's Message

As the nation's premier chemistry, earth sciences, and data analytics laboratory, Pacific Northwest National Laboratory advances scientific discovery, improves energy resiliency, and enhances national security. Our deep technical capabilities enable us to make scientific breakthroughs and deliver technological innovation to address our sponsors' highest priorities and our nation's greatest needs.

Key to advancing the Department of Energy's missions today, as well as tomorrow, is strong stewardship of our scientific and engineering capabilities. As part of our annual strategic planning process, we determine where to invest our discretionary research and development funding with an eye toward four specific outcomes—transformational science and technology, accelerated innovation, new partnerships, and enhanced core capabilities. Each year, we select a portfolio of Laboratory Directed Research and Development projects that explore new concepts aligned with these priorities.

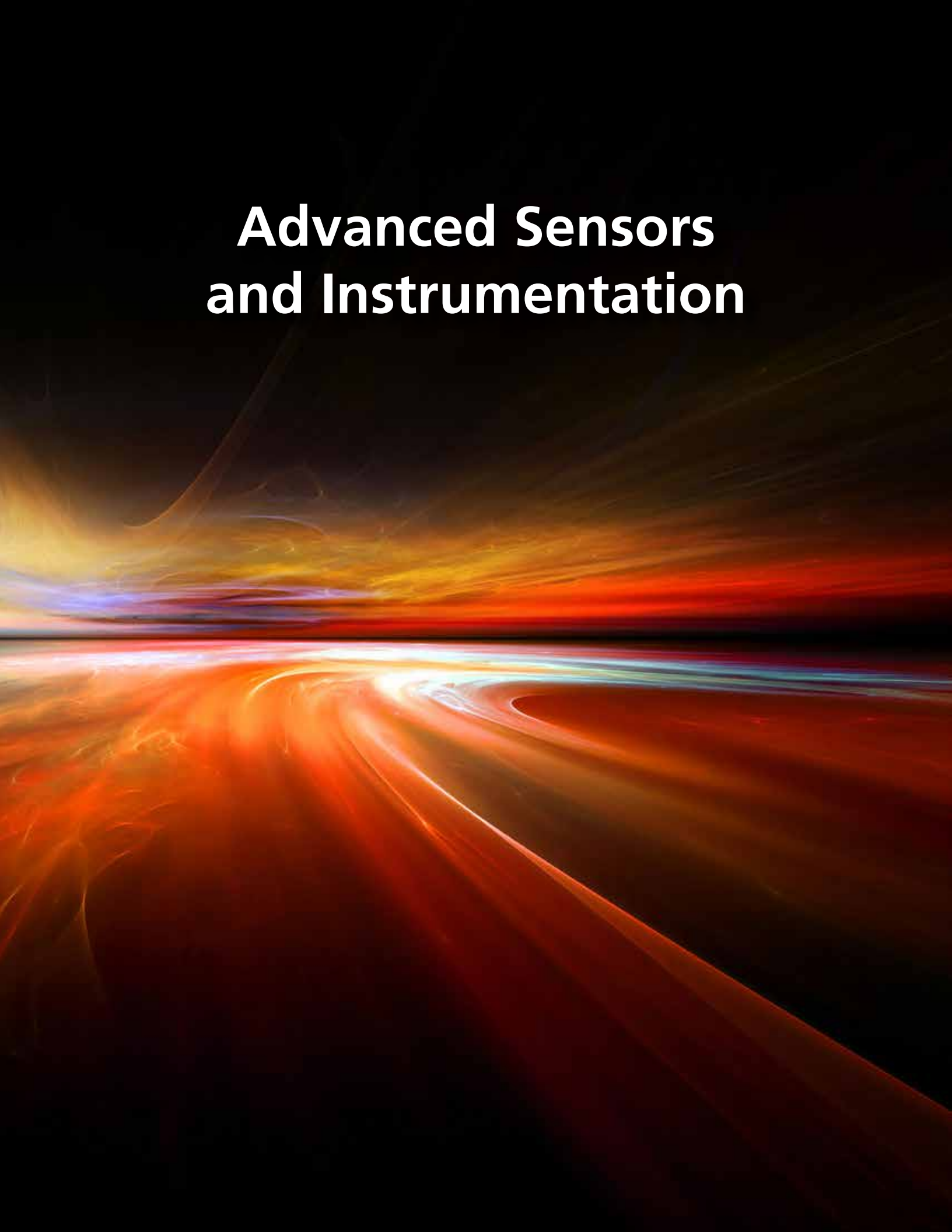
This report summarizes our FY 2017 LDRD program, highlighting more than 200 projects that span fundamental discoveries to applied solutions. We are proud to showcase the creativity of our scientists and engineers and the resulting accomplishments that reflect the impact of PNNL's research.

Thank you for taking the time to learn about the transformational science and technology underway at Pacific Northwest National Laboratory.

Steven F. Ashby

Dr. Steven Ashby
Director, PNNL

Advanced Sensors and Instrumentation



Active Millimeter-Wave Holographic Imaging and Instrumentation for Feature Localization

Doug McMakin

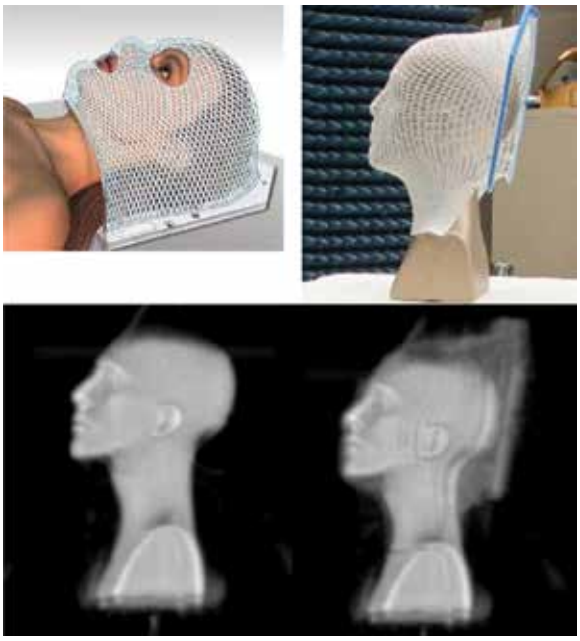
The objective of the project is to determine how precisely millimeter-wave (MMW) imaging can locate and characterize features in a mannequin (surrogate for humans) and whether MMW sensors can be used in managing and controlling moving objects in the form of a radiation gantry. The project will assist in the determination and applicability of MMW imaging for biometric feature identification, three-dimensional subject mapping, and capturing objects in motion under obscurations.

Medical radar sensors have been studied for respiratory gating and tumor tracking to better guide the LINAC (X-ray source) during radiation theory sessions to provide patient tumor motion information. Presently, radar imaging array sensors have not been utilized for this application. Alternatively, laser-based motion tracking technology is used now to provide a precision interface between the patient and the LINAC. However, lasers and Infrared motion tracking technologies are limited because they cannot penetrate through optically opaque patient constraining devices, which are starting to be used more extensively during radiation therapy sessions for patient constraining.

After reviewing our MMW body scanning work on security and biometrics, a Washington University medical physicist in St. Louis reached out to PNNL to determine if the MMW holographic imaging technology could be used for this application to provide better marker-less tumor tracking through patient constraining devices. This work focused on imaging studies to determine what accuracy can be obtained in active MMW holographic imaging in the human patient's bony anatomy or soft tissue by conducting feasibility imaging studies on metalized mannequins constrained by formfitting thermoplastic masks and other low-Z body constraining immobilization devices.

Radiation therapy for cancer patients requires high patient alignment to accurately target tumors and minimize damage to healthy tissue. This alignment technology also requires reproducible positioning from session-to-session to maintain maximum impact during radiation daily therapeutic sessions. Additionally, patient positioning is needed to expedite co-registration with the reference CT volumetric scanner to determine patient setup error, motion under treatment, and inaccuracy in the external surface alignment to be used as a surrogate for internal target position. This work demonstrated that MMW holographic imaging could be used in these applications.

Future work will involve modeling and simulation of MMW arrays in relation to patient placement on the LINAC therapeutic tables and developing real-time imaging techniques such as two-dimensional synthetic aperture radar and interferometric synthetic aperture radar to obtain instantaneous motion tracking.



Holographic radar imaging result of mannequin head with and without constraining mask.

PN17076/2966

A Probe-Based Microtiter Plate Assay for Characterization of Protein Binding Partners of Small Molecules

Aaron Wright

Glass and magnetic beads derivatized with chemical probes enable the rapid identification of protein-small molecule binding partners.

Small molecule-protein interactions are critical to cell communication, nutrient scavenging, inhibition or activation of metabolic pathways, as enzyme cofactors, and numerous other processes. Small molecules are also used to control protein functions relevant to energy and human health, for instance to trigger a biosynthetic pathway that produces hydrocarbons or to inhibit a bacterial pathogen. However, discovery of the interactions occurring between small molecules and proteins in diverse biological systems is not simple. *In silico* predictions are limited to proteins for which we have quality structural information (e.g., crystal structures), which are primarily limited to mammalian systems. Thereby, a novel method for rapidly characterizing small molecule-protein interactions has significant potential to transform our understanding of biological functions.

Human Health. The current state of the art in the development of new drug compounds centers around *in silico* docking studies of a target protein with theoretical compound libraries. These screens result in high-priority drug compound leads that are then synthesized in the lab and tested in binding assays to the target protein. If all is successful up to this point, they will start cell line and animal studies to see if the drug binds to the target protein and reverses whatever the negative cell phenotype is (e.g., cancer). The next step is a long and arduous process in animal studies to determine side effects, often without having any idea of the mechanism for observed side effects.

Energy and the Environment. Within microbes and microbiomes, it is difficult 1) to predict microbes with the potential to salvage small molecules, particularly nutrients such as B vitamins, and 2) to identify small molecule regulatory functions that impact cell physiology and functions. Instead of laborious chemical and computational screens, we present an alternative in which a chemical is derivatized to contain a reactive group that enables attachment between the chemical and a strong interacting protein. The chemical molecule is on a resin, which enables its easy placement in an array format. Analysis of small molecule-protein interactions is performed by mass spectrometry.

We started by making a chemical mimic of the antimicrobial agent, ketoconazole. This drug has recently been suggested as a therapy for tuberculosis treatment. We developed a chemical probe analog of ketoconazole, which we demonstrated had the same impacts on *Mycobacterium tuberculosis H37Rv* cell viability, with a similar minimum inhibitory concentration. We then determined how best to append the drug probe to glass and magnetic resins. We also evaluated an alternate approach in which the chemical probe is applied directly to the microbial cell, and following protein-small molecule interactions and irreversible binding, the complex is brought to the resin for subsequent washes and eventual proteomic analysis by mass spectrometry.

Using our approach, we revealed that the drug's presumed protein target is the primary and dominant target by quantitative proteomics analysis, but several exciting and unanticipated targets were identified. The assay was also used for screening of multiple conditions in a way that far outpaces conventional approaches. We identified many protein targets of the antimicrobial small molecule, both anticipated and unanticipated. We demonstrate that this approach can decrease the cost and timeline for drug development and target identification. Having this information allows decisions regarding dosage levels, combination therapies, and mechanistic understanding of on- or off-target activities of drugs. This information can be translated to improved clinical outcomes via the development of drug candidates with improved selectivity profiles. We are now translating this same approach to understanding how particular metabolites impact microbial community assembly and dysbiosis. The realm of possibilities for this approach is vast.

PN17055/2945

Design and Development of Coded Aperture for Video Compressive Sensing Acquisition System for Environment Transmission Electron Microscope

Libor Kovarik

The objective of this proposal is to develop a temporal compressive sensing (CS) acquisition system for transmission electron microscopy (TEM) imaging. CS has a potential to improve temporal resolution by an order of magnitude, enabling study of previously inaccessible dynamics of materials.

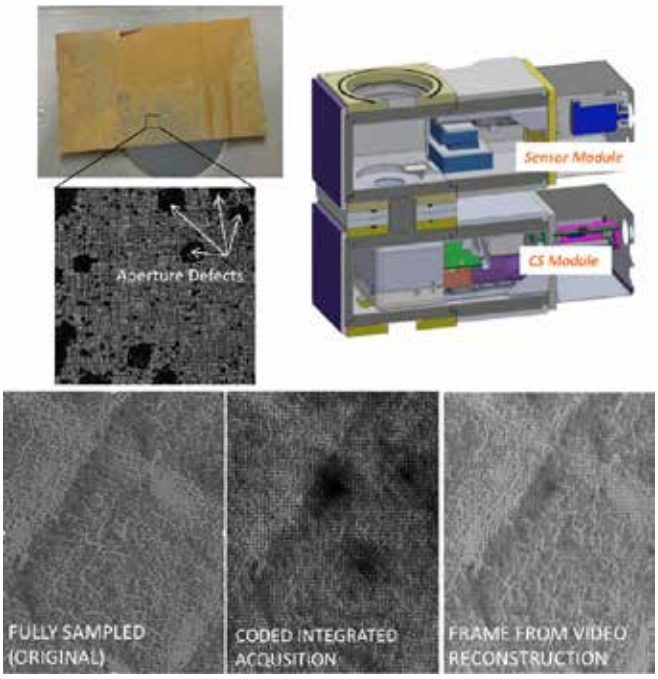
TEM microscopy represents a unique capability for probing the dynamics of energy, environmental, and biological systems under *in situ* conditions. While spatial resolution can now reach the ultimate, atomic-level limit, it is becoming well recognized that the speed and sensitivity of acquisition systems represents the main limiting factor for many experiments. Controlled by detector read-out rate, the temporal resolution is currently limited to mostly milliseconds.

In an effort to improve temporal resolution of TEM acquisition systems, this project focuses on developing video CS acquisition, which has potential to improve temporal resolution by an order of magnitude. CS represents a fundamentally new concept for enhancing temporal resolution. It is based on the idea that images/data can be acquired in sparse and coded form, which allows for faster acquisition. The sparsely acquired data can then be fully recovered using a suitable basis set using CS algorithms. In video CS, the sparse measurement is accomplished by modulation with a coded aperture (mask). The coded aperture video acquisition can be understood as a process of integrating multiple spatially coded frames into a single camera frame.

The work in FY 2017 focused mostly on completing the manufacturing of coded aperture. This work follows on our initial work in FY 2016, which involved overall development and implementation of CS. Due to shortcoming associated with previous manufacturing of coded aperture, the demonstration of video CS was not possible in FY 2016.

The key challenge for manufacturing of the coded aperture is maintaining structural stability at the required thickness of a few tens of μm and achieving well-defined aperture features at the pixel-level of the detector (few μm). The newly developed and applied electroplating manufacturing procedure lead, for the first-time, to fabrication of coded aperture suitable for testing in CS acquisition. Although the newly manufactured aperture contains several defects, the defect-free regions meet the design constraints of transmission/coding properties. The new procedure was developed in collaboration with the University of Louisville Micro/Nano Technology Center. The process involves electroplating of Au within an Si mold and subsequent etching.

Testing of the newly developed aperture was performed in a CS acquisition system built in FY 2016. The acquisition system was built as a modular camera, with one module housing a coded aperture controlled by piezo-stages and a second module housing the electron pixel sensor. The CS imaging was tested under as set of dynamic imaging conditions. At the present time, a successful reconstruction was demonstrated under a subset of these conditions.



The figure shows the newly developed coded aperture for video CS acquisition, a detailed view of aperture under white illumination, the prototype of the CS acquisition system, a comparison of TEM original images (view of standard Au grating sample) with coded acquisition of 16 frames, and an individual recovered frame from the integrated coded acquisition.

PN16056/2833

In summary, the newly developed coded aperture enabled us to demonstrate the concept of CS video TEM imaging. While the key concept was validated, there are several improvements needed for the aperture, as well as modification of the reconstruction algorithm, before this method can be applied for high-frame-rate TEM imaging.

Development of a New Mass Spectrometry Proteome Imaging Platform by Integrating Online Nano-Proteomics with Nanowell Sample Handling

Paul D. Piehowski

The objective of this project is to couple nanoproteomic sample handling technologies with laser capture microdissection (LCM) and nanoLC-MS to create a novel protein imaging capability that offers two orders of magnitude improvement in protein coverage with biologically relevant spatial resolution. This tool will allow scientists to interrogate the complex spatial heterogeneity of biological systems and tissues with unprecedented information depth.

Mass spectrometry imaging (MSI) is a powerful tool for mapping the spatial distribution of biological molecules across an area of interest. However, there are limitations to the current approach. First, molecules are transmitted directly from the surface to the mass spectrometer without separation, creating a large dynamic range of concentrations and limiting detection to the most abundant species. Second, the ionization process occurs in the context of the sample microenvironment, and since ionization efficiency is strongly influenced by this microenvironment, quantitative comparisons are challenging.

These factors are compounded when imaging proteins, which are present in significantly lower abundance than metabolites and lipids. Further, MS detection of intact protein species is challenging due to poor ionization efficiency and larger isotopomer envelopes.

To date, there is no MSI technology capable of imaging at the proteome level. Our platform aims to address these challenges by isolating tissue "voxels" into independently addressable nanowells for processing, followed by nanoLC-MS/MS analysis of each voxel. The resulting datasets are then stitched back together using an internally developed software tool called Trelliscope to create spatial maps of proteome expression.

Working with small proteomics samples, such as the voxels created for imaging experiments, faces two main challenges. First, proteins have a tendency to stick to surfaces, and thus, sample handling needs to be minimized to prevent losses to vessels and transfer pipettes. Secondly, low concentrations of proteins exhibit poor reaction kinetics, making the chemical modifications needed for analysis low efficiency.

To address these challenges, we compared two different technologies developed at PNNL: nanodroplet sample preparation and the simplified nanoproteomic platform (SNaPP). To do this, we created 100 μm^2 samples using LCM as a model system. We found that nanodroplet sample handling resulted in three-fold higher protein identifications when compared to SNaPP. Given this result, we decided to focus our efforts on optimizing the nanodroplet platform for laser dissected tissues. After testing multiple modified protocols, we came up with a new approach that increased protein coverage by 30% and reduced the number of handling steps required. We combined this with a novel tissue sample transfer approach to create a facile imaging platform.

With an approach established, we set out to benchmark the platform for proteome imaging experiments. First, we analyzed LCM-dissected tissues of various sizes to determine the relationship between spatial resolution and protein coverage achievable. We found that at 100 µm spatial resolution, we were able to identify > 1,400 proteins. This represents a 100-fold improvement over the current state of the art.

Second, we created an image of mouse liver tissue using a 5x5 pixel grid. Liver tissue was chosen for its known homogeneity on this size scale, making it a good model system for determining the technical variability of the approach. This experiment demonstrated a median protein variability of 20% across all pixels. This means that changes in protein concentration > 20% between pixels can be confidently identified using this approach.

In summary, we developed and demonstrated an approach for quantitative proteome imaging that improves protein coverage by two orders of magnitude. This technology has great potential for increasing our understanding of the molecular heterogeneity of biological tissues and how structural organization imparts function.

PN17050/2940

Dynamic 3D High-Resolution Imaging of Biogeochemical Activity in Intermediate-Scale Experiments

Timothy C. Johnson

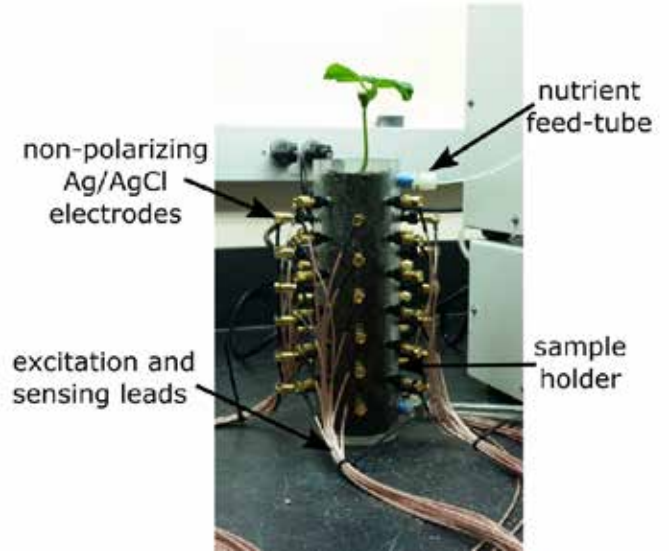
In this project, we developed the capability to autonomously monitor the frequency-dependent evolution of electrical properties in three dimensions (3D) at the laboratory-scale, using time-lapse complex conductivity tomography. This capability enables remote, non-invasive investigation of both the physical and biogeochemical interactions that govern changes in electrical properties such as root soil interactions, biodegradation, and engineered soil chemical alterations for subsurface remediation.

The low-frequency (0.01–10,000 Hz) complex conductivity spectrum is sensitive to the physical (e.g., grain size distribution, porosity), chemical (e.g., mineralogy, pore water chemistry), and biological (e.g., microbial abundance and distribution) properties of subsurface soils. Alterations in the complex conductivity spectrum in time and space are diagnostic for geochemical and biologically mediated processes occurring during both natural and engineered subsurface alterations (e.g., soil-root interaction, bioremediation, chemical amendment injection, energy resource extraction, geologic carbon sequestration, etc.).

One-dimensional measurements of complex conductivity are used extensively to non-invasively investigate numerous subsurface processes. In this project, we developed the capability to autonomously and remotely monitor the evolution of complex conductivity in 3D. This first-of-kind capability enables high-resolution investigation of the complex conductivity distribution within a medium, in contrast to existing capabilities where only the average of the entire medium is measured. By so doing, were are able to isolate localized alterations, thereby providing enhanced understanding of how subsurface biogeochemical conditions are evolving, and enabling accelerated scientific discovery.

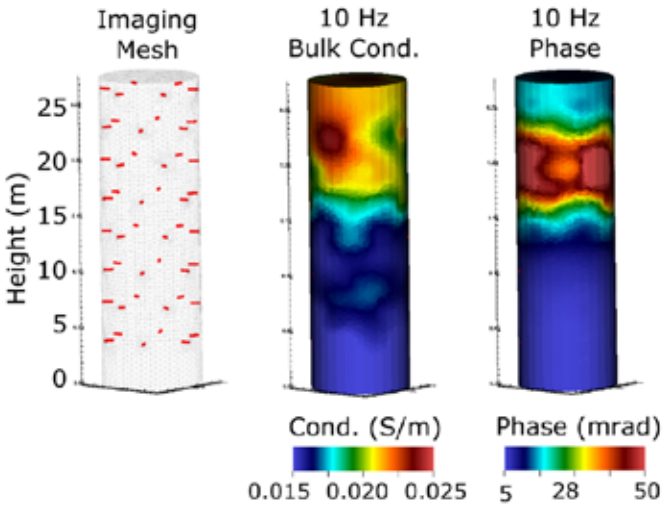
Our 4D (3D space + time) complex conductivity imaging capability is based on two recent advancements: 1) the development of a high-precision switching instrumentation elsewhere, and 2) creating the imaging system, which required developing a process for efficiently producing non-polarizing electrodes, codes to automate data collection, measurement protocols to limit data noise, the communication links between data collection and computing systems for autonomous tomographic inversion, and the numerical mesh generation routines to model sample holders, as well as fabrication of multi-electrode sample holders.

To test the system, we devised both biotic and abiotic processes as targets for time-lapse 3D imaging. These included imaging changes in complex conductivity associated with plant-root growth and during abiotic reduction and oxidation soils with high iron content.



Plant-root imaging column fabrication to image changes in low-frequency complex conductivity spectrum caused by root growth.

PN17030/2920



(Left) Computational mesh for 4D complex conductivity images. (Middle) Example image of change in complex conductivity magnitude caused by metallic inclusions. (Right) Example image of change in complex conductivity phase caused by metallic inclusions.

Overall, our system is able to image complex conductivity with designable spatial resolution. Time resolution is governed by the capabilities of the vendor-provided hardware system and is adequate given the time scales of evolution for most processes of interest.

Elemental Analysis of Rare Earths in Microfluidic Devices Capable of Employing Electrophoresis Based Separations

Amanda M. Lines

Our team aims to develop a novel microfluidic device to perform elemental separations upstream of analysis methods such as inductively coupled plasma mass spectrometry. Such a system would enhance the speed and throughput of environmental (e.g., Hanford waste tanks) sample analysis for rare earths as compared to existing methods.

Elemental analysis of rare earths is essential in a variety of fields, including environmental monitoring and nuclear safeguards. However, analysis often requires prior elemental separations that are time consuming, labor intensive, and costly. We aim to develop a novel system that significantly improves the throughput and reduces the cost of rare earth elemental separations. To accomplish this, microfluidic devices will be utilized and will allow for reductions in sample and waste volumes. Furthermore, these microfluidic devices will be designed to couple directly to analytical equipment for follow-on analysis.

Electrophoresis techniques have been successfully utilized to provide elemental separations of the lanthanides on lab-scale systems. This project aims to employ these techniques on the micro-scale. This will result in a system capable of quickly separating samples prior to follow-on analysis such as inductively couple mass spectrometry or contactless conductivity detection. Initial microfluidic device design and elemental separation optimization will be monitored via easy-to-couple analysis such as conductivity detection. Following this, the device will be modified to connect directly to a mass spectrometer for improved analysis.

During FY 2017, we identified and initiated all major procurements necessary to set up and test the microfluidic separations system. Schematics have been generated and lab space has been secured. Additionally, the first microfluidic chips have been designed and ordered.

Project work will continue in FY 2018. Initial testing will identify quality assurance and control parameters such as determining limits of detection and resolving power. Next, experiments to obtain and optimize the elemental separations of the lanthanides on the microfluidic device will be conducted. We anticipate this will be an iterative process, where parameters such as solution chemistry and microfluidic device design are altered to obtain the best level of separation as determined by the capacitively coupled, contactless conductivity detector. Ultimately, this system will be installed upstream of a mass spectrometer, which will allow for improved analysis of separation efficiency and resolution.

PN17100/2990

Embedding a Nervous System in Solid Metal Parts with Ultrasonic 3D Printing

Kayte M. Denslow

We are addressing science and technology challenges associated with ultrasonic three-dimensional (3D) metal printing, which allows for safe sensor embedment in 3D metal printed objects but requires research to address factors that govern component integrity. We are addressing challenges related to component integrity that need to be overcome in order for 3D printed metal objects to be adopted for mission-critical system monitoring and by-design ubiquitous sensing.

In the burgeoning age of ubiquitous sensing, solid materials will have their own "nervous system" of embedded sensors made possible by additive manufacturing (3D printing). The ability to safely embed sensors and ensure 3D-printed component integrity will be driving forces behind a paradigm shift that will advance ubiquitous sensing from passive and non-selective to *by-design*. Ensuring our national security will require the use of smarter, highly networked, and secure materials in strategic assets and critical infrastructure in order to enable real-time sensing, detection and tracking of threats, and new modes of information gathering.

Developing these smart materials requires us to embed temperature-sensitive sensors and associated electronics in metals without damaging them. To achieve this, the use of "cold welding" techniques is necessary to join metals instead of high-temperature techniques such as laser-sintering or thermal fusion. A promising and emerging 3D manufacturing "cold welding" technique that enables 3D printing of metal components with embedded sensors is 3D ultrasonic consolidation (3D-UC).

The parameters that result in high-quality 3D-UC metal components are loosely understood, and the lack of underpinning science results in the use of a trial-and-error optimization stage during fabrication. This inefficient aspect of 3D-UC represents the chief hurdle between its current state of use and its wider adoption, which would transform rapid prototyping and manufacturing of high-impact technologies (e.g., metal smart structures, wearable sensors, lab-on-a-chip). Better understanding and control of this nascent additive manufacturing technique is needed to overcome this challenge and enable efficient, reliable manufacturing of smart metal components and structures with "nervous systems" that can have high impact in national security and energy applications.

We procured, installed, and tested a bench-scale ultrasonic welder necessary to support research on factors that influence ultrasonically welded metal components.

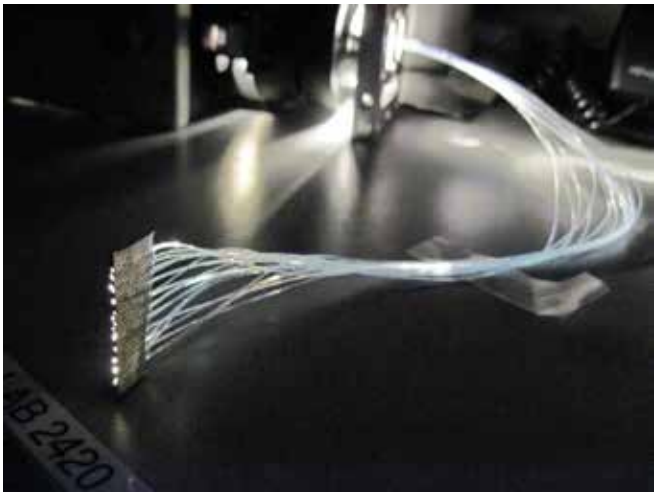
Test procedures were developed to enable quantitative measurement and *in situ* process monitoring. We began the process of designing and assembling multilayer mixed materials and built demonstration parts with aluminum, copper, titanium, stainless steel, and brass.

We aim to characterize effects on delicate sensor components within parts with *in situ* sensors. To that end, optical fibers were embedded in aluminum and copper parts. The optical fibers may be used to monitor pressure, temperature, and vibration during the 3D-UC process. With this information we will gain a better understanding of the characteristics of successful 3D-UC joining. We aim to better predict and control process outputs. Future tests will focus on collecting data through the optical fibers during the 3D-UC process.

Finally, we are exploring how to use the new capability to design and produce novel non-proliferation technologies. In the future, arms control and verification will require the use of smarter, highly networked, and secure materials in strategic assets in order to enable remote real-time sensing, detection, tracking, and new modes of information gathering. We are designing new materials that will enable inherently secure items with built-in identifying and tamper-indicating features. These technologies will enable better detection of treaty violations or other activities of proliferation concern.

In FY 2018 we plan to finalize and execute a test plan to improve our fundamental understanding of 3D-UC, collect interfacial data with embedded sensors, correlate data with component integrity test results, and identify data signatures that represent. Finally, we will embed temperature-sensitive objects (batteries, microelectronics, and energy harvesting materials) into 3D-UC materials.

PN17017/2907



Plastic optical fibers were non-destructively embedded in solid aluminum (lower left) with ultrasonic 3D printing.

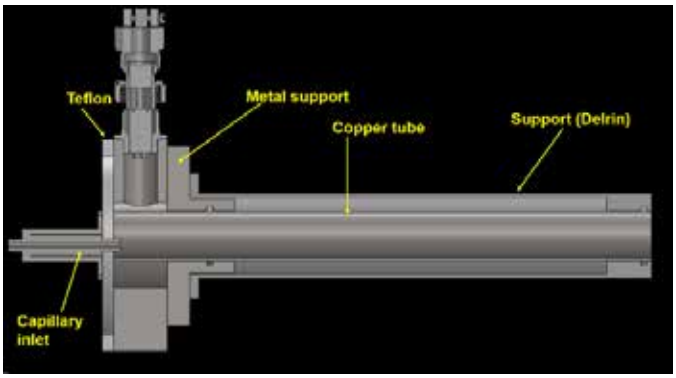
Integration of an Atmospheric Flow Tube Ionization Source (AFT) with a Novel Ion Mobility Analyzer

Robert G Ewing

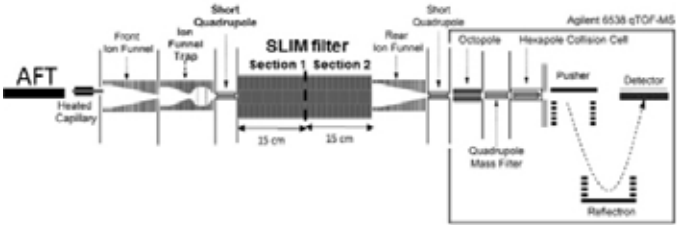
Field detection of illicit substances important to national security issues requires smaller instruments that are robust and have minimal power requirements, preferably hand-held. This project will investigate the development of a new generation of a field-deployable detection technology with improved sensitivity and selectivity over current capabilities.

Atmospheric Flow Tubes (AFTs) integrated with Structures for Lossless Ion Manipulation (SLIM) devices enable the capability for the detection of traces of chemical species in the atmosphere. More specifically, the real-time detection of illicit substances in the air is a potential benefit of the technology.

In FY 2016, it was reported that the SLIM-IMS module was assembled and interfaced to an Agilent QTOF mass spectrometer (MS). In that setup, a dielectric barrier discharge source was used. The setup was then optimized for selective ion transmission. In FY 2017, the AFT was designed and integrated with the SLIM device. A dielectric barrier discharge source was used. The AFT is equipped with a fan with controlled speed to adjust the air flow through the AFT. The copper tube is installed such that the capillary inlet is protruded into the tube for optimum ion sampling.



Cross section view of the AFT.



The experimental setup showing the integration of AFT with SLIM – MS.

The AFT was then integrated with the SLIM–QTOF mass spectrometer. The performance of the setup was evaluated with a few compounds, including tributylamine (TBA). The intensity and the stability of the spectra was found to be affected by both the air flow speed and also the position of the flow tube (copper tube) with respect to the capillary.

The ion current was measured at the interface between the SLIM module and the QTOF mass spectrometer. The measurement was done with the TBA sample and was found to be ~5 pA. This indicates that, with a suitable current amplifier, it would be possible to perform experiments without a mass spectrometer.

The sensitivity of the device was also evaluated with cocaine samples as a representative of illicit drugs. A thin layer of cocaine was deposited on a glass slide and was placed nearby the ionization source. The sensitivity of the system for cocaine was high and was found to increase as the air flow decreased. Although the setup was very sensitive, the intensity was not sufficient enough to measure the ion current with a Faraday plate.

PN16057/2834

Ion Manipulation at Atmospheric Pressure

Yehia M. Ibrahim

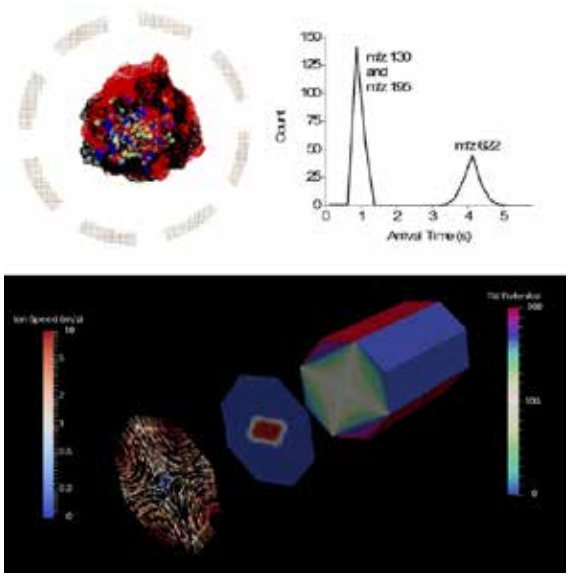
This project aims at advancing our ability to detect and characterize various molecular species relevant to environment, health, security, and new materials. When successful, this project will result in development of a sensitive and highly selective device and do so with reduced cost.

Portable ion mobility spectrometers have limited performance due to the constraints of size and the operating parameters. These constraints stem from the inherit nature of the constant electric field used in such devices and its inability to provide confinement of ions at atmospheric pressure. This work uses a different approach to confine ions at atmospheric pressure. Using traveling wave (TW) DC profiles with appropriate device geometry and TW parameters, ions can be efficiently confined and transmitted through the device. TW can also be used to affect ion mobility separation (IMS). Unlike constant field IMS, TW is a fixed amplitude revolving voltage profile enabling a significant increase in the resolution by increasing the path length. Since the ions can be confined efficiently at atmospheric pressure, one can achieve high IMS resolution without compromising sensitivity. This project is focused on optimizing the geometry of the device and its operating parameters using ion trajectory simulation to show the proof-of-concept ion transmission through the device.

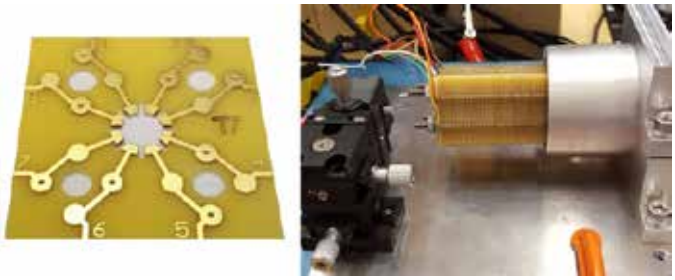
Different geometries of the electrode patterns were investigated using ion trajectory simulations, which provided insights on the optimum geometry and operation of the device. Results indicated the importance of electric field strength and its frequency to provide efficient operation of the device. Ion simulations also produced snapshots of ion velocity vectors, ion plume and voltage at an instant of time, which helped understand the ion confinement.

We designed and fabricated the lenses of the ion confinement at atmospheric pressure (APIC) device using printed circuit board (PCB) technology. The PCB fabrication method provided a flexible, fast, and affordable means to build the device. A TW is applied to the eight electrodes to confine ions.

The lenses were assembled into a 10-cm-long device. APIC was integrated with electrospray ionization source to provide ions from a standard mixture of compounds. Ion current was measured at different points throughout the device using a miniature probe. An ion current of ~50 pA was measured at the exit of the device using a Faraday plate. These results indicate the success of the device to confine and transport ions at atmospheric pressure. The experimental results also confirm the predictions provided by the ion trajectory simulations.



(Top Left) SIMION trajectory simulation showing ions confined within APIC device (Top Right) IMS separations simulated in SIMION. (Bottom Left) Velocity vectors of ion (Center) ion plume distribution (Bottom Right) Voltage profile at an instant of time.



(Left) Photo of individual lens showing eight electrodes forming a central confinement region. (Right) Photo of the assembled device integrated with an electrospray ion source.

PN17089/2979

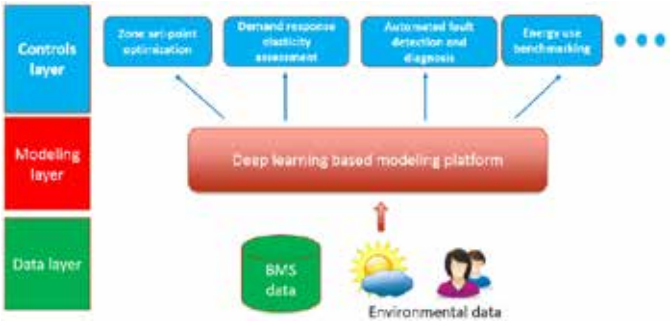
Learning Control for Building Systems

Abhinav Vishnu

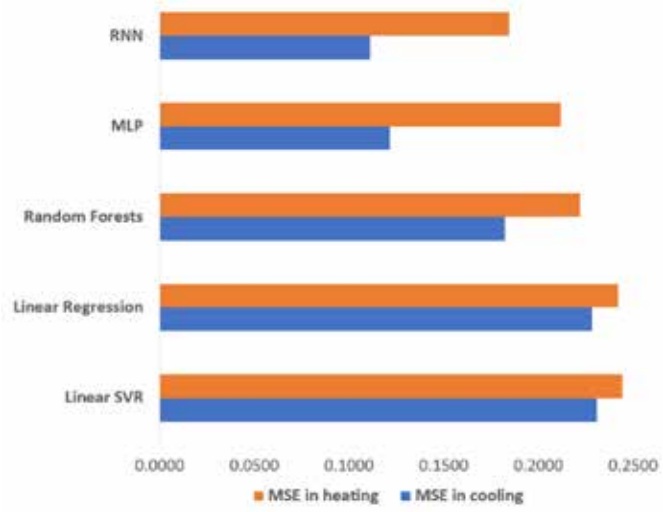
More than 40% of the total energy consumed in the United States can be attributed to buildings. This project investigates the use of machine learning techniques to leverage large amounts of data available from buildings in order to better predict and optimize their energy consumption and operational costs.

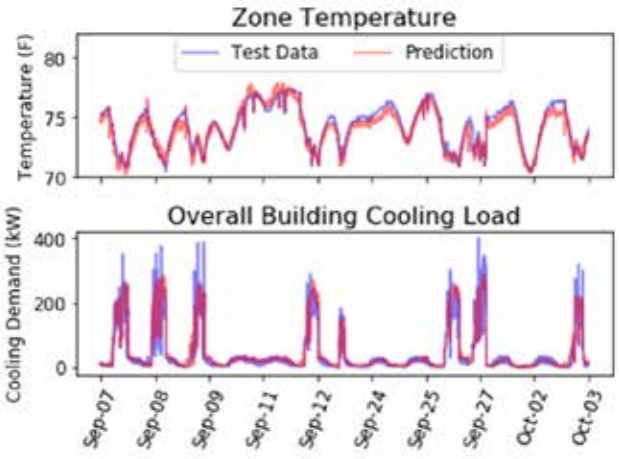
In order to optimize the operation of buildings, models that describe the relationship between energy consumption and "control knobs," such as set points, with high predictive capability are required. Physics-based modeling techniques are often labor intensive, requiring several physical parameters to be known beforehand. Data-driven modeling techniques are promising, but have been investigated to a somewhat limited extent for optimizing the operation and control of buildings. In this context, deep learning techniques such as Recurrent Neural Networks (RNNs) hold promise, empowered by advanced computational capabilities and big data opportunities.

We aim to investigate the use of deep learning for modeling and minimizing the power consumption and cost of operating buildings. Scalable modeling, control, and optimization frameworks, tested on data from real-world buildings, are the main technical deliverables targeted by the project.



We applied a recurrent neural-network-based modeling framework to an 8-month data set from a building on the PNNL campus. We used RNNs because of their ability to capture dynamic and nonlinear system characteristics. We focused on modeling the building heating and cooling demands and the indoor zone temperatures as a function of control inputs and exogenous inputs (e.g., the outdoor air temperature). We found that RNNs resulted in better mean squared error performance, an 8–52% reduction in error compared with other data-driven methods. Further, we found that a single hidden layer architecture with 128 nodes provided the smallest mean squared error.





In FY 2017, the modeling was performed on a data set from a single office building. While some preliminary findings were obtained with regard to the architecture of the network to minimize the prediction errors, a more thorough investigation of architectures using additional data sets is required as an area of future research.

In FY 2018, we will explore whether the knowledge obtained by training models on one data set can be transferred to other buildings of similar type. The utilization of machine learning models to minimize energy and cost requirements is also a future direction for research. In addition, development of use cases to enable these objectives will be a focus in FY 2018.

PN17025/2915

Signatures for Early Disease Detection: Application of a Non-Invasive Multi-Modal Sensor System for Bovine and Swine

Luke J. Gosink

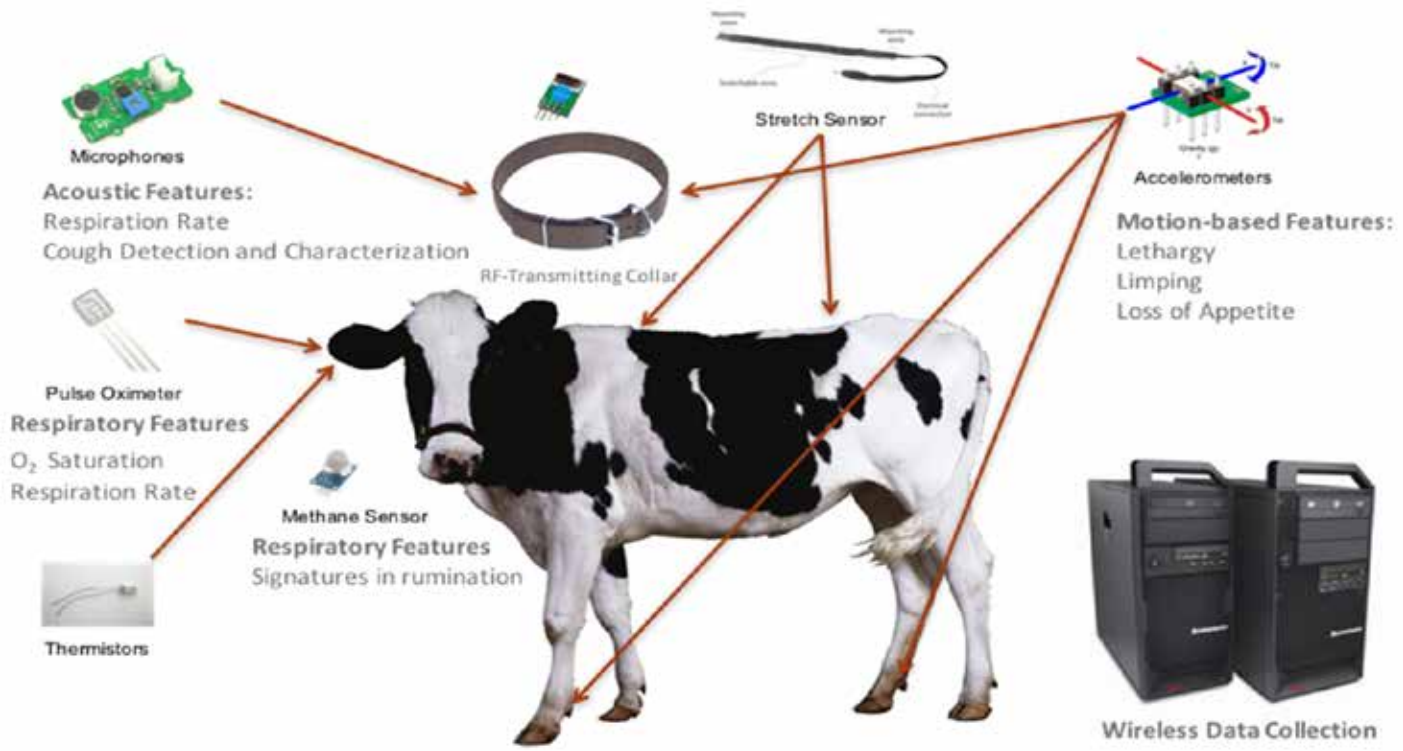
This effort will demonstrate a novel multi-modal, non-invasive, wearable sensor system, called Yellow Jacket, and its ability to monitor animals and detect early indications of bovine respiratory disease (BRD) in cattle.

The motivation for this work is based on the need to mitigate the devastating effects of panzootic or pandemic events. Whether a result of agro-terrorism or natural viral/bacterial evolution, such events have disastrous impacts to national security in terms of loss of life, immediate depletion of national food stores, and massive devastation to economic resources. For example, diseases like BRD are associated with a cumulative $4 billion annual loss to a $100 billion food industry.

The project will work with collaborators at Mississippi State School of Veterinarian Medicine to help evaluate the Yellow Jacket system. Through monitoring and early detection, this effort hopes to establish proof of concept for data-driven approaches that can be used to help better direct the use of antibiotics.

In FY 2016, Mississippi State used the Yellow Jacket system to monitor animal health in its challenge studies, which were sponsored and conducted independent of this work. While PNNL staff members were onsite to monitor the Yellow Jacket systems, they were not involved in working with any animals.

In FY 2017, PNNL completed the analysis of the data that was collected in FY 2016. This analysis identified several unique signatures that helped to classify early indications of lethargy and respiratory illness (e.g., activity levels, body position, height of head during activity times) and challenged respiration rates.



The Yellow Jacket sensor ensemble used in Mississippi State's challenge studies. Sensors used included accelerometer, microphone, thermister, and pulseoximeters.

PN16094/2871

Untethered – Coherent Millimeter-Wave Sensing Using Drones with Visual Motion Capture

Jonathan R. Tedeschi

The Untethered project has developed next-generation millimeter-wave (MMW) imaging techniques for national security and commercial applications with the use of drones and advanced MMW RADARs. Untethered enables autonomous, through-barrier, high-resolution, close-range imaging.

Previously developed imaging systems for use on low-altitude flying drones perform two-dimensional (2D) synthetic aperture radar (SAR) with resolution on the order of 3 meters from distances of kilometers. While this may be sufficient for air-to-air recon and terrain mapping, this does not achieve substantial resolution to detect personnel born explosives, improvised explosive devices/explosively formed penetrators, or infrastructure inspection. These technologies are inherently limited, as they must operate at high frequencies to achieve any lateral resolution, which limits penetration through barriers.

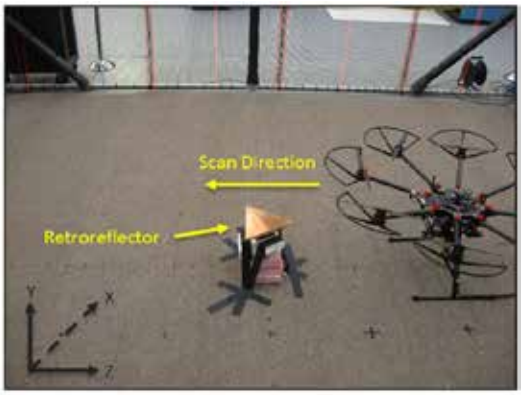
High-resolution through-barrier imaging has been achieved with MMW imaging technologies to the greater than 1mm level for security screening with standoff ranges of less than 1m to several meters, but this technology is limited to very rigid antenna array controls where unpredicted motion of either the object being scanned or the antenna array creates image distortion. The Untethered project aims to remove the barrier of rigidly controlled scan requirements and achieve high-resolution imaging at the centimeter level by the development of fusing multiple sensors with advanced MMW-focusing algorithms. This technology enables MMW imaging in areas that are logistically challenging or dangerous for a human operator to be present.

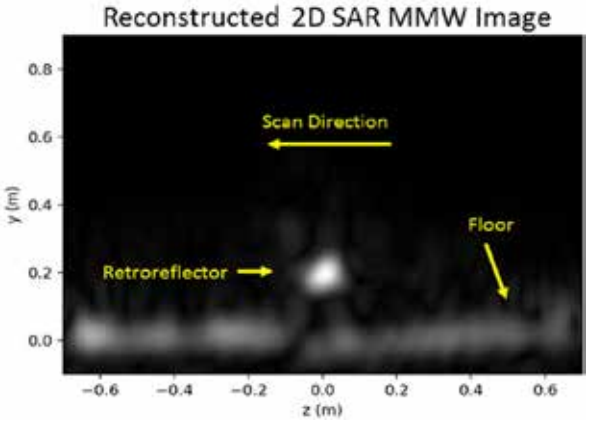
PNNL has developed an indoor operating environment for drone-based MMW data collection. The prototype configuration utilizes a drone with a max payload of 24lbs, allowing for antennas, RADARs, and additional control hardware to be installed locally on the drone.

The drone was flown over a retroreflector target. As the drone was in flight, the ground beneath it was illuminated, and the reflected signals were detected by the receive antenna.

Using the advanced MMW imaging algorithms developed in this project, the MMW data was focused, achieving high resolution. This 2D SAR MMW image with the new focusing techniques achieved 5-cm downrange and lateral resolution.

The FY 2017 efforts demonstrated the initial concept of 2D SAR high-resolution MMW imaging on a drone platform at close ranges. The research to be performed in FY 2018 will mature and expand the capabilities of the technology.





PN17002/2892

Utilizing High-Resolution Ion Mobility Separations in Multi-Omic Analyses of Biologically Important Isomers

Erin M. Baker

The objective of this research was to study small molecule isomers that have been implicated in phenomics but are known to be difficult to detect or separate with currently available technology.

Small molecules have very similar elemental and structural makeups; therefore, new technologies with high-resolution, sensitivity, and throughput are needed for better separations to occur. The overall objective of this study was to perform better structural separations of small molecules with currently available drift tube ion mobility separations (DTIMS) and then compare these to our high-resolution IMS separations in structures for lossless ion manipulations (SLIM) devices. We expected that the SLIM IMS analyses would unveil structures that were hidden from our currently available IMS systems.

To accomplish the aims of this project, first we created a collision cross section (CCS) database of more than 500 small molecules with the DTIMS instrument. To date, IMS metabolomic and exposomic studies have been limited by an inadequate number of accurate CCS values for small molecules, causing features to be detected but not confidently identified. In this work, we utilized DTIMS to directly measure CCS values for over 500 small molecules, including primary metabolites, secondary metabolites, and xenobiotics. Since DTIMS measurements do not need calibration like some other IMS techniques, they avoid the calibration errors that can cause problems for distinguishing structurally similar molecules.

All measurements were performed in triplicate in both positive and negative polarities with nitrogen gas and seven different electric fields, so that relative standard deviations could be assessed for each molecule and structural differences studied. The primary metabolites analyzed to date have come from key metabolism pathways, such as glycolysis, pentose phosphate pathway, and TCA cycle, while the secondary metabolites consisted of classes, such as terpenes and flavonoids, and the xenobiotics represented a range of molecules from antibiotics to polycyclic aromatic hydrocarbons. Different CCS trends were observed for several of the diverse small molecule classes, and when features were matched to the database, the addition of the IMS dimension greatly reduced the possible number of candidate molecules. This CCS database and structural information are freely available for download at http://panomics.pnnl.gov/metabolites/, with new molecules being added monthly.

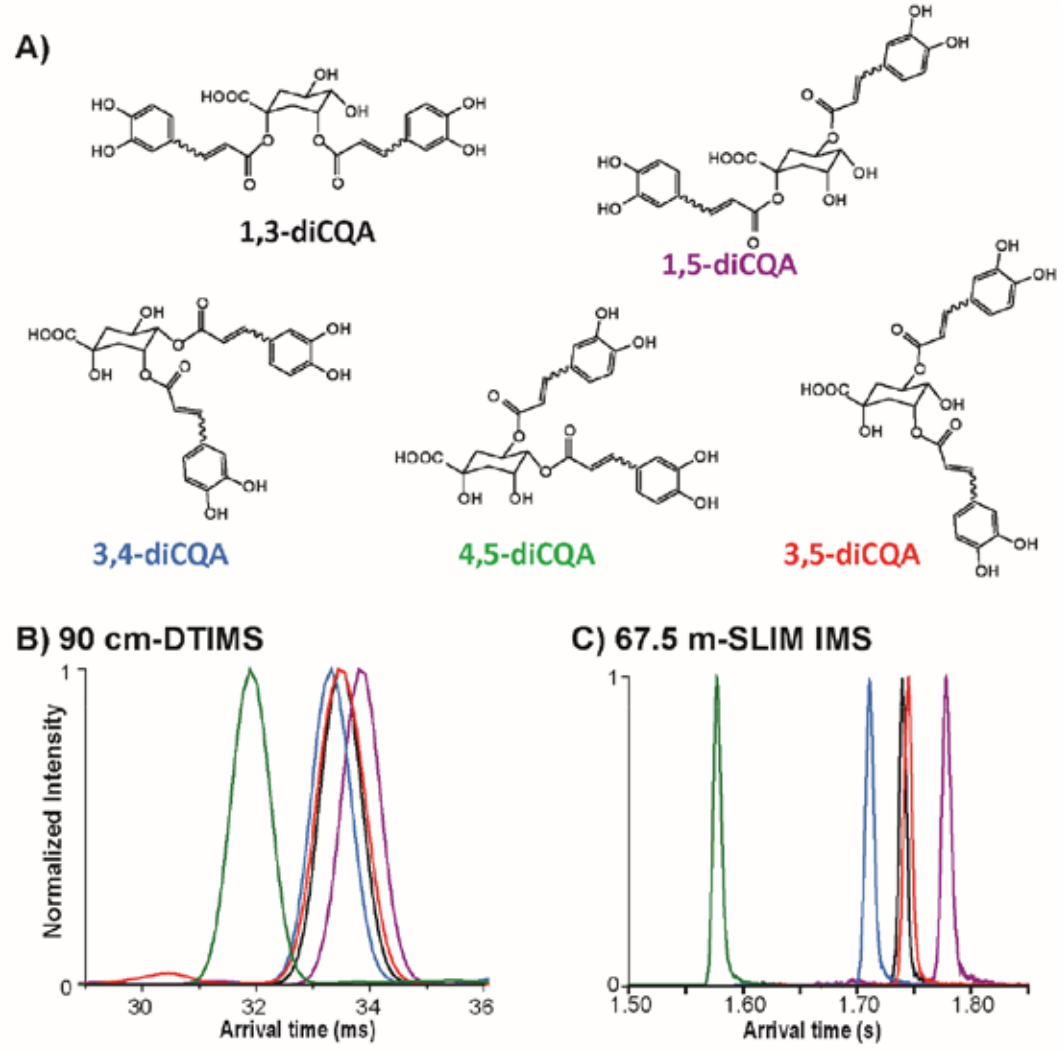
To perform our next aim of comparing the DTIMS instrument with the SLIM platform, we analyzed dicaffeoylquinic acids (diCQAs). This natural-product-based molecular class of secondary metabolites has recently been uncovered for having health benefits and anti-HIV activities, so understanding their structures and functions is of great interest for drug discovery efforts. DiCQAs are analytically challenging to identify and quantify since they commonly exist as a diverse mixture of positional and geometric (*cis/trans*) isomers.

In this study, we characterized the positional isomers for the five diCQA standards (1,3-diCQA; 1,5-diCQA; 3,4-diCQA; 3,5-diCQA; and 4,5-diCQA) using a 90-cm DTIMS-MS platform. Each diCQA isomer displayed a single IMS peak, consistent with the fact that these compounds naturally exist as trans/trans isomers. The arrival time distributions for these positional isomers were similar; however, 4,5-diCQA eluted first (indicating a smaller structure) and was baseline separated, while 1,5-diCQA had the longest arrival time and the largest structure. The CCS value for each diCQA isomer was measured by DTIMS. To better differentiate the positional isomers, the SLIM IMS-MS platform was then used to obtain ultrahigh resolution IMS separations. After five passes of the SLIM IMS-MS platform (67.5 m drift length), 4,5-diCQA; 3,4-diCQA; and 1,5-diCQA were well separated, and 1,3-diCQA and 3,5-diCQA partially separated.

This study tackled the difficult problem of identifying and quantifying positional isomers of phytochemicals in natural product and plant extracts, which to date, has been extremely challenging due to the vast structural diversity of the isomers present in the mixtures. The ultrahigh resolution SLIM IMS-MS technology enabled baseline separation of the isomers and was a powerful tool for analyzing the complex

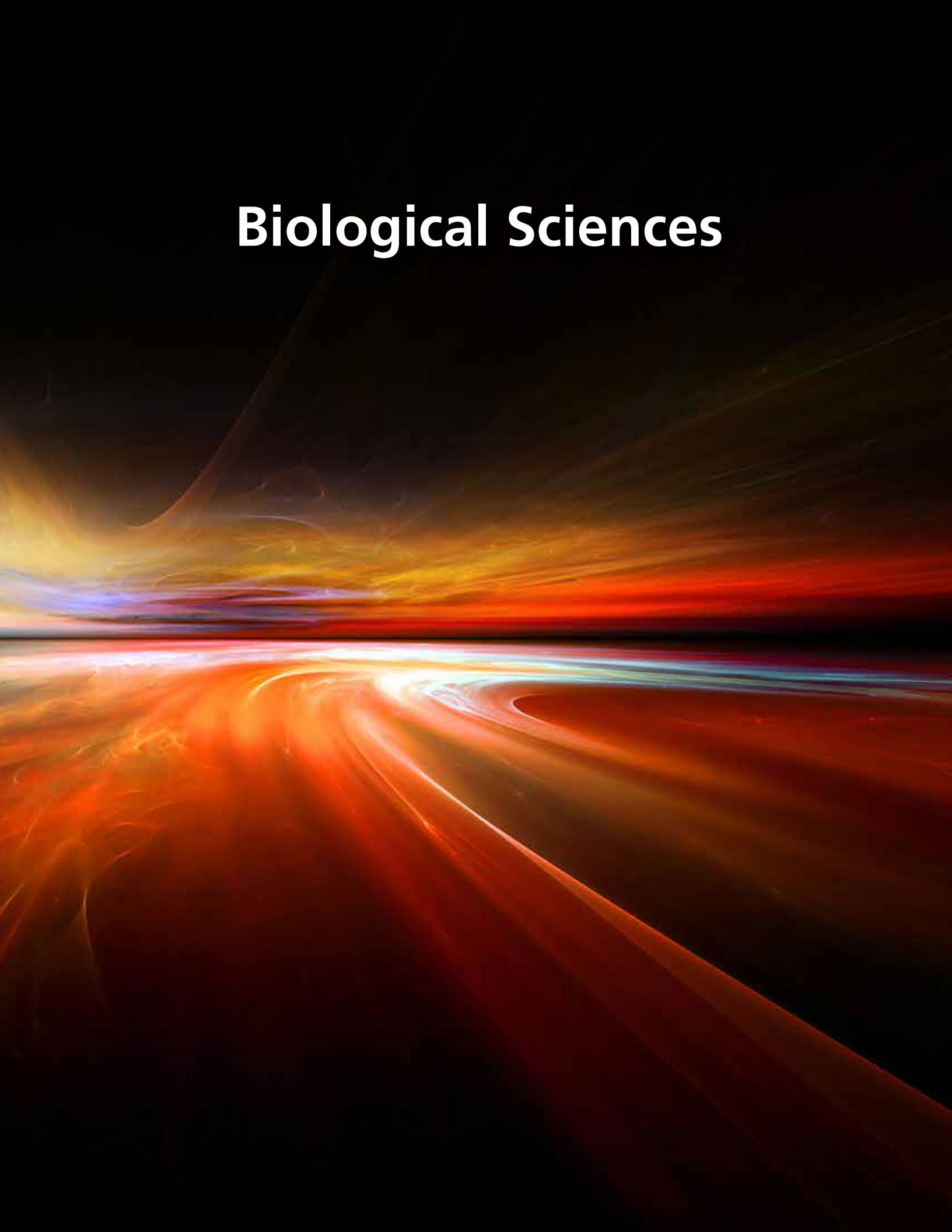
PN17047/2937

natural products and elucidating structure-activity relationships. While the DTIMS platform allowed separation, distinguishing the different components was difficult since they were not baseline resolved, therefore showing the great power of the SLIM platform.



IMS characterization of the dicaffeoylquinic acid positional isomers 1,3-diCQA; 1,5-diCQA; 3,4-diCQA; 3,5-diCQA; and 4,5-diCQA. **A**) The chemical structures of each isomer and ATDs using a **B**) DTIMS platform and **C**) ultrahigh resolution SLIM IMS.

Biological Sciences



Advancing Ecosystem Understanding of Carbon Turnover and Storage through Molecular Characterization

Kirsten S. Hofmockel

We are developing a mechanistic understanding into the coupling of carbon and nitrogen cycling through biotic and abiotic interactions in order to improve community-based land models. These models play an important role in helping to understand feedbacks between terrestrial carbon cycling and climate change.

The main goal of this project is to evaluate whether or not nitrogen inputs into the terrestrial environment alter the composition and availability of carbon residues for mineralization. The concurrently changing carbon and nitrogen cycles remain a key uncertainty in understanding feedbacks between the terrestrial carbon cycle and climate change, despite decades of research on decomposition responses to elevated rates of human-caused atmospheric nitrogen deposition. The coupling of carbon and nitrogen in Earth system models fundamentally alters critical feedbacks between the land biosphere and the global climate system, greatly attenuating ecosystem response to rising atmospheric CO_2 concentrations compared to carbon-only models. Yet, existing coupled models do not consider the full suite of linked carbon-nitrogen processes, particularly belowground. For example, the investigation of mineral and carbon interactions is not well understood. Here, it is generally accepted that nutrient additions increase belowground net primary productivity that should enhance physical protection of organic matter, especially in clay-rich soils, where acidification can increase the solubility and abundance of trivalent hydrolyzing cations (Fe^{+3}, Al^{+3}) that are particularly effective at stabilizing organic matter onto clays via cation bridging.

In order to complete the stated project goal, we have partnered with the Nutrient Network (an ecosystem scale nutrient addition field experiment) in obtaining soils from six field sites. Research plots at each of these field sites have received nitrogen amendments over several years. We are currently extending molecular knowledge about these sites to include the characterization and comparison of soil organic matter, mineralogy, metagenomics, and metaproteomics profiles between amended and control plots within each site. Further, to address our hypothesis that *N addition will increase the adsorption of plant and microbial residues onto mineral surfaces, particularly in sites with finer textured soils*, we are conducting soil incubations under controlled moisture and temperature conditions. In the absence of new carbon inputs, accessible carbon will be respired by microbial communities, leaving behind stabilized and/or chemically recalcitrant molecules, which will be identified and correlated to the mineralogy at each site.

Significant progress was made toward completion of the project goal and addressing the stated hypothesis. Of note was completion of the long-term controlled incubation experiment described above, which was conducted over 7 months. Respiration measurements demonstrated significant differences in the cumulative carbon mineralized, and the effect of nitrogen amendment on carbon respiration could be observed.

Additionally, during this experiment, incubation jars were sacrificed periodically and soils harvested for organic matter, mineralogy, and omics data generation, which is nearing completion. Results show distinct differences in soil organic matter profiles for each NutNet site, and preliminary results suggest a strong correlation between soil organic matter profiles and a site's mineralogy. This result will be confirmed with additional mineralogy measurements made during FY 2018.

A herculean effort was made in regard to the generation of metagenomic sequence data from soils harvested from the 70 incubation jars. These sequence data, along with the other omics generated data, represent the largest set of its kind for this type of experiment and will be a valuable resource for mechanistically understanding the dynamics and response of the soil microbiome as the soil becomes carbon depleted. Marginal soils are often starved for labile carbon, and this information will be highly informative as to role the soil microbiome and mineralogy play in the development of this condition.

PN16073/2850

At the Fringe of a Shifting Carbon Paradigm with Climate Change: Unlocking the Organo-Mineral Controls on the Bioavailability of Carbon at the Terrestrial-Aquatic Interface

Ryan S. Renslow

Through our creative, experimental approach of simulated pedogenesis and unique combination of technologies that span across fields and disciplines, this work will advance our understanding of the mineralogical and biological mechanisms that regulate carbon cycling along the terrestrial-aquatic continuum in soils and sediments.

Recent evidence highlights the importance of organo-mineral interactions in regulating the source or sink capacity of soil. High surface area soils, such as allophane-rich or clay-rich soils, retain organic matter (OM) via sorption to mineral surfaces, which can also contribute physical isolation in interlayer spaces.

In addition to the mineralogical influence, stabilization of OM can be influenced by development of reduced conditions (e.g., high water content and solubility and transport limitations on oxygen availability) that inhibit decomposition. Though recent literature for soils suggests that mineral interactions are important in governing the fate of OM, the capillary fringe can be permanently or seasonally inundated with water, making it hydrologically similar to marine sediments, where mineral adsorption is considered the governing process for persistence of OM.

Since near-surface capillary fringe sediments would be subject to pedogenic processes, it is unclear what the contribution of surface area and adsorption is to stabilization of OM in these sediments. Most of the work to date, which relates surface area to stabilization of OM, has focused on investigations of natural soils where OM has been incorporated onto mineral surfaces before sampling and details on the mechanisms for this dynamic system can only be inferred, limiting our ability to properly predict the source sink capacity of soil. This research is taking methodical steps toward elucidating these mechanisms.

To test the mechanisms of mineral surface area protection of OM, we are facilitating secondary precipitation of aluminosilicates in the presence of OM held at two different temperatures in natural Nisqually River sediments (from the Mt. Rainier area). A 3-month reaction was completed, which is intended to simulate early pedogenesis.

We coupled high-resolution carbon classification with microbiome characterization and CO_2 emissions in a short-term laboratory incubation to test the combined effects of temperature (4˚C versus 20˚C) and moisture (50% versus 100% water-saturated) in capillary fringe sediments maintained under different mineralogical conditions (mineral-amended versus non-amended).

Our results show that less CO_2 was produced in mineral-amended sediments compared to non-amended sediments at 20˚C and only in water-saturated sediments at 4˚C. The abundance of condensed hydrocarbons and compounds less thermodynamically favorable to microorganisms were also relatively depleted in mineral-amended sediments compared to non-amended sediments. In these mineral-amended sediments, the reduction in thermodynamic favorability was especially enhanced in water-saturated versus non-saturated sediments. Shifts in the microbiome were most pronounced between water-saturated and non-saturated samples from mineral-amended sediments (with little effect of moisture content in non-amended sediments). Our results show that increased mineral content accentuated the effect of moisture on CO_2 production, carbon composition, and on the microbiome, suggesting that microbial access to carbon is governed by both mineral sorption and moisture content.

Furthermore, we found that the increase in temperature showed a net increase in iron coatings, mineral surface area, and formation of colloids less than 200 nm in diameter. The result was a leaching rate of the higher temperature conditions soils that was two to three times higher than the lower temperature. This information helps fill knowledge gaps on mechanisms for carbon release from soil at the terrestrial-aquatic interface, thus improving our understanding of carbon cycling in general.

PN16075/2852

Biological Threat Signatures for *Bacillus anthracis*

Owen P. Leiser

Being able to distinguish laboratory-adapted pathogens from those that are naturally occurring remains a challenge for the research and biodefense communities. This work will enable a more rapid identification of the source of disease outbreaks and a more effective response.

Previous work at PNNL has been focused on identifying and characterizing protein, carbohydrate, and other cellular signatures of pathogens grown in laboratory media. In addition, we have shown that the commonly used genomic tools for identification and characterization of pathogens are insufficient for providing information regarding whether an isolated pathogen has been cultured in a laboratory, which would be a likely prerequisite to its use as a bioweapon. This project was initiated to provide an in-depth investigation of both wild and laboratory strains of *Bacillus (B.) anthracis* using proteomics to understand the protein signatures that differentiate benign from possibly intentional infections.

In FY 2015, two of our team members traveled to the Emerging Pathogens Institute at the University of Florida to meet with Dr. Jason Blackburn and compile a list of global wild and laboratory *B. anthracis* strains from his collection, to be grown for proteomic analysis for this research project (a total of 80 samples). Dr. Blackburn also visited PNNL in July 2015 to give a seminar and have additional project discussions. His laboratory prepared the *B. anthracis* samples for proteomic analysis at PNNL and verified that the samples were sterile through extensive testing. However, a Centers for Disease Control and Prevention transport moratorium on *B. anthracis* delayed the shipment of samples to PNNL until FY 2016.

Also in FY 2015, a team member traveled to Ventura, California, for the Gordon Research Conference on Chemical and Biological Terrorism Defense and presented data about *Yersinia (Y.) pestis* threat signatures, which were generated during a previous PNNL effort that led directly to this project. During the meeting, a partnership with Northwestern University was established that expanded our project capabilities and included a proteomic characterization of *Y. pestis* and *Y. pestis* infection and allowed additional proteomic studies with *Y. pestis* while we were awaiting *B. anthracis* samples.

We were provided samples of lung fluid from *Y.-pestis*-infected and uninfected mice, as well as cultured *Y. pestis* samples for proteomics analysis (all samples were inactivated and non-infectious prior to sending to PNNL). This work investigated proteomic signatures of *Y. pestis* isolated from mouse lungs after infection and *Y. pestis* processed without any laboratory culture steps. As expected, we observed significantly increased protein abundance in virulence determinants in the *Y. pestis* isolated from mouse lungs after infection compared to the cultures grown under laboratory conditions, and we also observed unexpected changes in central metabolism and stress response. The host response to *Y. pestis* infection was also investigated and can serve as a baseline for future proposed studies using additional strains of *Y. pestis* from different evolutionary lineages. These data offer additional insight into protein expression in laboratory-adapted pathogen strains during infection, which also informs our main efforts investigating *B. anthracis*.

At the end of FY 2015, we invested a significant amount of time further analyzing existing unpublished data that supports project goals and drafting manuscripts. Specifically, we drafted two manuscripts describing the work to study long-term adaptation of *Y. pestis* to laboratory conditions, we modified a manuscript and added data that investigated the use of proteomic signatures of *Y. pestis* grown in different laboratory media for forensics, and we revisited work that examined the differences in proteomic signatures of *B. anthracis* spores produced in laboratory media versus soil. The two manuscripts discussing adaptation of *Y. pestis* to laboratory conditions were ultimately merged into a single manuscript, which was published in December 2015.

In FY 2016, we received the delayed 80 autoclave-inactivated *B. anthracis* biomass samples and began trial proteomic preparations to determine optimum experimental conditions. Samples were processed in randomized batches over the course of about 3 months, beginning in January 2016, and analyzed using liquid chromatography (LC)-mass spectrometry (MS)/MS bottom-up proteomic techniques. All 80 samples were run on the LC-MS/MS instrument in duplicate by July 2016.

PN15061/2736

Training was acquired in tissue culture techniques at Biosafety Level 2 (BSL2) using commercially available human primary and immortalized cell lines and an attenuated strain of *B. anthracis*. This training provides the foundation for future work at both BSL2 and BSL3 controls and safeguards, which will be critical to success in pathogen signature science.

At the July 2016 Cascadia Proteomics Symposium in Seattle, Washington, preliminary findings from this project were presented, which included discussion on the utility of proteomics techniques for forensic applications.

This project also led to collaboration in FY 2016 with another project using proteomics data to train a machine-learning algorithm, the Lasso regression classifier, to distinguish wild from laboratory strains of *Y. pestis*. The same approach was applied to proteomics data generated from the *B. anthracis* cultures.

The classifier was able to differentiate wild and laboratory strains of *B. anthracis* with over 90% accuracy. In addition to training and data production, these data were extensively analyzed and a manuscript prepared detailing the findings of the work. These results show that wild strains of *B. anthracis* express proteins essential for sporulation and metabolic processes associated with sporulation and, therefore, survival in the environment at significantly higher levels than their laboratory cousins. These results will serve to inform characterization of unknown strains and potential differentiation between routine and nefarious disease outbreaks.

We also worked closely with Dr. Jason Blackburn at University of Florida to prepare a manuscript presenting genomic and proteomic data regarding a recently identified Nigerian strain of *B. anthracis* that possesses metabolic qualities relevant to vaccination efforts in West Africa, in comparison to the virulent Ames strain and Sterne vaccine strain.

Copperphilic Coating for Development of an *In Situ* Antimicrobial Coating

Eric M. Winder

This project uses PNNL's expertise in biofilms and materials development as a means to investigate novel material to combat biological fouling.

Continued anthropogenic release of antifouling materials, including copper, into the environment is concerning. Instead of applying toxic, metal-laden coatings onto ships and other marine structures, the proposed research aims to utilize the copper already present in the water (0.25 ppb open ocean; 5.4 ppb in San Diego Bay Yacht Basin; 1.13 ppb mid-Columbia river surface water) to create antifouling coatings. This would permit development of an antifouling coating while also aiding in removing free copper from the aquatic environment. Without additional copper input, this 1.13-ppb copper concentration in the Columbia river is already above levels shown to effect juvenile Coho salmon.

This *in situ* development of an antifouling coating would be accomplished through painted application of a copperphilic material. When applied to a boat or structure, the coating would have zero toxic components, but upon placement in water, the paint would begin to adsorb and concentrate copper (c.f., electroless copper plating, but with eco-friendly biological substrates rather than harsh industrial chemicals).

While many heavy metals are only toxic in a soluble, "bioavailable" form, copper has been shown to be toxic/inhibitory as both a soluble cation and as a solid or bound-cation. Additionally, while the proposed copperphilic materials would have binding constants well below the available concentration of copper and could be generalized as a copper chelator, it would be expected that the bound copper would rapidly transition from bound to unbound and back to bound states quickly and continuously as the microenvironment changes. Ultimately, this bound copper would provide a means to delay fouling through inhibitory effects on fouling organisms.

For this proposed project, scientists at the Sequim and Richland PNNL locations have been working to produce and/or acquire the copperphilic materials and test their physical properties.

These chemical materials and the biological materials (e.g., laccase) obtained will be tested by Microtox® assay testing at the Marine Science Laboratory (Sequim, Washington) to determine their toxicity to *Aliivibrio fischeri* as indicated by disruption of luminescence. Down selected, non-toxic materials will be exposed to raw seawater in the coming months to test their antifouling/microbial capability.

Select Sorbent Physical Characteristics

Materials	Sorbent Characteristics	
	Surface Area (m^2/g)	Pore Diameter (nm)
Organic Ligands-Based Sorbent		
SH-NP Silica	791	*
SH-NF Silica	190	*
IDAA-NP Silica	*	*
Diphos-SH-NF Silica	123	-
PropPhos-NP Silica	364	6.9
IDAA Styrene-SH-NF Silica	89	*
EDA-NP-Silica	*	*
Inorganic-Based Sorbent		
Mn-Fe_3O_4 MNP (8 nm)	280	-
Fe_3O_4 MNP (8nm)	160	-
MnO_2-NS Composites	193	11.6
Inorganic-Based Sorbent		
Actinide Resin	102	*
Diphonix Resin	6	1.4
Ln Resin	64	3.8
Chelex 100 Resin	3	9.4
GT 74 Resin	55	*
AGMP 100-200 (SAX) Resin	2	1.9
CG50 (WAX) Resin	62	*
MnO_2 Resin (Eichrom)	3.2	26.1
Activated Carbon	1400	4.2

*Denotes materials testing in progress.

PN17091/2981

Cultivation-Independent Untangling of Microbial Gene Regulation Networks

Aaron T. Wright

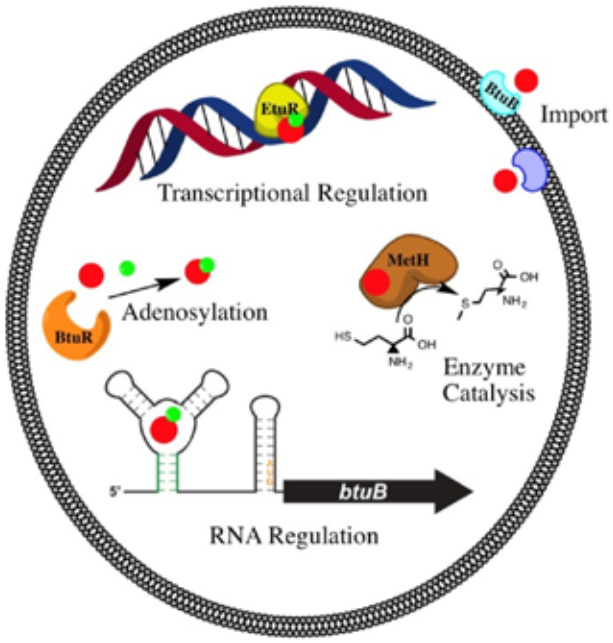
We designed a metabolite-based probe approach to identify the genomic targets of transcription factors in uncultivated microbes.

Currently, less than 1% of the world's microbes are thought to have been isolated and cultivated in the laboratory, despite significant efforts. Our inability to separate the remaining more than 99% stems largely from the fact that, in nature, microbes are almost never found on their own but rather exist as members of diverse microbial communities. As community members, they typically interact strongly with each other, exchanging resources as they cooperate and compete with one another. In contrast to macro-organisms, interactions in microbial communities occur at the molecular scale and are difficult to decipher. The mechanisms by which species and the community's functions change over time remain an ongoing investigation.

Microbes regulate their gene expression in response to environmental cues and the presence or absence of specific metabolites—chemical compounds involved in growth and survival. This regulation occurs by the binding of specific metabolites to proteins known as transcription factors. Once the transcription factor binds its respective metabolite, it changes shape and can turn "on" or "off" gene expression. When changes in gene expression are detected by metatranscriptomics, it can be difficult to interpret the exact mechanistic cause or transcription factors involved in the change. This limits our understanding of gene regulation in microbial communities and hampers our ability to predict how these communities will respond to perturbation.

In this project, we first tested the feasibility of this approach using the well-known repression of tryptophan biosynthesis in *Escherichia (E.) coli*. The transcription factor TrpR is a transcriptional repressor, binding to the DNA when the intracellular concentration of Trp is high. Using a bioassay, we tested a series of tryptophan probes, demonstrating that several are transported into living cells, can bind TrpR, and repress gene expression. This bioassay helped us determine that it will be difficult to develop a probe able to form a tertiary TrpR-probe-DNA complex that facilitates complex enrichment. The synthesis of the tryptophan probes and their application to biology was recently published and featured on the journal cover of *Organic Chemistry Frontiers*.

We developed a similar bioassay to demonstrate that our vitamin B_{12} molecular probe (B_{12}-ABP) binds the transcription factor EutR to induce expression of genes in the ethanolamine utilization pathway. Results demonstrated that the B_{12}-ABP is inducing gene expression in *E. coli*, like the unmodified Cyano-B_{12} vitamin. This is the first time ever that a molecular probe, containing a diazirine cross-linking motif, has been shown to form the tertiary protein-metabolite-DNA complex and be transcriptionally active. With this assay, we tested how well a series of different B_{12}-probe structures could induce transcription. We determined the best B_{12}-ABP probe structure and could enrich and identify EutR via proteomics. We now have the technical challenge of enriching the DNA that is bound to the transcription factor and probe.



B_{12}-ABP interactions validated include import by btuB; enzyme catalysis of MetH, NrdJ, and EtuBC; adenosylation by BtuR; RNA regulation of the btuB riboswitch; and transcriptional regulation of EutR.

PN15010/2685

Deciphering Microbial Communication through Metabolites

Thomas O. Metz

This project will establish a Metabolomics Center at PNNL, which will develop the advanced measurement capabilities necessary to define the metabolic processes of microbiomes and determine how these processes change due to perturbation.

PNNL seeks to address scientific questions that include 1) what are the impacts of perturbations on microbial community metabolic processes, 2) what are the impacts of changes in microbial community metabolic processes on ecosystem function and health, 3) can a systems-level understanding of microbial community interactions and metabolic processes be developed, and 4) can improved models be developed to better predict the impact of perturbations on microbial community metabolic processes.

With these questions in mind, understanding microbial metabolism in a community context is essential. Metabolomics measurements will play a key role in defining the metabolic processes of microbial communities and in determining how these processes change due to perturbations. Examples of microbiome perturbations include transitions of the human microbiome due to industrial processes associated with energy extraction and production (e.g., exposures to radiation, oil spills, pesticides, and fertilizers for biofuel feed- stocks) and transitions of environmental microbiomes due to climate change.

A related aspect of microbial metabolism is communication via metabolites and other small molecules. These molecules are used by microbes to communicate with each other, their environments, and their hosts. However, the majority of microbial metabolites have not yet been identified, nor have their roles in interactions between community members been characterized. The objectives of the Metabolomics Center are to 1) determine the relative influence of direct (e.g., via secondary metabolites) versus indirect (e.g., via changes in microbial community metabolite pools) microbial communication and interaction through metabolites and other small molecules; 2) increase the identification coverage of the metabolomes of microbial communities, their hosts, and their environments; and 3) identify the mechanisms by which metabolites involved in communication and interaction are generated, transported, and sequestered within microbial communities.

In FY 2017, this project had the following goals: 1) create a knowledgebase of identified metabolites, unidentified metabolite features, their chemical information, and associated microbiome metadata and 2) develop a pipeline for structural elucidation of important, novel, or unidentifiable molecules detected in soil and gut microbiomes.

The first goal was realized by the development of a prototype microbiome metabolomics knowledgebase that houses gas chromatography-mass spectrometry and nuclear magnetic resonance (NMR) spectroscopy data from a metabolomics study of wetlands soil subjected to differing hydration regimes. The metabolomics data is complemented by comprehensive experimental metadata, which enables users to query the variations associations of metadata and metabolomics data in order to test existing and develop new hypotheses.

Work on the second goal focused on development of a novel pipeline for identification of unknown metabolites in complex samples. The samples are first analyzed using in-house Fourier transform ion cyclotron resonance mass spectrometry capabilities in order to annotate as many molecules as possible using molecular formula. Next, those compounds whose formula cannot be matched to relevant metabolite databases are then isolated and enriched using offline liquid chromatographic separations. The enriched compounds are then subjected to NMR analysis in order to achieve confident chemical structure characterization.

For FY 2018, work will continue to develop the metabolite isolation, enrichment, and identification pipeline, and biochemically novel metabolites that are identified will then be used as the basis for synthesis of activity-based protein probes for enrichment of proteins that bind to the novel metabolites. The proteins will then be identified using proteomics analysis, and their protein sequences used to map to the associated genome sequences of relevant microorganisms.

PN16019/2796

Decision Support Framework for Assessing Stream Temperature Effects on Fisheries Resources in the Columbia River Basin

Ryan A. Harnish

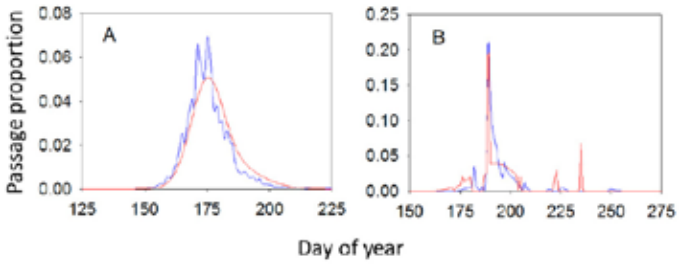
Warming temperatures have the potential to extirpate Pacific salmon from the Columbia River Basin. Our modeling framework allows for the identification of the life stages and habitats at which directed management actions will be most effective at conserving these species of great socioeconomic importance under future climate conditions.

We developed a decision support/modeling framework for identifying climate-induced alterations to water quantity and quality that may jeopardize the persistence of anadromous salmonids in the Columbia River Basin. Because operation of the Columbia River hydrosystem is regulated to facilitate passage of anadromous salmonids, climate change may also affect the future of hydropower in the basin, making this an important ecological and economic issue. Identifying the times and locations in which aquatic environments are altered beyond the thermal tolerances and habitat requirements of these important species is the first step toward developing potential mitigation solutions.

Although similar analyses have been conducted previously to evaluate the potential effects of climate change on anadromous salmonids of the Pacific Northwest, past efforts have failed to quantitatively assess the effect of environmental conditions on the timing of life stages, of which the initiation and duration are at least partially dependent on stream temperature and discharge. In addition, each life stage has unique physiological tolerances to environmental conditions. Therefore, in order to accurately depict the effects of climate change on anadromous salmonids, it is critical to understand how alterations to the river environment may affect life stage phenology.

Previous efforts have modeled future climate conditions at too coarse a resolution to evaluate the effects of climate change on fish passage conditions at hydroelectric dams. A separate PNNL project is underway to link multiple climate, hydrology, and hydrodynamic models to produce high time (e.g., hourly) and space (e.g., 30 m) resolution, spatially distributed future water quantity (i.e., discharge) and quality (i.e., temperature, total dissolved gas) projections for the Columbia River Basin. These high-resolution projections of future stream conditions applied to our modeling framework, which allows for the plastic response of fish to environmental conditions, will provide a tool set for more accurately and precisely evaluating the probable location, timing, and duration of events that are likely to affect the migration, growth, survival, and life stage phenology of anadromous salmonids in the Columbia River Basin.

Functions relating environmental conditions (i.e., temperature and discharge) to significant life stage events (e.g., adult river entry, hydrosystem migration, spawning, egg hatch, emergence, smolt migration) were identified for several Columbia River Basin salmon populations (i.e., Hanford Reach and Deschutes River fall Chinook, and Wenatchee, Okanogan, and Redfish Lake sockeye) using published relationships and historic environmental conditions and observational fish data. Modeled relationships were validated by comparing multiple years of observational data of adult and smolt passage counts at hydroelectric dams to those predicted from the model.



Example of validation results displaying the model-predicted (red) and observed (blue) adult sockeye salmon passage timing at Bonneville Dam on the Columbia River (A) and at Zosel Dam on the Okanogan River (B) in 2016.

This model, used with high space and time resolution projections of future water quantity and quality, can help to identify the geographic locations and salmonid life stages that may become at risk due to climate-induced increases in stream temperature or reductions in streamflow. Identifying the locations at which management actions will be most effective optimizes resource allocation and the chance of success.

PN17058/2948

Decomposers in Transition

Kirsten S. Hofmockel

Microorganisms played a central role in forming the atmosphere of the earth, yet we are unable to accurately model how microbes affect current and future climates. This work will provide fundamental biological information required to understand and model how microorganisms transform plant residue from the soil to the atmosphere.

Microbially explicit carbon cycling models are based largely on homogenous liquid cultures that do not represent the heterogeneous soil environment.

To build a robust, process-driven, predictive model of such a highly complex and heterogeneous system, we must create an experimental framework that is reduced enough to be controlled, yet complex enough to relate to the actual soil environment.

Heterogeneity is the hallmark trait of the soil and, therefore, fundamental to our understanding of soil microbial ecology. Spatial heterogeneity generates a range of microbial niches that influence microbial community composition and function. Moreover, the spatial structure of soil is dynamic and can change in response to seasonal dynamics, resulting in the formation and turnover of soil aggregates, roots, flowpaths, etc. This spatial structure is the foundation for interpreting microbial responses to changing soil moisture, which impacts the connectivity between microorganisms and their exchange of nutrients and growth substrates. Yet, the effects of spatial heterogeneity on microbial ecology are largely ignored, because it is so challenging to incorporate into an experimental and modeling framework.

As such, we have developed an experimental platform that varies the spatial complexity of the microbial habitat to test how the connectivity of microbes to each other and to substrates affects microbial metabolism of soil carbon. Using activity-based protein profiling, in combination with metatranscriptomic analysis, will enable us to accurately identify the key organisms and enzymes regulating carbon decomposition.

By explicitly confronting spatial structure in both our model and bench experiments, we aim to accomplish our goal to determine how microbial interactions and community responses differ under well-mixed and structured conditions and, thus, inform microbially explicit models, which are currently the cutting edge of carbon decomposition modeling.

Growth conditions were developed during FY 2016 for soil microbes cultured in broth, uniform structural and non-uniform structural microcosms representing homogenous and heterogeneous spatial environments. Measurements were made to determine microbial community composition, protein composition (proteomics), extracellular enzyme production, carbon substrate metabolism, and carbon-use efficiency.

Microcosms were also inoculated with cells cultured from soil sampled from an agricultural field site in Iowa. Soil microcosms were amended with individual carbon substrates ranging from simple sugar to recalcitrant crystalline cellulose.

To decipher the metabolic strategies of the microcosms communities under the given conditions, we made progress utilizing an activity-based cell sorting approach to sort cells based on a given function. The activity-based approach we applied utilized chemical probes designed to covalently tag proteases, glycoside hydrolases, and redox sensitive proteins. Once cells were probe-labeled *in vivo* and subsequently fixed, fluorescent markers were appended to the probe-labeled proteins and sorted via flow cytometry. DNA was isolated from probe positive, probe negative, and unsorted cells from each condition and sent out to be sequenced. These results will inform us on which organisms have the specified function and are active within the growth parameters.

From our measurements, we contributed to empirical parameters to the microbial cell model JAM—Just Another Microbial—and modified it to fit our batch soil microcosm experiments. Through collaborative interactions with data scientists, statisticians, microbial ecologists, chemists, and modelers, we now have incorporated empirical experimental data to determine growth and individual carbon source consumption rates to constrain parameters of the existing JAM model. These experiments were designed and implemented so that mixed carbon sources could be used for incubations in order to stimulate growth and enhance community diversity to reflect a population that echoes the native soil while generating data required to improve predictive modeling.

PN16092/2869

Determining Mechanisms of Microbial Metal Mobilization in Coastal Wetland Environments

Rene M. Boiteau

Environmental and land use change is expected to significantly alter patterns of terrestrial temperature and hydrology, which affect the availability of micronutrient metals. The purpose of this project is to understand feedbacks between hydrologic and climatic factors and metal availability that impact the rates at which organisms produce and break down organic carbon in soils and sediments.

Many biologically essential trace metals, such as iron, cobalt, nickel, copper, and zinc, are scarcely soluble in aquatic environments. Previous studies have demonstrated that organisms regulate the solubility and bioavailability of these metals through the production of organic metal chelating agents with a range of metal binding strengths and specificities. Determining what compounds are produced where and when is central to understanding how microbial communities adapt to changes in metal supply that result from external environmental factors such as hydrology or pH.

The goal of this project is to characterize metal active compounds directly from soil environments using hyphenated liquid chromatography mass spectrometry and determine which organisms are utilizing them by identifying the microbes that possess the biosynthetic pathways for these compounds in order to address the following questions: 1) What metabolites bind and solubilize metals in soil/waters across environmental gradients and how do they impact metal bioavailability? 2) What organisms produce and take up these metal species in the environment? 3) What chemical or environmental triggers activate specific metal acquisition strategies? 4) What metabolic pathways are co-regulated with metal metabolism and what is the potential impact of these changes on rates of organism growth and carbon cycling?

During the current year of this project, the capability of analyzing trace metal speciation at PNNL using trace metal clean chromatography hyphenated with inductively coupled plasma mass spectrometry has been established to achieve sufficient detection limits to detect metal complexes at naturally occurring concentrations. This analytical tool was used to detect metal binding species (Fe, Cu, Ni, Zn) in environmental samples, which were then identified by coupling the same chromatography to high resolution electrospray ionization mass spectrometry (Orbitrap or Fourier transform ion cyclotron resonance) to obtain parent ion and fragmentation information, and by capturing the fractions and analyzing the isolated compounds by direct infusion with Fourier transform ion cyclotron resonance mass spectrometry to obtain molecular formula based on accurate mass.

Numerous compounds bound to these elements were detected and identified in grassland soils. These compounds included plant siderophores involved in iron acquisition by grass species and siderophores that are commonly produced by basidomycota and ascomycota fungi, which dominated the fungal taxa at this site based on internal transcribed spacer sequencing. In addition, several compounds were metal complexes that have not previously been observed but appear to be analogues of plant metabolites. Many of these compounds were complexed to metals other than iron, suggesting alternative roles for these compounds in acquisition of other micronutrient metals. The affinity for different metals was compound specific. This further implies that suites of molecules that compete for metal binding are used in metal regulation and highlights previously unrecognized interdependences of metal acquisition pathways.

In addition, microbial metagenome mining was used to investigate which organisms produce these compounds. Metagenomes from Kansas Native Prairie soils were screened for the presence of genes involved in metal metabolite production. An entire assembled cluster containing multiple pyoverdine biosynthesis genes was found. This gene cluster closely matches those of *Pseudomonas*, an abundant taxa within these soils, suggesting that pyoverdines may potentially be important for iron acquisition in these environments. Follow up chemical characterization experiments revealed that sorption of pyoverdines to soil particles reduces their solubility in bulk soils—a characteristic that is not true of plant or fungal analogues that appear in the same soils. This result raises new questions about the microenvironments in which *Pseudomonas* species actively secrete these metal chelators compared to the plant and fungal species. During FY 2018, additional experiments will be conducted

PN16106/2883

using a model strain of *Pseudomonas* to understand which conditions promote pyoverdine production, which members of the rhizosphere community of Kansas soil are capable of utilizing pyoverdine-bound iron, and whether pyoverdine biosynthesis aids the growth of grass species in calcareous soils.

In order to evaluate the effect of these metal chelating metabolisms on growth rates, energy expenditure, and interactions between organisms, models of metal uptake must be constructed. Metal addition experiments with the compounds isolated from Kansas soil suggested that the kinetics of metal exchange are likely slow (hours to days), and thus metal speciation is likely controlled not by equilibrium binding processes, as is typically assumed in environmental systems, but by kinetic association and dissociation processes. A platform for generating kinetic models of metal speciation with multiple ligands and metals has been generated in R and will be expanded during FY 2018 to include multiple plant and microbial chelators and multiple metals. This tool will be used to evaluate hypotheses about the interrelationship between metal uptake pathways and environmental conditions.

Compound turnover rates are important parameters that determine the efficiency of chelation strategies, but are currently unknown in environmental systems. These will be determined by ^{13}C incorporation into chelate pools in laboratory experiments. To achieve this goal, a novel computational algorithm was developed to identify compound labelling rates from chromatographically resolved ultra-high resolution mass spectrometry data in an untargeted fashion. As a model system to develop this approach, switchgrass was grown in calcareous soils and incubated with isotopically labelled carbon dioxide for two days. These tools simultaneously determine the turnover rate of other major metabolite groups within the rhizosphere.

Development of High-Throughput Metabolomics Technologies: Application to Studying the Flowering Time in *Arabidopsis thaliana*

Kerem Bingol

In this project, we develop novel approaches for rapid identification and quantitation of unknown and known metabolites in complex metabolite mixtures. We apply these new methods to understanding developmental timing of plants in response to atmospheric [CO_2] rise, which will have major implications on a global scale, since [CO_2] is expected to disrupt carbon cycling within ecosystems. We envision that the project will lead to unprecedented discoveries of new biomarker metabolites related to developmental timing of plants.

The total number of different metabolites in our universe could range up to 200,000 or more. Currently, we are limited to routine analysis of only less than 1% of them; the remaining 99% have unknown structures. Mass spectrometry (MS) and nuclear magnetic resonance (NMR) spectroscopy are the two major experimental analysis techniques in the analysis of metabolites. The future trajectory of metabolomics and its applications largely depend on the analytical capabilities of NMR and MS and their combination. Although the use of both NMR and MS methods is increasing, in the majority of these studies, the two methods are essentially implemented independently from each other. This approach does not fully capitalize on the complementary strengths of these two analytical techniques.

Our goal is to develop advanced sample preparation, hyphenation (e.g., liquid chromatography- [LC-]MS) and computational approaches with enhanced integration of NMR and MS. We will use these capabilities for accurate and high-throughput identification and quantitation of unknown and known metabolites in the model plant organism, *Arabidopsis (A.) thaliana*. Previous studies have shown that sugars act as signaling molecules that can influence the downstream expression of flowering genes. Elevated [CO_2] generally increases the sugar profile of plants. This may be the fundamental mechanism for how elevated [CO_2] influences flowering time. In order to fully test this hypothesis, comprehensive profiling of all carbohydrates and metabolites is needed, which is only possible by identification and quantitation of all unknown and known biomarker metabolites in *A. thaliana*. Therefore, this case study will be excellent demonstration of our newly developed technologies.

In this project, we hope to identify new biomarker metabolites and metabolic pathways regulating the flowering time of *A. thaliana*.

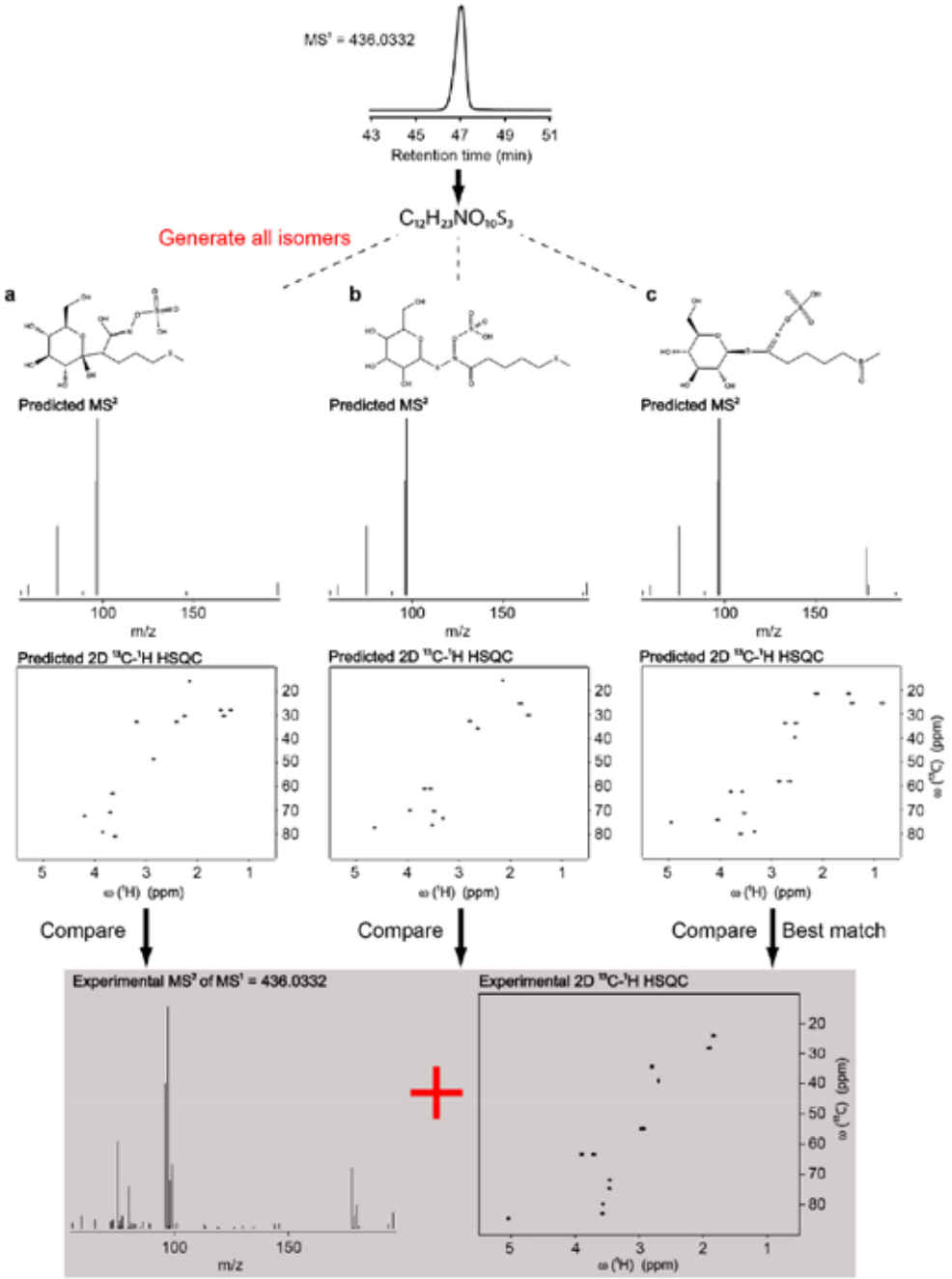
In FY 2017, our first task was to develop an optimized sample preparation protocol for comprehensive study of plant metabolome. The next step was to develop a combined LC-MS and NMR protocol for characterization of plant metabolites. Moreover, we further integrated this with a new cheminformatics approach, named as ISEL NMR/MS^2. ISEL NMR/MS^2 allows rapid identification of unknown metabolites in metabolomics samples by combining MS/MS and NMR predictions. We tested the ISEL NMR/MS^2 approach in *A. thaliana*. This allowed identification of a new secondary metabolite (glucoraphanin) enrolled in plant signaling.

During FY 2018, our focus will be on developing advanced quantitation techniques for combined NMR/MS analysis and its application. Quantitative analysis in metabolomics is often accomplished by multivariate statistical analysis techniques such as principal component analysis and partial least-squares regression. Although combined use of NMR and MS is becoming popular, the majority of these studies perform separate multivariate analyses of their respective NMR and MS data sets. Only at the end are the results combined in an attempt to enhance the total information content. Such analyses completely ignore the highly informative correlations between NMR and MS data sets and, therefore, are not ideal to handle data analysis of combined NMR/MS studies.

We will develop the capability for quantitative analysis of combined NMR/MS studies. We will test and compare currently available spectral processing, alignment, binning, fitting, and statistical analysis platforms. We will adapt the best performing approach to the analysis of NMR/MS datasets of *A. thaliana* metabolome. We expect our quantitative data analysis approach will provide better separation of groups and greater levels of model interpretability as compared with multivariate statistics applied to NMR or MS datasets alone.

PN17011/2901

The unknown biomarkers obtained by this analysis will be identified by our ISEL NMR/MS^2 approach. Ultimately, all of these efforts will lead to discovery of new metabolic pathways in *Arabidopsis*, regulating the flowering time in response to atmospheric [CO_2] rise.



Application of ISEL NMR/MS^2 on *A. thaliana* metabolite extract having a single unknown metabolite in the NMR spectrum. The approach starts with an unknown LC-MS feature, and then combines the experimental MS/MS and NMR information of the unknown to effectively filter the false positive candidate structures based on their predicted MS/MS and NMR spectra. This allowed identification of glucoraphanin, which was not present in the experimental metabolomics databases (i.e., unknown/uncatalogued metabolite). ISEL NMR/MS^2 is the first metabolomics approach that combines MS/MS and NMR predictions for unknown metabolite identification.

PN17011/2901

Dynamic Multiscale Modeling of Complex Biosystems (DMMCB): A Framework for Multiscale Metabolic Modeling

Garrett B. Goh

This project envisions creating an interpretable deep learning framework for computational biology and chemistry applications, working towards a data-driven research paradigm that uses artificial intelligence to accelerate conventional research and development.

The rise and fall of artificial neural networks is well documented in the scientific literature of both computer science and computational chemistry. Yet almost two decades later, we are now seeing a resurgence of interest in deep learning, a machine learning algorithm based on multilayer neural networks. Within the last few years, we have seen the transformative impact of deep learning in many domains, particularly in speech recognition and computer vision, to the extent that the majority of expert practitioners in those fields are now regularly eschewing prior established models in favor of deep learning models. Unlike traditional machine learning, deep learning distinguishes itself in its use of a hierarchical cascade of non-linear functions. This allows it to learn representations and extract out the necessary features from raw unprocessed data needed to make the necessary predictions, all without the need of expert knowledge.

In computational chemistry, the state of deep learning is extremely new. Based on the results from a few recent studies, we have observed the broad application of deep learning in many sub-fields of computational chemistry, including computer-aided drug design, computational structural biology, quantum chemistry, and materials design. In almost all applications we have examined, the performance of deep learning models are frequently superior to traditional machine learning algorithms Nevertheless, we have identified that almost all applications of deep learning in computational chemistry, and also in other fields in general, have been using deep learning as a classifier or a regression tool, leveraging its representation learning ability.

The objective of this research is, therefore, to develop a framework for elevating deep learning (deep neural networks) from being yet another machine learning tool to a hypothesis or concept generation tool. In order to achieve this, we will lay the groundwork to establish interpretable deep learning models for computational chemistry.

In FY 2017, we developed Chemception, a deep convolutional neural network that uses simple 2D molecular drawings for predicting a broad range of chemical properties. Despite being trained on unsophisticated raw data, we demonstrated how Chemception, with minimal expert knowledge, is able to perform at parity to more established models in the literature that rely on sophisticated engineered features such as molecular fingerprints and descriptors.

Next, we explored the effects of changing the amount of chemical information provided to Chemception. Our results indicate that with a small addition of basic chemical information, an augmented Chemception model consistently outperforms established models. In addition, it also demonstrates proof-of-concept that first-principles understanding of the research problem is not a pre-requisite for deep learning models to make accurate predictions. *In other words, deep neural networks can be used to make predictions from data even in situations when no underlying theory has been developed*, and this significantly expands the type of research challenges that can be potentially pursued.

Instead of using image data (molecular drawings), we also explored the possibility of using chemical text data (SMILES) to predict chemical properties. The resulting model, SMILES-2vec, was shown to perform at a level similar to augmented Chemception. In addition, we developed an explanation mask that identifies the most important part of the string responsible for its prediction. We concluded that the network identifies chemically relevant functional groups for solubility predictions, thus indicating that deep neural networks are capable of learning technically accurate scientific concepts.

Lastly, we also explored the effectiveness of using transfer learning for chemical property prediction. The resulting model, ChemNet, has been shown to be a universal, general-purpose, deep neural network that can be quickly adapted to predict a broad range of small-molecule chemical properties.

In FY 2018, we have identified two objectives. The first is continued development of deep learning models for chemistry applications, such as combining different modalities (data formats) in training deep neural networks, and leveraging of graph representations of molecules. The second objective is to demonstrate proof-of-concept that this data-driven, domain-agnostic approach that we have developed can be applied to other technical fields besides chemistry.

PN16003/2780

Dynamic, Multimodal, Molecular Imaging of Live Biological Systems

Julia Laskin

We are developing a unique, modular platform for quantitative, dynamic, multimodal imaging of multiple metabolic pathways in living cells and tissues with subcellular resolution and exceptional molecular sensitivity and selectivity.

Molecular-level understanding of complex dynamics of biological systems, such as signaling or metabolic pathways over lengths and time scales, is a grand challenge in biology. Many existing individual imaging modalities are typically limited to specific molecular properties and biological processes. Therefore, multimodal imaging is necessary for predictive understanding of complex and adaptive living biological systems. In this project, we aim to develop a unique platform that combines electrochemical microscopy (ECM), nano-DESI mass spectrometry imaging (nano-DESI MSI), hyperspectral optical microscopy (HOM), and super resolution fluorescence microscopy (SRF), thereby providing an unprecedented depth of information about molecules involved in communication within and between living cells in their native environment.

Nano-DESI enables minimally invasive quantitative MSI of hundreds of endogenous molecules in biological samples using localized liquid extraction. ECM is an emerging technique for imaging biological systems that provides several powerful modalities for contact-free mapping of the cell topography, studying protein dynamics at the nanoscale, and quantitative imaging of selected redox-active species in complex biological systems.

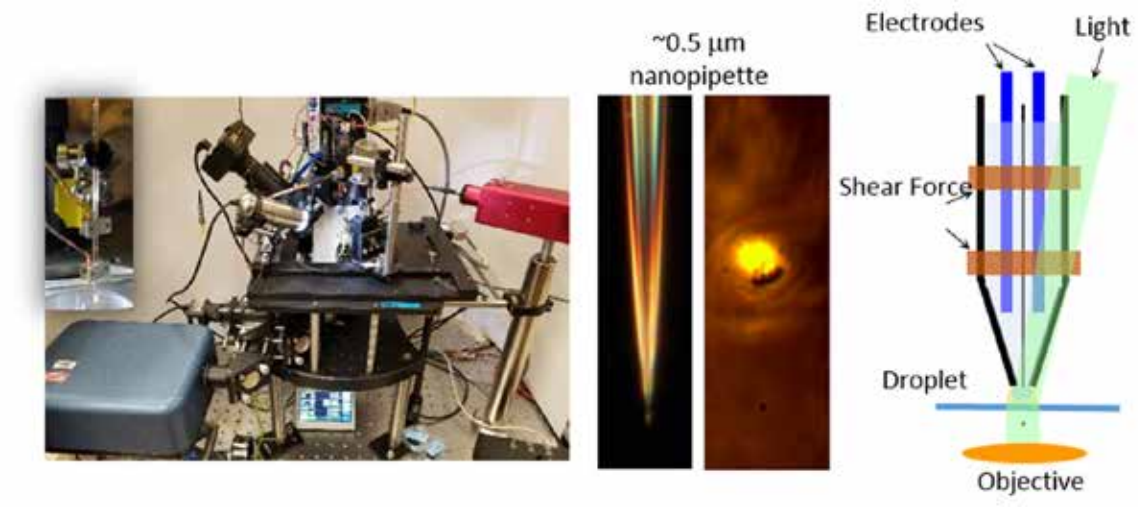
Coupling with highly complementary SRF and HOM techniques will enable quantitative imaging of endogenous chromophores using spectrally (377-1051 nm, $\Delta\lambda < 5$ nm) and spatially (sampled at 50 nm^2/pixel) resolved optical absorption microscopy.

In FY 2017, we have successfully combined the ECM and HOM modalities. This was achieved using a finely pulled, precisely positioned nanopipette containing an optical that facilitates surface topography mapping along with optical and electrochemical measurements. This system allowed us to obtain high-resolution images of stomata in the epidermal layer of *Portulacaria afra*. A substantial change in absorbance in and around stomata indicates the presence of different chromophores in these regions of the plant tissue. The metabolites around stomata also exhibit natural fluorescence.

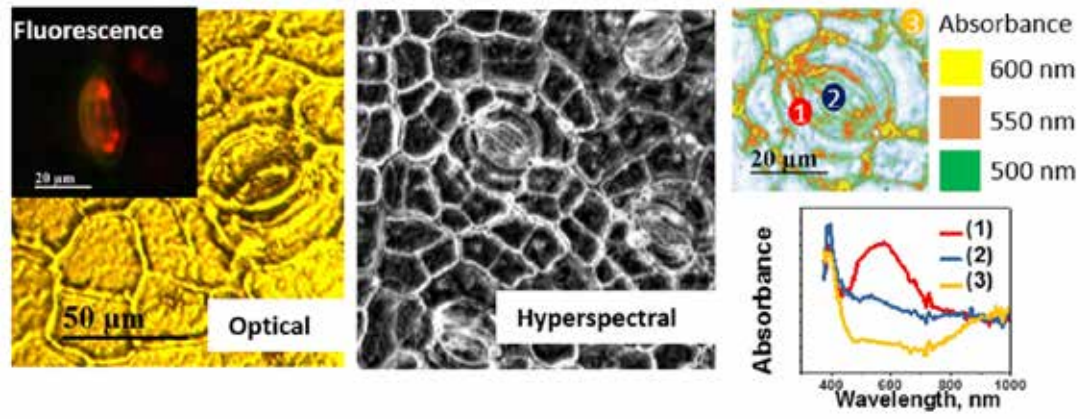
In FY 2018, we will complete the construction of the multimodal imaging platform by adding the nano-DESI modality, and we will develop data acquisition, integration, and analysis tools. The proposed work will provide initial insights on molecules involved in microbe-plant interactions.

Proof-of-principle experiments will be performed with model plant systems to demonstrate the capabilities of the multimodal system. Numerous metabolites will be spatially mapped using nano-DESI MSI. Meanwhile, ECM will provide information on the topographic changes in the regions of interest in the samples and inform on the dynamics of protein complexes. Finally, HOM will report on localization of distinct proteins, organelles, and subcellular eukaryotic structures in the living cells with nanometer precision.

PN17004/2894



Prototype of optical/electrochemical/shear force imaging platform.



Prototype of optical/electrochemical/shear force imaging platform.

Flexible Microbial Biogas Conversion Technology for Sustainable Animal Nutrition

Alexander S. Beliaev

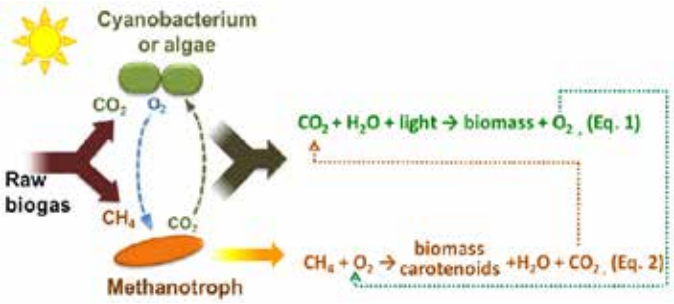
Anaerobic digestion (AD) is an attractive waste treatment technology; however, its economic benefits emerge only when the process is carried out at a large scale. Many AD systems used by the agriculture and food industries for waste treatment are relatively small, thus forcing the operators to burn the produced biogas for heat. Our technology extracts high-value products from biogas, thus improving the feasibility of waste treatment using AD and capitalizing on the abundance this renewable resource.

Biogas, which is a mixture of CH_4 and CO_2, represents an important renewable energy source that can be upgraded to a gaseous transportation fuel or combusted to generate electricity. However, because of its low pressure, high upgrade costs, and presence of contaminants, a large fraction of the produced biogas is being flared. To capitalize on the abundance and low cost of this renewable resource, new approaches are needed that can use biogas as a feedstock for production of high-value products.

Our technology is capable of efficient capture and conversion of raw biogas into high-value nutritional compounds, such as single-cell protein and carotenoids, using engineered methanotroph- algal co-cultures. At the heart of our technology lies metabolic coupling that involves mutualistic cooperation between oxygenic phototrophs and methanotrophic bacteria, whereby the CH_4 uptake and conversion to biomass and CO_2 is enabled by photosynthesis. Successful implementation of the technology has a potential to significantly improve the economics and broaden the application of AD through effective capture and conversion of biogas into value-added products.

Our group has previously demonstrated a successful metabolic coupling between *Methylomicrobium alcaliphilum* 20Z and a unicellular cyanobacterium *Synechococcus* sp. PCC 7002, which grew as a co-culture under photosynthetic conditions using methane and carbon dioxide gas mixtures of varying compositions. Here, we expanded the range of organisms capable of growing in a co-culture and optimized out binary system for the production of single-cell protein and beta-carotene. Our initial efforts focused on evaluating the growth rates and biomass yields using small-scale batch cultivation. Secondary screening targeted total protein and carotenoid contents as a function of nutrient availability, irradiance, and salinity.

As a result, we assembled and optimized a new binary culture consisting of *Methylomicrobium alcaliphilum* 20Z and a cyanobacterium *Spirulina platensis* UTEX LB 2340, which demonstrated ability to efficiently grow on a methane/carbon dioxide gas mixture at high pH, thus dramatically reducing the risk of biological contamination. The assembled co-culture not only demonstrated stability and robustness through a 3-week continuous cultivation in a 7-L photobioreactor but also accumulated of high protein and beta-carotene concentrations. The achieved biomass productivity of the co-culture was more than 35 mg L^{-1} h^{-1} compared to only 2 mg L^{-1} h^{-1} when the cyanobacterium was cultivated as a mono-culture.



Schematic representation of the co-cultivation approach that utilizes partnering between microalgae or cyanobacteria and methanotrophic bacteria to capture mixed CH_4/CO_2 gas streams. The metabolic synergism arises from the concurrent metabolic activities: cyanobacterial light-driven reduction of H_2O to fix CO_2 and production of biomass and O_2 (Eq.1) and aerobic oxidation of CH_4 to CO_2 by methanotrophs, which derives reductant for biomass production. Because the end products of oxygenic photosynthesis (e.g., O_2, see Eq. 1) and aerobic CH_4 oxidation (e.g., CO_2, see Eq. 2) serve as substrates for each other, growth of each organism in the co-culture is balanced and, ultimately, self-regulated.

PN17053/2943

Furthermore, we uncovered a critical dependence of the *Spirulina* growth and photosynthetic rate on the availability of vitamin B_{12} (cobalamin). Because *Methylomicrobium* naturally produces this vitamin, albeit at low quantities, the co-cultivation of these organisms provides an avenue to further enhance the co-culture yields. Our strategy for that included genetic engineering of the methanotrophy for enhanced B_{12} production via deregulation and overexpression of the cobalamin (*cob*) biosynthetic pathway. Specifically, we used targeted mutagenesis to delete the RNA-based regulatory sequences suppressing the B_{12} biosynthesis; simultaneously, additional copies of the *cob* operon were introduced to *Methylomicrobium* under the control of a high activity promoter.

Successful implementation of the technology has a potential to significantly improve the economics and broaden the application of AD through effective capture and conversion of biogas into value-added products. Valorization of biogas will pave the way toward broader deployment of AD for efficient capture of stranded waste carbon and its conversion to value-added products.

Fungal Solid State Fermentation for Citric Acid and Enzyme Co-Products that Derive Value from Agricultural Waste

Jon K. Magnuson

Aspergillus (A.) niger ***is a fungus that easily converts sugars from corn or sugar cane to citric acid to supply the large existing global citric acid market. We are developing a low-cost solid state fermentation capability using advanced strains of*** **A. niger** ***to produce citric acid and valuable enzymes from agricultural (food processing) waste found in hundreds of locations across the United States, rather than valuable sugars.***

Aspergillus niger is a fungus that has been utilized for 100 years to convert sugars, such as glucose and sucrose, to a valuable product found in soft drinks and other products, citric acid. This fungus is also used by industry to produce other valuable products, enzymes that can break down non-food plant biomass into sugars. This project aims to develop high-performance strains of *A. niger* for use in solid state fermentations, which means the fungus grows on top of the moist solid plant residues (food waste) instead of in large, vigorously stirred, and aerated tanks of sugars found in submerged fermentations.

The value of the solid state method is that it does not require a lot of expensive equipment like submerged fermentations, and can be established at sites where relatively small amounts of plant waste are generated, like food processors. This translates to economic and environmental benefits by using these waste streams that would be a disposal cost to, instead, generate valuable products. Our advanced strains of *A. niger* produce higher levels of citric acid on these food processing waste substrates and do not require addition of valuable sugars or other nutrients.

We have obtained a variety of agricultural residues from the local fruit juice and winemaking industry, namely, grape and apple pomace. These are the solid, fiber-rich residues (basically, pulp) remaining after the juice has been extracted from the fruit. These residues have low to no value in their current state, so they would represent an inexpensive substrate (source of carbon, nitrogen, and other nutrients) for a bioprocess such as the citric acid process we have proposed.

We have tested three *Aspergillus niger* strains for growth and production of citric acid on apple pomace hydrated with varying amounts of water, citric acid production medium, or only the salts (nitrogen and trace elements) found in citric acid production media. The three strains of *A. niger* that have been tested include two proprietary PNNL strains, one of which has a specific gene deleted and the other a specific gene inserted. The isogenic control for the study (aka the parental strain) is *A. niger*: Δ*kusA*, which has the *kusA* gene deleted, thus allowing targeted gene integration or deletion. This is a necessary precursor strain in order to create the other engineered strains.

We have made two exciting and practical discoveries. The first discovery is that our gene over-expression strain performs better than the control strain with respect to citric acid production under these solid state fermentation conditions, so we have a high-performance practical strain. The second discovery is that the pomace and water cultures performed just as well as those with added nutrients. This means that no expensive nutrients need to be added, thus making the process more economical. The fact that we only needed one gram of water per gram of substrate (pomace) is also important, since water usage is an economically and environmentally expensive part of such processes, so our minimal water usage is very encouraging.

In the first two quarters of FY 2017, we focused on finishing this project by optimizing the conditions for production of citric acid by examining additional culture configurations and the addition of up to 50% more water, but with no added nutrients. We also examined growth on grape pomace in parallel with apple pomace. Finally we examined the production of valuable enzymes at the same time as citric acid being produced. These enzymes are important for breaking down biomass, such as pomace, and increasing the ability of the fungus to efficiently utilize the carbon. The enzymes are also a potential product for the very industries from which the agricultural residues originate, so a valuable loop in the process can be created.

PN17082/2972

Fungal Synbio Platform for Plug-and-Play Consortium Engineering

Alexander S. Beliaev

Current metabolic engineering strategies for conversion of lignocellulosic feedstocks are based on single-organism chassis modified to reduce sugar preferences. Our strategy is to improve conversion efficiencies through simultaneous use of pentose and hexose sugars from the deconstruction stream via the division-of-labor strategy. The outcome of this project paves way for the development of "designer" consortia capable of single-step conversion of lignocellulosic feedstocks into a range of biofuels and bioproducts.

The ability to simultaneous convert multiple carbon substrates present in lignocellulosic deconstruction streams is a highly desirable but extremely challenging goal, as it largely relies on the metabolic capacity of the production chassis. Deconstruction products resulting from conventional acidic thermochemical pretreatment are complex and contain pentose and hexose monosaccharides, polysaccharides, and inhibitors (furfural, hydroxymethyl furfural, aliphatic acids, and phenols).

Existing production platforms do not simultaneously use pentoses and hexoses in an efficient manner, and most of them are sensitive to contaminants and inhibitors found in the deconstruction stream. While genetic engineering can improve productivities of the microbial chassis, these modifications come at the expense of the organism's fitness and viability, with many production strains dependent on stringent cultivation environments.

To that end, conversion systems based on multi-organism consortia, which feature division of labor through metabolic specialization, have a tremendous potential to efficiently transform complex sugar and aromatic streams into a broad portfolio of specialty fuels and commodity products. However, biotechnological application of consortia is limited by the availability of organism chassis that are not only metabolically versatile and physiologically robust but also highly amenable to genetic manipulations for improved control of cellular functions and behavior. Therefore, the main objective of this proposal is to develop a highly tractable fungal chassis that is capable of producing a range of specialty fuels and value-added products from deconstructed lignocellulosic material, thus enabling the design of innovative, consortial-based bioprocess technologies. Successful execution of the proposed research tasks will lead to important outcomes, which include 1) improvements in lignocellulose conversion rates and efficiencies through simultaneous use of both pentose and hexose sugars founds in the deconstruction stream and 2) enabling development of "designer" consortia capable of single-step conversion of lignocellulosic feedstocks into a range of biofuels and bioproducts.

As our ultimate objective is to develop and test the performance of fungal consortia for production of fuels and chemicals, we first constructed strains that can be used as chassis for building a prototype co-culture. We chose to engineer two fungal chassis for the production of 3-hydroxypropionic acid (3HP). The first one, *Yarrowia lipolytica*, was chosen for its genetic tractability, allowing us to rapidly test the designed 3HP pathways and its growth as a haploid single-celled organism, opening the door to future adaptive evolution experiments.

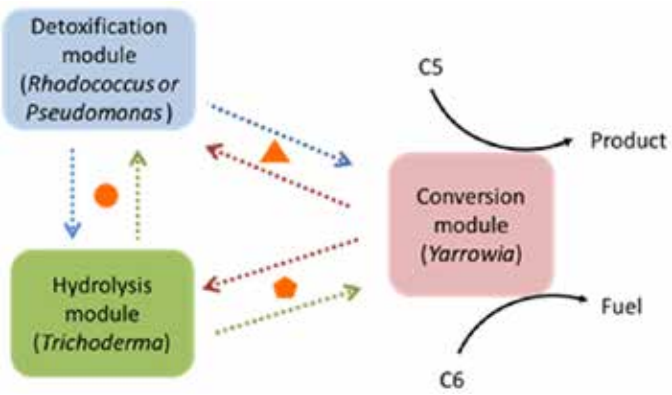
The second chassis, *Aspergillus pseudoterreus*, was chosen as a real-world chassis with the potential to make higher titers of 3HP in an acidic environment. This filamentous fungus is less genetically tractable and transformation tools and protocols needed to be developed. To support this effort, we designed and constructed novel, high-level expression constructs capable of targeted genome integration and tested transformation and selection of these plasmids via the protoplast method.

To enhance the selection and optimization of production strains, we also have engineered a metabolite sensor circuit that enables the detection of 3HP. To enable expression, we also incorporated a nuclear localization signal via site-directed mutagenesis at the beginning of the LysR gene. These new constructs provide an ability to monitor the production of 3HP by our engineered strains and facilitate optimization of yields and titers via directed evolution.

Finally, to test the design of consortia, we used the filamentous fungus, *Neurospora crassa*, to take advantage of existing

PN17044/2934

mutants. Several different strategies were employed for constructing strains incapable of utilizing pentose and hexose sugars and included mutants with inactivated sugar transporters, sugar-acting enzymes, or carbon catabolite repression machinery. As a result, we isolated several *Neurospora crassa* strains with considerably decreased xylose metabolism by removal of enzymes that allow xylose to be broken down. Similarly, we have targeted enzymes in glycolysis using crossing techniques to inactivate glucose catabolism. The resulting strains with preferential utilization of pentose or hexose sugars provide a basis for testing the efficiency of consortia-based approach to conversion of acid-hydrolyzed lignocellulosic stream.



Conceptual design of modular multi-organism conversion platform where detoxification, hydrolysis, and product synthesis are carried out separately by highly specialized organisms. The division of labor across the organisms allows direct conversion of acid pre-treated lignocellulose feedstock into a range of desired products.

GoBrachy – Developing a Metabolite-Trait Association Network Model for Carbon Allocation in *Brachypodium*

Christer Jansson

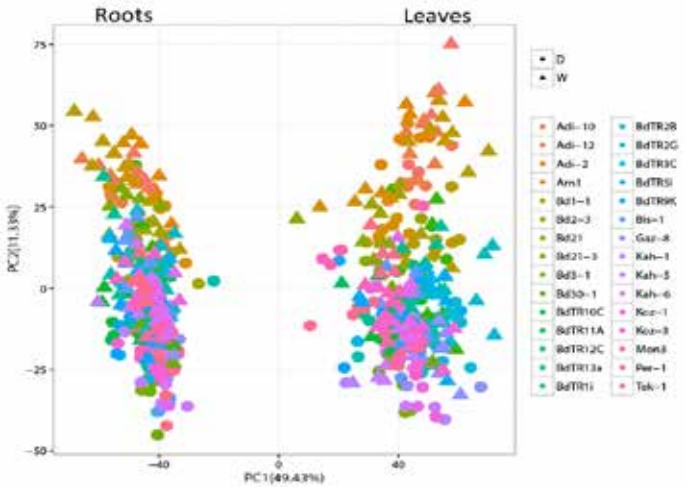
We are building capabilities to predict the trait and responses of model grasses and energy crops in ecosystems to environmental perturbations using high-coverage metabolomics.

The main goal of this work is to develop a Metabolite-Trait Association Network, a predictive model that links metabolite profiles in the grass model plant *Brachypodium* to quality traits, such as aboveground biomass (AGBM) and belowground biomass (BGBM), using high-coverage metabolomics. This will enable the predictive understanding of how the quality traits of plants are shaped as a function of plant genotype and its environment.

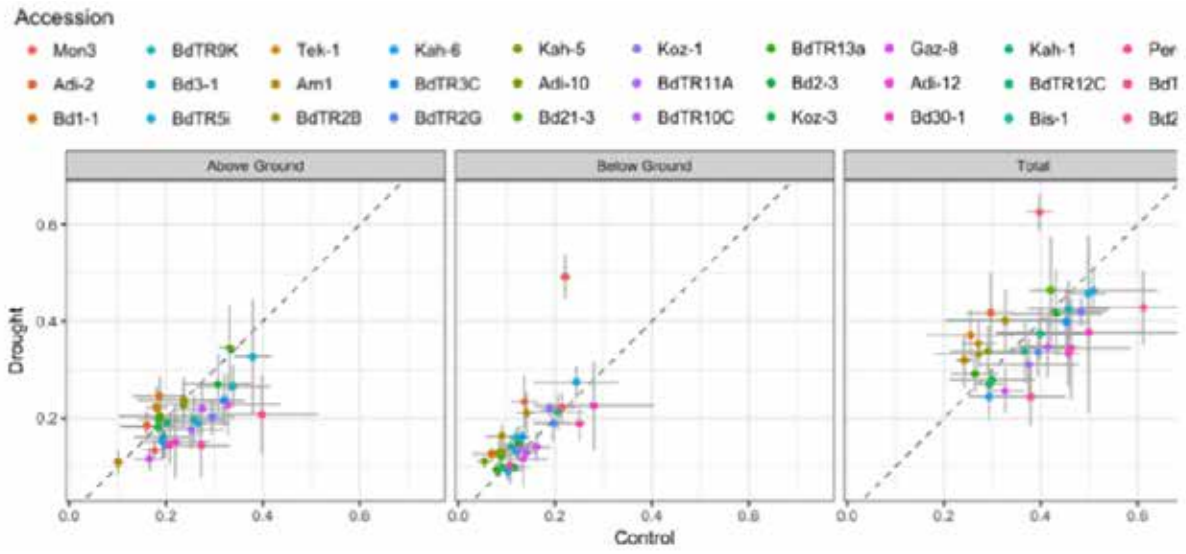
GoBrachy experiments were designed to provide 1) a mechanistic understanding of genotypic diversity for trait and responses, 2) an understanding of how the plant genotype interacts with the environment to inform emergent traits and responses across scales—from molecular to plant to the integrated plant systems level, and 3) a predictive, mechanistic understanding of ecosystem responses to environmental perturbations.

In the last quarter of FY 2016 and first quarter of FY 2017, we completed the screening of 30 *Brachypodium* lines selected for this project under five environmental perturbations: non-perturbed conditions (control), drought, elevated atmospheric CO_2 levels, increased temperature, and elevated CO_2 levels + increased temperature. Dry AGBM and BGBM were recorded for each of the five perturbations.

A large-scale, non-targeted metabolomics study was performed with the samples harvested from control and drought treatments. Five biological replicates of both leaf and root metabolites of the 30 lines were analyzed by both gas chromatography and liquid chromatography mass spectrometry.



Principal component analysis of the total metabolomes of roots and leaves of the 30 Brachypodium genotypes.

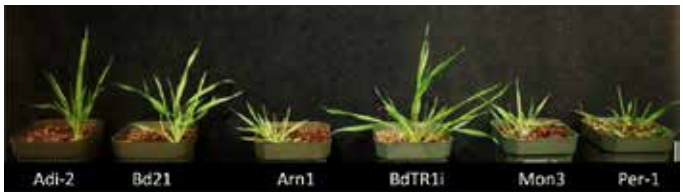


Dry biomass distribution of the 30 *Brachypodium* genotypes under drought and control conditions.

PN16098/2875

This revealed a clear metabolite shift following drought treatment and a clear separation between organ-specific metabolomes, regardless of the genotype.

Based on the dry biomass data, six of the most interesting lines with striking and opposite responses to drought were selected for detailed characterization.



Above ground phenotype of the six selected Brachypodium genotypes.

The drought experiment was repeated with the selected six lines and tissue harvested for proteomics and transcriptomics studies. Proteins were extracted from leaves and roots of drought and control samples in triplicate from the six lines. These samples were labeled with isobaric tags for relative and absolute quantitation (iTRAQ) reagent in 8-plex iTRAQ experiments. Labeled samples are ready for mass spectrometric analysis, and the resulting data will be used for global proteomics.

PN16098/2875

Gut-on-a-Chip for Multi-Omic Studies of the Gut Microbiome

Joshua N. Adkins

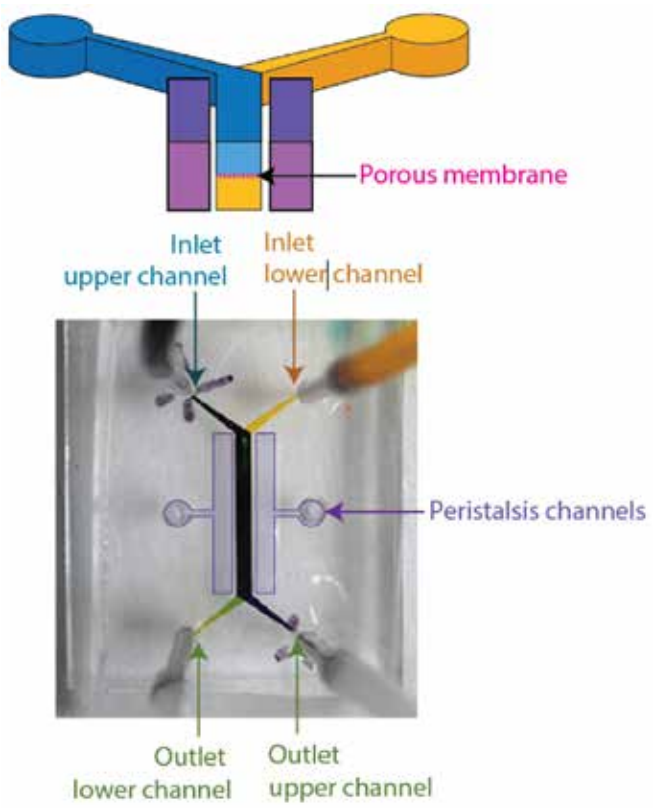
There is a need for reproducible, comprehensive characterization of biomolecules from complex microbiome research. The development of a gut-on-a-chip with coupling to biomolecular "omics" measurements will allow for accelerating hypothesis development and testing for gut microbiome research.

Alteration of the human gut microbiota composition and function has been associated with many human diseases. The impacts of perturbation resulting from pathogens, antibiotics, or toxins on the microbiome, despite great interest, remain largely unknown.

A critical challenge to effectively study the effect of perturbations on microbiota is the lack of well-controlled methods to sustain bacteria in direct contact with a living epithelium. The "gut-on-a-chip" is a microfluidic technology that provides a microenvironment mimicking many aspects of the gut's environment *in vitro* and in a small, controllable platform.

During FY 2017, we designed our first microfluidic gut-on-a-chip devices based on the gut-on-a-chip previously designed by Donald Ingber group at Wyss Institute. Our chips comprise key design changes to address several of our specific needs. Specifically, the gut-on-a-chip device is fabricated from polydimethylsiloxane (PDMS) using a soft lithography technique.

The chip design consists of three layers. The upper and lower layers contain flipped microchannels for supply of cells and regents, as well as aligned reagent chambers for exchange of medium across the membrane. The middle layer has a PDMS membrane containing circular pores. The upper and lower layers also have hollow vacuum chambers on either side of the reagent chambers for creating peristaltic motion of membrane. Two sets of Tygon® tubing are used to control the flow of cell and culture medium and for the pumping and evacuating vacuum chambers. The cell and culture medium flow and air for peristaltic motion of membrane using vacuum chambers is computer-controlled through in-house developed software for *in vitro* mimicking of peristaltic motions and fluid flow inside a gut.



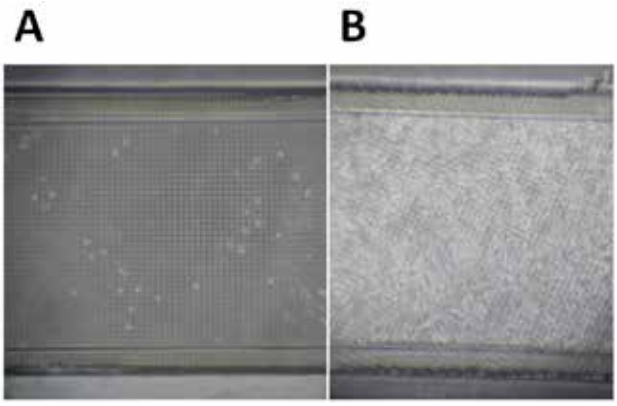
Design of the microfluidic gut-on-a-chip device. The top panel shows a vertical section of the microfluidic device comprising four microfluidic channels: two full height side channels (purple) surrounding a central growth channel split in two (blue and yellow) by a microengineered membrane (deep pink). The bottom picture is a top view of the device and shows food dye exchanged between the upper and lower microfluidic chambers.

We have tested multiple options to control the fluidic fluxes in the central microchannel. We first tried to use pressure-induced fluxes; this method resulted in overpressure in parts of the device, causing cells to burst and subsequently impairing cell-growth. This method was replaced by the installation of a syringe pump directly within the cell-culture incubator facilitating the simultaneous control of the media temperature, gas ratios, and media flux.

PN16108/2885

In order to allow the attachment of the cells to the microengineered membrane and to induce cell polarization, the microengineered porous membrane is coated with an extracellular matrix solution. We have improved the coating protocol. In particular, we have included a step of functionalization of the PDMS plastic with (3-Aminopropyl) trimethoxysilane to enhance the attachment of the extracellular matrix solution. The protein composition of the extracellular matrix was optimized to increase cell adhesion.

We are now able to grow Caco-2 Bbe1 cells within the chips, and we will soon characterize the tissue formed in presence and absence of peristaltic movement using automated nano-proteomics technologies that we have developed.



Demonstration of growth within the chips. Panel A is a micrograph of the chip membrane right after seeding. Panel B shows the chip membrane 48 hours after seeding.

PN16108/2885

Hydraulics of the *Brachypodium*-Root-Soil System Under Variable Environmental Conditions

Thomas W. Wietsma

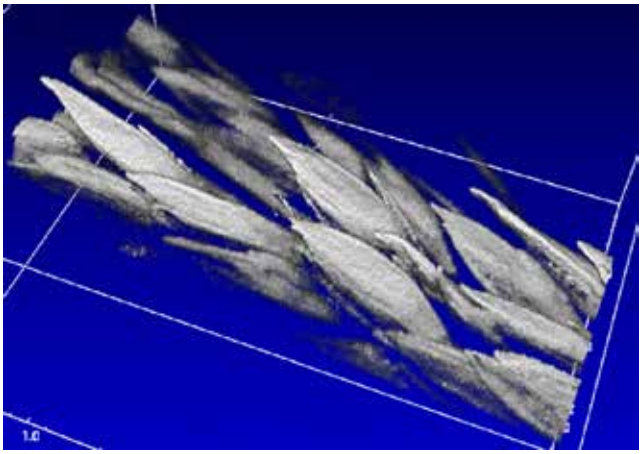
The goal of this project is to develop experimental capabilities to study soil-plant hydraulics in combination with root system architecture (RSA).

Numerical models for root water uptake in soil-plant systems have been developing rapidly, increasing the demand for laboratory experimental data to test and verify these models. Most of the increasingly detailed models are either compared to long-term field crop data or do not involve comparisons at all. Ideally, experiments would provide information on dynamic RSA in combination with soil-plant hydraulics such as water pressures and volumetric water contents.

The objectives of the research are to 1) support the development of the Virtual Plant-Atmosphere-Soil System (vPASS2.0) model with the development of novel laboratory rhizo-columns, optical coherence tomography (OCT), and improved X-ray-computed tomography (XCT) imaging; and 2) provide hydraulic and chemical characterization of the soil being used in a variety of different experiments. The vPASS2.0 development will focus on *Brachypodium sp.* as a model for bioenergy grasses.

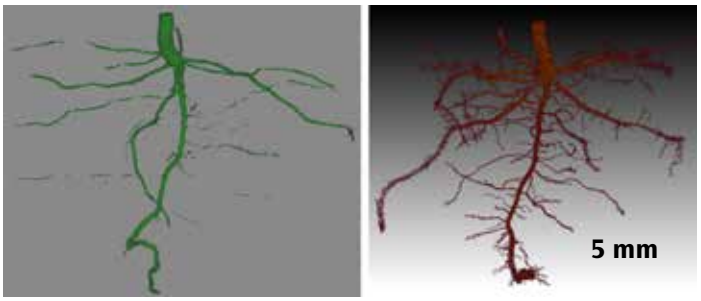
The novel rhizo-columns developed allow for a wide range of measurements, including temperature (thermistors), water pressure head (tensiometers), and moisture content (spectral induced polarization). Gas samplers allow for the measure of the impact of hydraulic redistribution on the microbial community by measuring the soil respiration rate at different locations. The rhizo-columns are equipped with polymer tensiometers with the ability to measure soil water retention relations all the way down to the permanent plant wilting point. Furthermore, water flow into the rhizo-columns is controlled by imposing a constant head boundary at the bottom that is in equilibrium with evapotranspiration from the top.

OCT capability was developed to provide non-destructive visualization of fine features of *Brachypodium* roots, stems, and leaves in order to identify phenotypic traits associated with growing conditions. OCT was used to resolve surface and sub-surface features of plants with 3 µm vertical resolution (4–8 lateral µm resolution). OCT imaging plant structures was shown to be compatible with the laboratory rhizo-columns.



Three-dimensional depth-resolved images of *Brachypodium* seeds.

The rhizo-columns were developed to be easily transferred for XCT scanning. Enhanced segmentation capabilities in soil/root imaging (extraction of root image and metrics) were developed by adapting two new segmentation tools: Waikato Environment for Knowledge Analysis segmentation in ImageJ and the Region Growing Tool in VGSTUDIO MAX. This capability allows for more accurate results in less time.



Improved XCT imaging.

To support the development of the vPASS2.0 Soil Hydrobiogeochemistry Module, the hydraulic and chemical properties of the selected soil were experimentally obtained. These properties consist of the particle size distribution, porosity, hydraulic conductivity, soil water retention, carbon/nitrogen ratio, pH, cation exchange capacity, and elemental nutrients.

PN17007/2897

Impact of Environmental Stressors on 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.

The concepts of stability and resilience are key to promoting ecosystem sustainability in the face of significant climate change. The same fundamental principles are relevant to the maintenance of homeostasis (health) in complex organisms subjected to environmental perturbations. Our objective is to specifically target the flow of information, from the genome to the functional level defined by proteins and metabolites, with emphases on the dynamic response to environmental perturbations and the communication mechanisms that allow complex systems—whether communities of microbes or organized tissues in multi-cellular organisms—to adapt to and mitigate perturbation.

We will achieve these goals by applying systems biology approaches to quantitative data generated with unique PNNL instrumentation in MS, nuclear magnetic resonance, and imaging to understand the effects of environmental stressors on multi-cellular systems. We will select novel model systems that represent unique challenges in understanding the effects of environmental stressors on the health and resilience of complex biological systems.

Use of synthetic peptoids to study intracellular trafficking by single particle tracking. We designed and synthesized photonic nanomembranes with different emission wavelengths (from approximately 400 nm to 780 nm) and synthesized peptoid nanomembranes that contain small-molecule sensors suitable for H_2S sensing. Peptoids with various charges and folic acid densities were tested for their influence on cellular uptake pathways and kinetics, as well as the ability of nanomembranes to escape the lysosome, using single particle tracking techniques. The resulting highly photostable nanomembranes with enhanced lysosome escape will be further tested for RNA delivery. One peptoid was constructed to contain an extracellular matrix binding domain (RGD), assembled into nanotubes and tested for cellular adhesion and uptake. By exploiting the side-chain diversity of peptoids, we have also synthesized peptoid particles for RNA delivery. Our preliminary results show that fluorine-containing peptoid particles exhibit enhanced efficacy of RNA delivery, a key capability for synthetic biology applications.

Evaluating the effect of dietary carbon sources on gut microbiome communities. We are investigating the effect of the Specific Carbohydrate Diet (SCD), which restricts the types of carbohydrates consumed, on the metabolic action of the gut microbiome in patients with inflammatory bowel disease. A pilot study has been performed, which evaluated clinical presentation and the microbiome community composition before and after adoption of the SCD, and shows marked improvement of symptoms. To investigate corresponding changes in the metabolic function of the microbiome, metabolomic and proteomic analyses are being performed on these same samples. Data collection (supported by subcontract) is ongoing.

To build a robust capability at PNNL for analysis and interpretation of the anticipated microbiome multiomics data, we have performed a literature review identifying the state of the field for human gut microbiome research relating to diet and inflammatory bowel disease, as well as the most current protocols and techniques for constructing comprehensive references databases for optimizing analysis of proteomics data and integrating proteomic and metabolomic data. Because human microbiomes are highly variable between individuals, we are building a workflow that will facilitate construction of reference databases that are customized to individual samples. This will increase both the amount of information we are able to extract from a sample and the accuracy of that information.

Establishing the CRISPR/Cas system for manipulation of signal transduction pathways. This task established the knowledge base for transforming mammalian cells with Clustered Regularly Interspaced Short Palindromic Repeats activating (CRISPRa) components against specific gene targets and developing a screening assay to determine the effects of CRISPRa transformed cells. PNNL-compatible protocols were developed from the published literature and empirical research. The primary goal achieved in this effort was a screening approach for determining the effects of CRISPRa

PN15016/2691

in a breast cancer cell line (Hs-578T) using the epithelial growth factor pathway as the model system. We selected known regulatory targets for this pathway and increased their expression using the developed CRISPRa techniques (expression levels were validated using quantitative polymerase chain reaction). We developed a flow cytometry based method to screen for activity of phosphorylated extracellular signal regulated kinase (pERK), a downstream target of each of the selected CRISPRa targets. We effectively demonstrated the coupling of CRISPRa transformation protocols with the flow cytometry screening method to assess pERK activity, thereby enabling sensitivity analysis of each selected regulatory element for this pathway.

Each of the tasks in this project has enhanced our ability to manipulate complex adaptive systems and accurately measure and analyze the response to perturbation. New tools in RNA delivery and gene editing were developed and applied to model microbial communities of relevance to both environmental and human health.

Integrated *In Situ* Chemical and Topographic Optical Imaging of Live Microbiomes in Transition

Curtis J. Larimer

A novel "bioscope" that combines white light interferometry (WLI) and hyperspectral optical microscopy was developed. This new imaging technology was used to simultaneously capture physical (three-dimensional [3D] topology) and chemical (spectrally resolved absorbance micrographs) images of microbes and their communities.

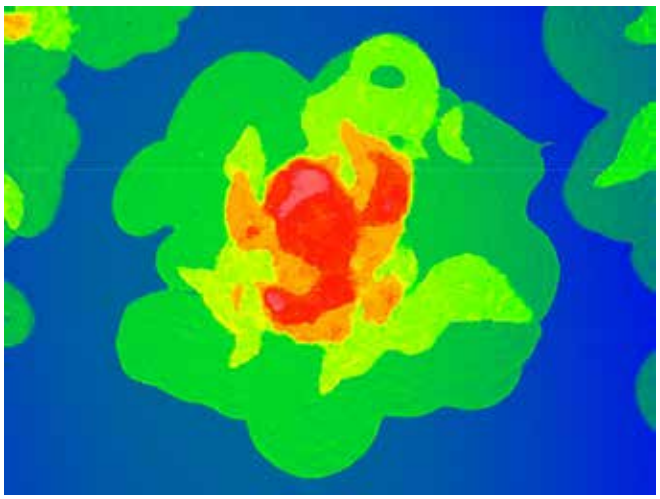
Many microbial communities form elaborate multicellular structures, but the rules governing their architecture are not well understood. Quantitative microscopy of microbiomes in transition requires next-generation instruments that are capable of generating label-free cell-level maps of host-microbe interactions in a non-destructive fashion. We worked to develop a unique capability for *in situ* imaging, allowing multidimensional topographic and chemical analysis of live microbiomes in transition over a broad range of length and time scales.

A new microscope with multiple imaging modalities was designed and built. WLI has the highest resolution of any optical imaging technique (3 nm axial resolution) and can measure volume changes with femtoliter precision, but its use in bioimaging has been very limited.

Hyperspectral optical microscopy was used to complement WLI, because it can assess chemical composition and—like WLI—is a non-destructive technique with a fast acquisition time that is well-suited to imaging unlabeled samples. This unique combination was used for non-destructive monitoring of microbial systems in order to identify the key transition period(s) in microbiome dynamics. The focus of this project was to demonstrate system capabilities with a simple biological system. The integrated microscope was validated with an organism known to develop multi-level topography as it grows (Bacillus thuringiensis). In the future, we aim to image more complex microbial consortia that are important to the environment and human health. An environmental flow cell was designed and built to enable imaging of bacteria in a minimally disruptive process.

During the course of this research, it became clear that the novel imaging capability that was developed could also have broader application beyond imaging microbes. We collected proof-of-concept data demonstrating the ability to observe fingerprints left on surfaces of interest. With further study, the technology could be used for non-destructive collection of forensic evidence. The wealth of information collected by sensitive 3D and hyperspectral information could also lead to an ability to deduce new information about a person of interest.

In the future, we aim to use the technology developed to better understand spatial variation in redox conditions in microbiomes, cell signaling in microbial communities, and growth rates of competing/cooperating cell groups.



3D image of a model microbiome growing on an agar plate.

PN16109/2886

Marine Mesocosms for the Study of Climate Effects on Eelgrass and Microbial Communities

George T. Bonheyo

This project included the design and construction of a coastal research platform consisting of multiple mesocosms to enable the parallel comparison and analysis of varied physical, chemical, or biological parameters on ecosystem productivity and resilience. An intense design phase was carried out with multiple researchers to create a unified platform for supporting fundamental through applied research with sediments, marine plants, and microorganisms in service of a diverse set of needs.

Eelgrass is the dominant dicot plant found in the coastal waters of the U.S. west coast and northeast coast.

Seven outdoor tanks were installed along the shoreline at the Marine Sciences Laboratory. An eighth tank, located indoors, provides even greater flexibility for controlling lighting, chemistry, and temperature.

The coastal zone is one of the least understood components of regional and global carbon and nutrient cycling and is poorly represented in ecosystem and integrated Earth system models. Throughout the northern hemisphere, eelgrass (Zostera marina) provides habitat and food, improves water clarity, and sequesters carbon. The experimental mesocosms were designed and constructed to allow the manipulation and control of populations, communities, and environmental conditions within a more complete and longer-lived experimental ecosystem than could be studied in the lab or in the field.

The tank systems, including plumbing, feed/mixing tanks (for control of water properties such as salinity, temperature, pH), and lighting systems (for indoor use), were designed to allow for maximum flexibility. A variable intensity, full-spectrum LED lighting system was constructed to accommodate work with an indoor tank where wider range of thermal, illumination, and chemical parameters could be controlled. A multiparameter, custom sensor system for the eelgrass tanks was designed for *in situ* monitoring of pH, dissolved oxygen, electrical conductivity, temperature, oxidation reduction potential, visible light spectrum, photosynthetically active radiation, turbidity, water flow, depth, and chlorophyll a/b concentration. Research projects using the tank system are examining eelgrass metabolomics, rhizome and sediment microbial community development following planting, and the relative contribution of the eelgrass and the epiphytic microbial communities it supports on carbon/nutrient cycling, as well as primary and secondary productivity. Data is logged locally to an embedded computer, with access via a Web portal to support offsite collaboration.

PN17071/2961

Microbiome-Exposome Interactions

Aaron T. Wright

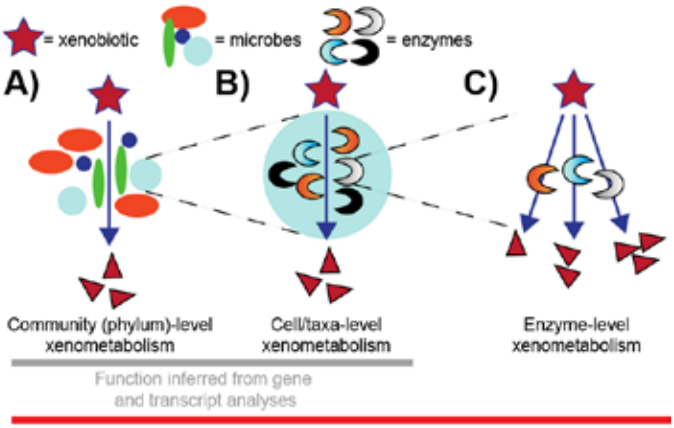
This project is addressing key gaps in our understanding of how the composition and function of microbial communities (microbiomes) are impacted by exposures to environmental agents and how these changes impact host susceptibility to the agents.

The broad diversity of microbes residing in the body greatly expands the synthetic and metabolic capacities of the human genome. The microbiome produces essential amino acids and vitamins, metabolizes otherwise indigestible dietary compounds, and shapes patterns of host immunity and tolerance. Emerging studies also show that the microbes within the intestinal tract can mediate the chemical transformation of xenobiotics such as pharmaceuticals and environmental contaminants associated with energy production. The impact that these environmental exposures have on the microbiome and, in turn, how the perturbed microbiome influences host susceptibility to environmental agents, is poorly understood. Work in this area is complicated by the large number of microbial species and enzymes capable of performing certain metabolic functions. A better understanding of these functions will be critical to the development of microbiome-based therapeutic strategies. In order to gain this understanding, innovations in approaches to identify the specific proteins associated with a given biochemical activity are needed.

This work goes beyond current metagenomic strategies to gain a mechanistic understanding of the specific organisms and proteins involved in a given microbiome activity. To achieve this goal, we are employing activity-based probes (ABPs) and proteomics profiling, coupled with new computational approaches for prediction of microbiome functions. This unique capability allows us to address two interrelated aims: 1) determining the impact of environmental exposures (e.g., chemicals and diet) on the composition and diversity of gut microbiomes using mouse intestinal microbiomes as a model, and 2) developing and applying activity-based protein profiling (ABPP) methods to identify proteins that have metabolic activity toward foreign chemicals (xenobiotics) and the microbes that express these activities.

Characterizing gut microbiota metabolism using ABPs. Our previous work has yielded a pipeline for isolating, probing, and identifying microbes based on their capacity for a given enzymatic activity. We have now identified a suite of microbes that produce enzymes capable of reversing glucuronidation, a major step in xenobiotic metabolism. Through coupling cell sorting and DNA sequencing, we have unambiguously identified microbes that are responsible for this activity within the gut. This was the first instance isolating and identifying microbes based on a specific metabolic activity.

Tracking metabolic shifts in microbial communities using functional annotation. Our group, and others, has shown that upon perturbation of microbial communities, the burden for a specific metabolic activity can shift among organisms. The identity of the members that take on this responsibility and the molecular tools that they use to compensate for the loss of a given metabolic activity are not immediately apparent using metagenomics approaches. ABPP provides a direct functional link for the protein encoded by the often poorly annotated genomes of the microorganism housed in the gut. We are interested in perturbations caused by external factors, including diet, environmental exposure, disease, and drug administration.



Functional activity measured at all levels by Activity-Based Protein Profiling

We determined the effect that dietary fiber intake has on glycoside hydrolase activity in the gut microbiota. Dietary fibers are complex polysaccharides that cannot be degraded by the host, but can be degraded by microbial enzymes. We are currently focused on profiling how changes in diet influence the production of these enzymes by the microbiome and how this, in turn, influences metabolite availability for the host. This study is providing unique insight into how carbon availability influences enzyme activity within a complex commu-

PN15073/2748

nity, and is also the first demonstration of protein identification via ABPP within the gut microbiome.

In addition to nutrient availability, the microbial community living in the animal gut influences host physiology by providing important signaling cues. One such cue is provided by bile homeostasis. Bile acids are produced by the host and are continuously modified by the gut microbiota, which affects the net hydrophobicity of the total bile acid pool and also modulates host signaling pathways. Arguably, the most important way in which microbiota modifies bile is through de-conjugation of bile salts through bile salt hydrolase (BSH) activity, which is a prerequisite for all further modification by the microbe. BSH is contained in many widely consumed, over-the-counter probiotic formulations. Despite the potential therapeutic application of this enzyme, there is no simple assay for continuous monitoring of BSH activity, and there are no non-destructive means of characterizing its activity in whole cell or microbial community samples. To address this gap, we have developed a novel, continuous fluorescence assay that can be used for characterization of BSH activity with purified protein, cell lysate, whole cells, and in human gut microbiome homogenate. This assay is used to show the first *in vivo* characterization of BSH activity, as well as the first example of *in vivo* characterization of a small molecule BSH modulator. As the tools to study the microbiome become increasingly refined, so will our ability to understand and manipulate the relationship between host and microbial community.

Microbiome Models Across Scales – from Metabolism to Succession: A Framework for Modeling, Simulation, and Theory Development for Microbial Ecology

William R. Cannon

We are developing new computational and theoretical methods that are more predictive than the current state of the science to understand the adaptation, acclimatization, and response of microbial communities in a changing environment.

This project is modeling the dynamics of microbes in complex environments, soils, and the gut to understand how a changing environment affects the microbes residing in it and how these microbes feedback to and shape the environment.

This year, we demonstrated the relationship between kinetic formulations of biochemical reaction rates that use rate constants and a statistical thermodynamics formulation that uses chemical potentials, and presented a new theorem for coupled reactions based on chemical potentials, which can provide relative rates for all sequentially coupled reactions under all conditions, steady state or not. This new development obviates the need for rate constants in many applications and is enabling complex, large-scale biological systems, such as metabolism, to be modeled with kinetics and thermodynamics. This will result in more accurate predictions of microbial in changing environmental conditions and will shed fundamental insight into the thermodynamic coupling between living organisms and their environment.

This new framework allowed us to analyze data in a fundamentally different way, linking data analysis to the fundamental principles of microbial energetics. Using chemostat data, we characterized the energy and power generated by microbial cells as they grow and reproduce. A typical heterotrophic cell requires about 500 picojoules per cell to replicate and generates about 1 Watt/g as it grows. The power generated by bacteria is comparable to the most efficient fuel cells developed to-date. This information will be used in microbial models to determine which species adapt to new environmental conditions by out-competing their neighbors.

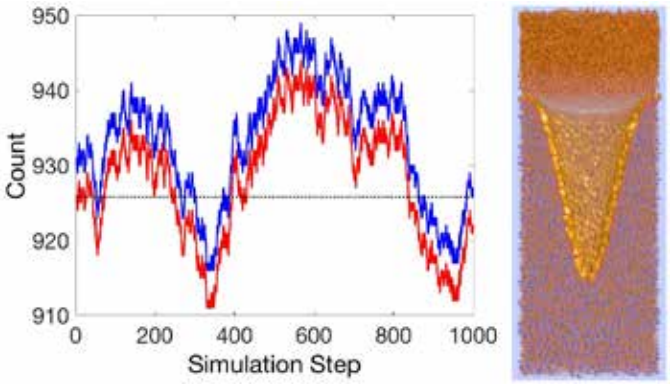
To model this competition, we implemented a series of functional guilds—models of microbes in which summary reactions represent their metabolic capacity—which grow aerobically on glucose, acetate, succinate, glycerol, and ferment glucose to succinate, ethanol, acetate, formate, and lactate. These models use the elementary modes of metabolism that capture 90% of the metabolic capability while, at the same time, being much simpler than genome-scale metabolic models.

We demonstrated the coupling of metabolism to the environment in a model of the gut microbiome. In these simulations, gut microbes produce short-chain fatty acids, such as acetate, which are taken up by the gut epithelial cells. The gut cells use short-chain fatty acids for growth, and the production of these short-chain fatty acids by the microbes determine the density of villi in the gut.

As a result of this project, we are ready to implement a realistic model of the dynamics of microbes as they adapt to changes in the environment and to model how these microbial transitions in turn modify the environment.

Select Sorbent Physical Characteristics

Growth Rate	0.11	0.49	PEMFC Fuel Cell
Power (Watts/cell)	5.6×10^{-14}	5.8×10^{-13}	NA
Power-to-Weight (Watts/g)	0.26	1.15	1.00



(Left) Statistical thermodynamics simulations (blue) produce the exact same trajectory as kinetic rate law equations. (Right) Simulation snapshot of gut epithelial crypt interacting with microbes.

PN15069/2744

Microbiome Responses to Hydrologic Regime Shifts and Subsequent Alteration to Ecosystem Function

Vanessa L. Bailey

The overarching goal of this project is to determine how changing redox conditions impact the coupling of soil biogeochemical cycles. The use of multiple 'omic analyses can reveal processes that are more sensitive to change than those observed through direct end-product measurements. Understanding these processes in soils under changing hydrologic regimes is critical to improving models of climate change effects on the land-climate system.

Soil microbiomes strongly influence the functioning of terrestrial ecosystems, from the pore-scale to the global-scale, by transforming materials in the environment. The availability of water strongly influences the rate and nature of microbiome-driven material transformations. Soil water content is a dynamic variable influenced by shifting hydrologic regimes due to natural variation (e.g., seasonal and weather-driven precipitation) and human-influences (e.g., land-use and climate change). Determining how soil water content dynamics—and changes in those dynamics—impact soil microbiome composition and function is fundamental to understanding global carbon and water cycles.

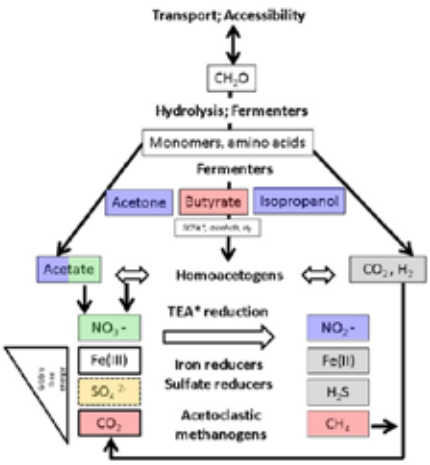
Changes to the water cycle alter the balance among aerobic and anaerobic metabolic pathways for CO_2 and CH_4 production, but the underlying dynamics and water sensitivities of the soil microbiome functions that control these net fluxes are currently unknown. We sought to mechanistically test the redox controls associated with coupled biogeochemical cycles on anaerobic carbon decomposition processes.

The research focuses on a tidally influenced wetland near the mouth of the Columbia River in Washington State. Climate change is altering the hydrologic regime in this, and similar, watersheds such that historically snowmelt-driven moisture inputs are being replaced by rain events, and the timing of snowmelt and precipitation is changing. Aquatic-terrestrial transition zones like wetlands will experience the immediate consequence of changes in the magnitude and dynamics of hydrologic flow. Coastal wetlands are particularly susceptible to storm events, salt-water intrusion and increasing water levels. Pulsing hydrology can activate and de-activate microbial populations and their functions that quickly alter the nature and rate of biogeochemical transformations; extended saturation conditions can drive methane production, whereas drying will expose soil organic matter to rapid microbial decomposition. As fluctuating redox conditions alter terminal electron acceptor availability and recovery strengths in soil, a disproportionate release of carbon dioxide can also stem from alternative anaerobic degradation processes like sulfate and iron reduction, in addition to methanogenesis.

We hypothesized that extreme moisture regimes will select for microbial taxa with life-history strategies adapted to desiccation or moisture stress. We conducted a microcosm study subjecting soils to drying (oxic) and saturation (anoxic) for an extended period (75 days) at 21°C to interrogate changes in the microbial community composition in the light of ecological theory and elucidate the relationship between community response and geochemical signatures. Only through 'omic analyses did we detect internal elemental cycles that were not evident in our physiological analyses.

Notably, rapid changes in hydrology increase substrate availability for both aerobic and anaerobic decomposition processes, and fluctuating redox conditions altered terminal electron acceptor and donor availability and recovery strengths. Relative changes in microbiome composition, in tandem with biogeochemical rates, indicate cryptic processes such as homoacetogenesis suppressed C fluxes out of the soil. We also found that disproportionate emissions of CO_2 came from alternative anaerobic processes, such as sulfate and iron reduction, compared to methanogenesis.

These internal biological cycles are not reflected in most models, yet disruption of these processes change net ecosystem function. Thus, resolution of microbiome processes is critical.



Carbon process responses to fluctuating soil moisture. CH_2O = organic matter; TEA = terminal electron acceptor; SCFA = short chain fatty acid; Blue = increased at the end of dry-wet cycle; Gold = increased at the end of wet-dry cycle; Red = increased under saturation; Green = no difference between dry-wet and saturation conditions. Grey = species were not measured.

PN15078/2753

Molecular Mechanisms of Drought Mortality and Survival

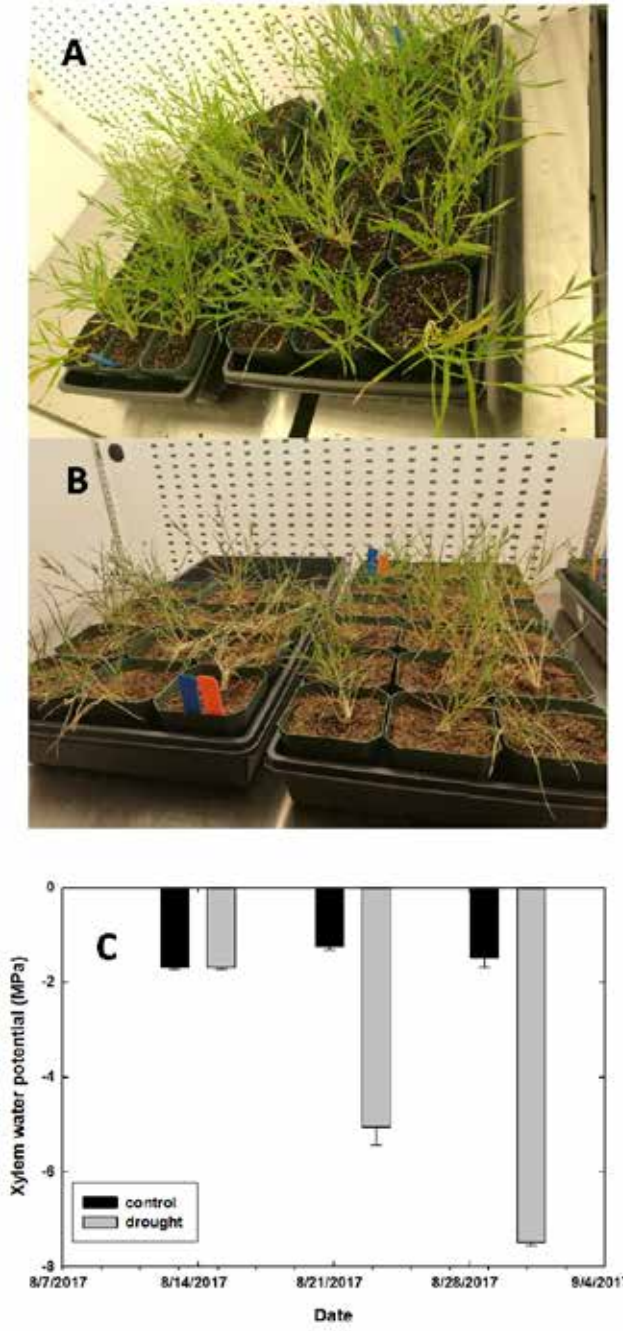
Nate G. McDowell

This project allows, for the first time, direct understanding of the molecular underpinnings of drought-induced mortality in agricultural species. This enables future application of omics solutions to minimize drought-induced impacts to yield and survival of crops critical to the United States and its citizens.

The carbon starvation and hydraulic failure theories advanced the field of ecology by stimulating intense research on the coupling of plant carbon and water metabolism. In parallel, omics work on model plants has led to great understanding of carbon and hydraulic metabolism in relation to stress, but mortality theory has not yet benefited from this type of study. There is great debate regarding the regulation of carbon and hydraulic metabolism during drought-induced death; these debates could be addressed with one experimental study combining drought-induced mortality with proper omics measurements.

This experiment has successfully established the protocols for growing, droughting, and killing *Brachypodium* grasses. All real-time measurements, specifically dry mass, anatomical data, and plant water stress, have been completed. Samples are now being prepared for omics measurements.

The most exciting result thus far is the extreme water stress that *Brachypodium* experienced while still surviving. This in itself is a large result, which will be enhanced in understanding and impact once the omics measurements are complete in FY 2018.



A) Well-watered *Brachypodium* (control plants); **B**) droughted *Brachypodium*; and **C**) water potential measurements (and standard errors) for the two treatments during the experiment. More negative values mean greater water stress.

PN17094/2984

Molecular Phenotyping of *Brachypodium* to Provide Metabolic and Functional Linkages to the Plant-Atmosphere-Soil-Systems (iPASS) Model

Kim K. Hixson

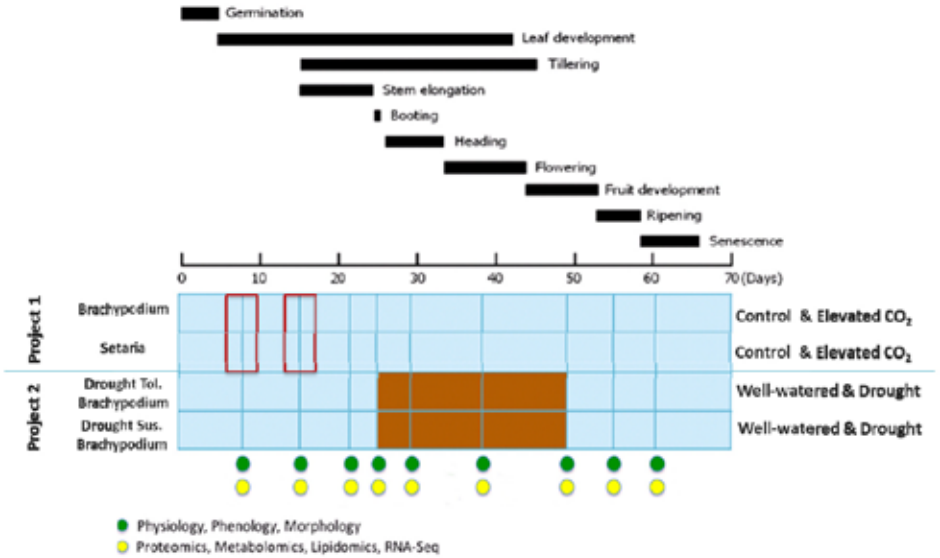
In this work, we will develop the fundamental molecular characterization necessary to build a model that ultimately links phenotypic measurements to the atmospheric processes and soil biogeochemistry associated with* Brachypodium *(C3 plant) and* Setaria *(C4 plant) during its lifecycle and under conditions of drought stress and increased atmospheric CO_2 concentrations.

During FY 2017, two plant growth and sampling projects were undertaken. Project 1 encompassed the growth and harvest of both a single reference line of *Brachypodium* (C3) and *Setaria* (C4) grown under ambient or elevated (800 ppm) CO_2 levels. Roots and leaves were harvested at each major developmental stage. After set-up and testing of plant lines and growth chambers, this project was started in mid-January, and we collected six developmental stages of *Setaria* and eight developmental stages of *Brachypodium*.

Project 2 involved the growth of *Brachypodium* and *Setaria* grown in collaboration with Michigan State University in their dynamic environmental phenometric (DEPI) chambers. The DEPI chambers were capable of non-destructive measurements involving growth and photosynthesis parameters. From Project 1 and Project 2, we extracted metabololites, proteins, and ribonucleic acid (RNA) from all samples. All proteomics data has been acquired, and data analysis is currently underway.

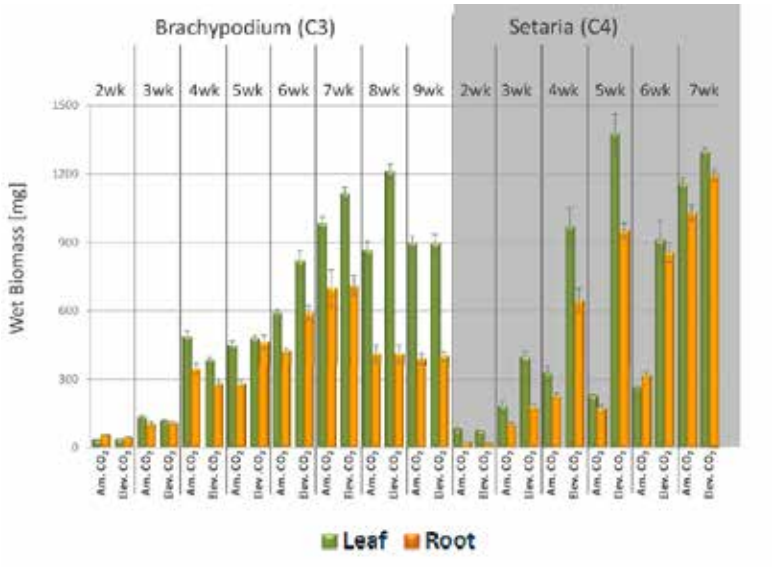
An important technological development is the further enhancement of a next-generation metabolomics platform. The SPE-IMS-MS platform was initially developed for mammalian system metabolites. We are building upon that established technology to enhance its capability to include plant and microbe secondary metabolites. In support of this goal, we have obtained secondary metabolite standards that represent major classes of plant secondary metabolism (e.g., phenylpropanoids, terpenoids, hormones, phenolics, glycosides, glucosinolates, fatty acids, alkaloids, wax/resin compounds). These compounds have been analyzed on the SPE-IMS-MS system to validate predicted mass, drift times, and collisional cross sections, which is an identifying dimension to this next-generation metabolomics system.

Additionally, we have constructed an *in silico* library using NWChem, which was used to calculate predicted collisional cross sections and accurate masses for experimentally verified and theoretically determined metabolites produced by *Setaria* and *Brachypodium*.



Growth and harvest plan for *Projects 1 and 2*. Red boxes highlight plants harvested to date.

PN17006/2896



Wet biomass of leaves and roots collected from the first two stages of *Setaria* and *Brachypodium* collected under ambient or elevated CO_2.

Monitoring of Terrestrial Aquatic Ecosystems with Hyperspectral Imagery

Luke J. Gosink

Hyperspectral cameras can record complex chemical and physical systems with tremendous spectral and spatial resolution. The proposed project will leverage and extend PNNL-developed analytics to provide a persistent, interpretable, and information-rich characterization of the wet/dry dynamics of terrestrial aquatic ecosystems.

Data from hyperspectral systems have been used to monitor a variety of application spaces, including evolving microbial communities in coastal and inland aqueducts, oil spills and cleanup, lakes and watershed basins, anticipating crop yield, and drought impacts on ecology.

However, despite the wide-spread use of hyperspectral cameras, contemporary uses are either overly general (e.g., clustering and segmentation strategies such as principle-component analysis) or overly specific (e.g., template approaches that rely on spectral libraries to support very narrow, myopic analysis of hyperspectral imaging data).

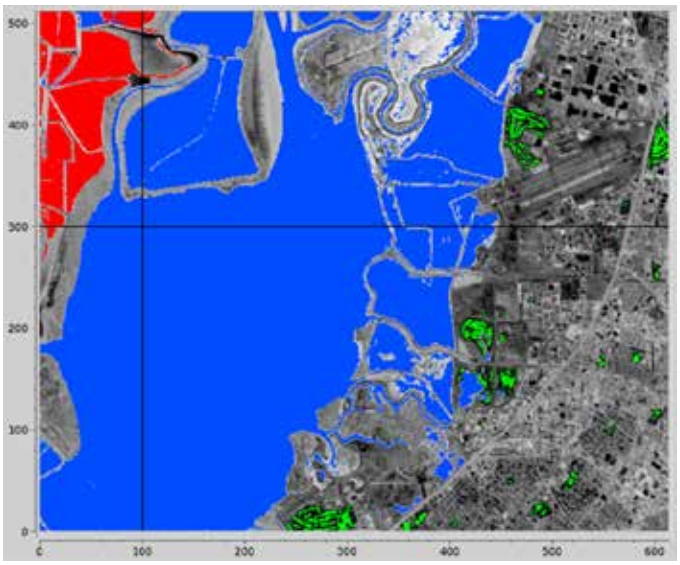
The proposed project will integrate proprietary methods in compressive sensing, statistical design of ensembles, and machine learning to provide a persistent, interpretable, and information-rich characterization of the wet/dry dynamics of terrestrial aquatic ecosystems. We will specifically answer the following questions: 1) What spectral signatures characterize and differentiate ecosystems that can adapt and remain healthy amongst increasing salinity and drought conditions? 2) How do these spectral signatures help validate (or identify new) biogeochemical rates, microbial functions, and vegetation physiologies for the mechanistic processes that model these ecosystems?

During FY 2017, we purchased hyperspectral instrumentation required to conduct lab experiments.

During FY 2018, we will complete a full spatial-temporal spectral characterization of *Brachypodia* lifecycle based on the following four types of conditions: healthy, drought conditions, salinated soil, and both drought and salinated soil.

We will capture hyperspectral imagery of both soil and plant during these studies to identify key spectral signatures that can indicate and differentiate the four different states.

Based on the success of this study, we will repeat these experiments in a larger, outdoor setting to test the developed signatures in a less controlled environment.



An example of a segmented hyperspectral image. Red regions indicate evaporation ponds filled with brine-shrimp, blue regions indicate water, and green regions denote golf courses (separate from lawns).

PN17104/2994

Multiplex Pathogen Detection

Janine R. Hutchison

This project uses PNNL's expertise in bacterial detection and microfabrication to develop detection methods for more than one bacterial pathogen. Multiplex pathogen detection can potentially reduce the time needed from sample collection to bacterial identification and can provide a more comprehensive approach to understanding disease potential.

Detecting and responding to the presence of bacterial pathogens or potential biothreat agents is important for understanding the breadth and impact of naturally occurring and intentional biological events, allowing an informed response to save lives, minimize impact, and monitor situations effectively. Current technologies for detecting and identifying bacterial targets from indoor environments require trained personnel, labor intensive sampling, tedious sample preparation, and hours to days to identify and quantify organisms. Therefore, improved recovery, quantification, and detection technologies are needed to effectively respond to an unknown biological agent(s) of interest.

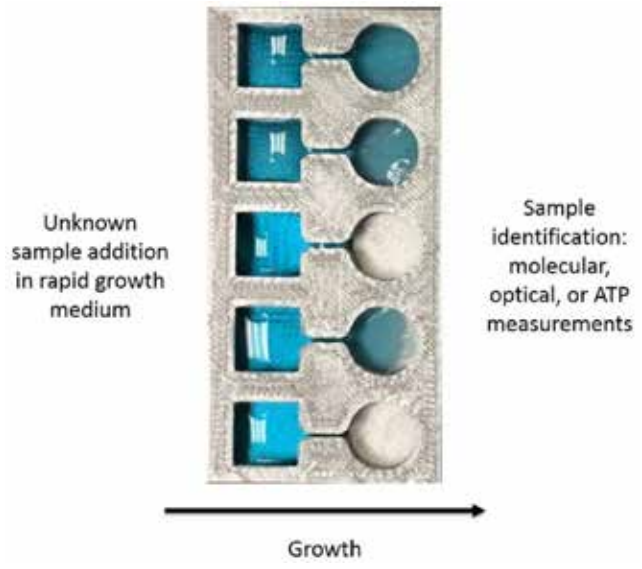
A multidisciplinary team was assembled with expertise in rapid bacterial growth, cellular separation and concentration, microfabrication, and organism identification to address the technical challenges surrounding multiplex pathogen detection.

We have focused our efforts on the development of a multiplex pathogen detection device that will execute the following steps: 1) separate and concentrate the bacterial cells from a complex sample such as a swab or a wipe, 2) increase the total number of cells to meet the detection limits using PNNL's custom rapid bacterial growth media (CLEM), 3) quantify and determine the viability, 4) provide a presumptive bacterial identification, and 5) report the results to the user.

We have demonstrated that a microfluidic incubation device can be 3D-printed rather than produced by photolithography. Utilizing 3D printing allows unique materials to be printed and incorporated into the device and design modification to be made relatively easily for rapid prototyping.

A major challenge for multiplex pathogen detection is determining which type of assay is necessary to identify a true unknown with high confidence. The current device design allows for multiple analysis types to occur, and the multichannel format allows for positive and negative controls to be included on one small device. By incorporating multiple methods for sample identification, such as isothermal polymerase chain reaction (PCR), immunoassays, optical imaging for cellular morphology, antibiotic sensitivity, and adenosine triphosphate (ATP), multiple measurements can be made to improve confidence in sample identification.

Our research for the next fiscal year includes focusing on reducing the number of manual steps for sample identification by exploring integrated optical measurements using a smartphone microscope and isothermal PCR for molecular identification. Overall device design will be streamlined to reduce the need for manual manipulation of the sample.



An unknown biological sample is added to the sample detection reservoirs (squares) containing a custom growth enhancing medium. The sample is transported to the analysis reservoirs (circles) through microchannels. Sample identification and viability assessment occurs after a short growth step using molecular methods (organism identity), optical (presumptive morphological identification), or ATP measurements (viability).

PN17049/2939

Multi-Scale Processes Controlling Spatial Variation in Greenhouse Gas Emissions in a Subarctic Watershed

James C. Stegen

This project combines field measurements with simulation modeling to deepen understanding of and reduce uncertainty in predicted greenhouse gas emissions across permafrost to non-permafrost transition zones. This is vital, as increased greenhouse gas emissions due to permafrost thaw could produce positive feedbacks that accelerate the rate of climate change, yet there is a high-degree of uncertainty in model predictions, because fundamental processes are poorly represented in climate models.

Microbial communities play a central role in the functioning of natural ecosystems by heavily influencing biogeochemical cycles. A major scientific challenge is to use knowledge of those influences to improve the ability of earth system models to robustly predict future climate change by incorporating feedbacks between environmental change and biogeochemical rates. Understanding how shifts in the environment are tied to shifts in biogeochemical rates via changes in microbial communities is particularly relevant in high-latitude terrestrial systems underlain by permafrost due to vast carbon stocks currently stored within thawing permafrost.

The primary research objective of this study is to gain new knowledge of the factors that govern observed patterns in the rates of greenhouse gas (CO_2) emissions associated with permafrost to non-permafrost transition zones. The project further aims to use that knowledge to improve representation of soil microbiology in the Community Land Model (CLM).

To pursue these objectives, field studies are being used to generate data needed to first test CLM predictions and, in turn, provide hypotheses for how the model may be improved through inclusion of additional biotic (e.g., microbial community composition) and abiotic (e.g., organic carbon composition) features. The field site, near Fairbanks, AK, is characterized by spatial transitions in the presence/absence of permafrost and in active layer depth (ALD; the depth of soil overlaying permafrost). Across these spatial transitions, greenhouse gas emission rates are being directly measured and biotic/abiotic features of the associated soil environment are being characterized.

From an initial sampling trip to the Alaska field site in the fall of 2013, CO_2 emissions were characterized across spatial gradients in ALD. Results show a strong relationship between ALD and CO_2 emission rates but only across the lowest elevation sites with the thinnest ALDs. This result points to a highly non-linear influence of ALD on CO_2 emissions at the landscape scale. Multi-scale field sampling designs were implemented in the 2014, 2015, and 2016 field seasons to better characterize this non-linear relationship, understand how other environmental features contribute to the non-linear behavior, and to evaluate the degree to which primary drivers of CO_2 emissions are maintained across spatial scales.

To go beyond the influence of ALD, soil cores were taken across the spatial gradients in active layer depth. Soil core material from 2013 has been characterized in terms of numerous biotic and abiotic characteristics, including microbial community composition and organic carbon composition. These data resulted in a peer-reviewed publication that aims to overturn existing approaches for characterizing microbial diversity and linking the resulting measurements to ecosystem properties.

Soil cores from 2014 sampling have been processed using similar characterization methods. Coupling belowground data to CO_2 flux from 2014 samples revealed that the relationship between ALD and CO_2 flux is indirect. The best predictor of flux rates—and presumably the variable that most directly indicates the control point over CO_2 flux—was a proxy for the degree to which soil C has been microbially processed. CO_2 flux was also coupled with aboveground variables, and tree-stand basal area showed the strongest relationship with CO_2 flux, though this relationship was scale-dependent. Those analyses led to a second peer-reviewed publication.

In addition, CLM has been set up and simulations have been run for a variety of conditions informed by field observations. For example, we observed that a key plant functional trait (specific leaf area) varies significantly with ALD.

In future work, CLM will be run across multiple spatial scales across the field system. This approach takes CLM—a global-scale model often run on 10km^2 grid cells—down to the scale of point field measurements, thereby reaching across an extreme disparity in scale. To support this integration we quantified spatiotemporal variation in CO_2 emissions—using the 2015 and 2016 field seasons to sample at a larger spatial scale relative to 2014—and net primary production, vertical profiles of numerous key variables such as soil moisture, pH, texture, C content/composition, and microbial biomass, composition, and potential enzyme activities have been characterized. These data will enable a multi-scale comparison between CLM predictions and empirical observations.

PN14027/2611

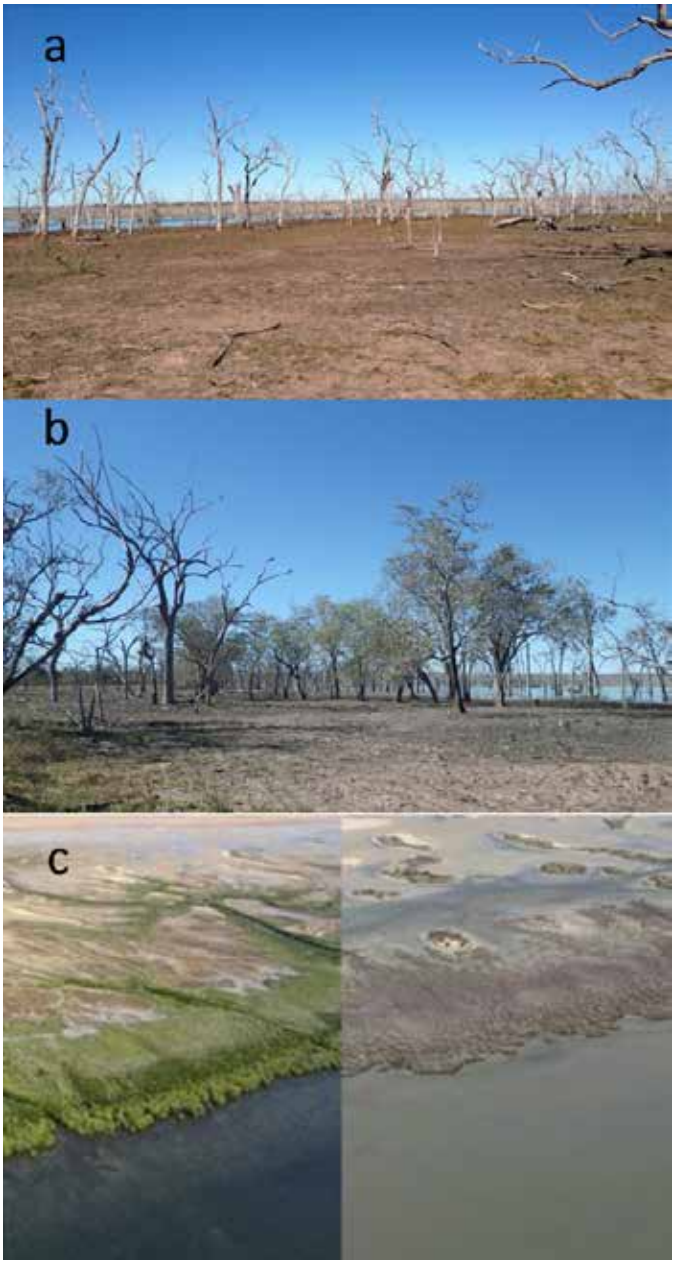
Patterns and Mechanisms of Coastal Forest Loss in Relation to Sea-Level Variability, Drought, and Heat-Waves

Nate G. McDowell

This project allows, for the first time, understanding of the physiological underpinnings of drought induced mortality in coastal species located at terrestrial-aquatic interfaces. This informs on the risk of mortality these ecosystems face under a changing environment and links directly to understanding the impacts of changing vegetation at these ecosystems upon ecosystem resources.

Sea-level extremes, precipitation reductions, and heat-waves are all increasing throughout much of the globe. Extreme impacts, including widespread mortality, have been documented when these driving variables coincide. These driving variables are anticipated to increase in frequency over future decades, thus widespread impacts on coastal ecosystems are likely. The mechanisms of such impacts are unknown, thus realistic predictions of future impacts are impossible. Our primary objective is to test, for the first time, if sea-level shifts, drought, and climate warming are causing impacts on growth and survival of coastal ecosystems.

This experiment was initiated in July 2017 and has focused on obtaining instrumentation for the success of this project in FY 2018. We have successfully obtained the majority of equipment required for water use measurements, tree ring analyses, and soil sampling. We also devoted significant time to identifying field sites for FY 2018 visits, which we have now narrowed down to our key choices.



a) > 1,000 km of severe mortality of Australian mangroves after extreme variation in sea level combined with heat and drought. b) Local survival, potentially due to ground-water access. c) Before/after photos of mangrove loss.

PN17105/2995

Permafrost Microbiome Responses to Hydrologic Perturbation and Subsequent Alteration to Ecosystem Function

James C. Stegen

This project uses laboratory incubations and molecular characterization of field-collected permafrost to deepen understanding of how permafrost microbiomes respond to hydrologic perturbations and connects these responses with shifts in greenhouse gas emissions. This is vital as increased greenhouse gas emissions due to permafrost thaw could produce positive feedbacks that accelerate the rate of climate change, yet there is a high degree of uncertainty in model predictions because fundamental processes are poorly understood.

Microbial communities play a central role in the functioning of natural ecosystems by heavily influencing biogeochemical cycles. A major scientific challenge is to use knowledge of those influences to improve the predictive ability of microbially explicit models by incorporating feedbacks between environmental change, microbiome responses, and biogeochemical rates. Understanding how shifts in the environment are tied to shifts in biogeochemical rates via changes in microbial communities is particularly relevant in permafrost systems due to vast carbon stocks currently stored within thawing permafrost.

The primary research objective of this study is to reveal processes governing the response of permafrost microbiomes to hydrologic perturbation. The project further aims to provide that knowledge and associated data to modeling efforts. Those modeling efforts aim to build a predictive modeling framework that explicitly represents key members of soil microbiomes and the redox-associated processes they drive and feedback with.

Field-collected permafrost is being characterized using advanced molecular tools (multi-omics) and incubated in the laboratory under different redox and hydrologic conditions. In addition, permafrost is being thawed *in situ* across a naturally occurring hydrologic gradient. Laboratory experiments are specifically designed to link permafrost microbiomes and associated biogeochemical processes to shifts in soil moisture conditions.

We conducted an initial scoping experiment that provided key insights into microbiome response-to-thaw and other information (e.g., variability across replicates and the timescale of microbiome responses). This was used to design follow-on experiments.

The initial experiment did not attempt to manipulate moisture, but instead thawed permafrost under either aerobic or anaerobic redox conditions. The reason was to mimic redox conditions that might arise under different moisture conditions; drier soil generally leads to aerobic conditions, while saturated soil generally leads to anaerobic conditions.

Permafrost cores were incubated across a time series that allowed biogenic greenhouse gases to be repeatedly sampled. Replicate cores were also destructively harvested at multiple time points. The results showed that redox conditions had a strong influence over microbiome response-to-thaw and biogenic gas emissions. In particular, under aerobic conditions, the microbiome became dominated by one species: *Rhodoferax ferrireducens*. This organism can grow across a broad range of conditions and likely contributes significantly to system-scale biogeochemical function. One important aspect is that it can couple fermentation products to the reduction of Fe(III). This suggests that it has the potential to inhibit production of CH_4—a process that also relies on fermentation products. In turn, conditions that favor *R. ferrireducens* likely shift biogenic gas production toward CO_2 and away from CH_4. This provides new insight into factors governing the balance between CO_2 and CH_4 following permafrost thaw, which is a critical uncertainty due to CH_4 being a far more potent greenhouse gas than CO_2.

In addition to the microbiome itself, we observed significant shifts in metabolite profiles following permafrost thaw. The data revealed a significant increase in protein-like metabolites associated with phosphorylation molecular transformations under conditions where *R. ferrireducens* became dominant. This suggests that *R. ferrireducens* is likely a dominant driver of biogeochemical cycling and microbe-microbe interactions in thawing permafrost, at least under conditions that facilitate its growth. We have hypothesized that conditions in which there is significant spatial or temporal variation in redox potential facilitate *R. ferrireducens* due to its ability to grow both aerobically and anaerobically.

We tested this hypothesis with an additional experiment, which revealed that redox conditions are indeed important, but that pre-thaw geochemistry is potentially more influential over *R. ferrireducens* response-to-thaw—this points to a strong influence of environmental history.

PN16030/2807

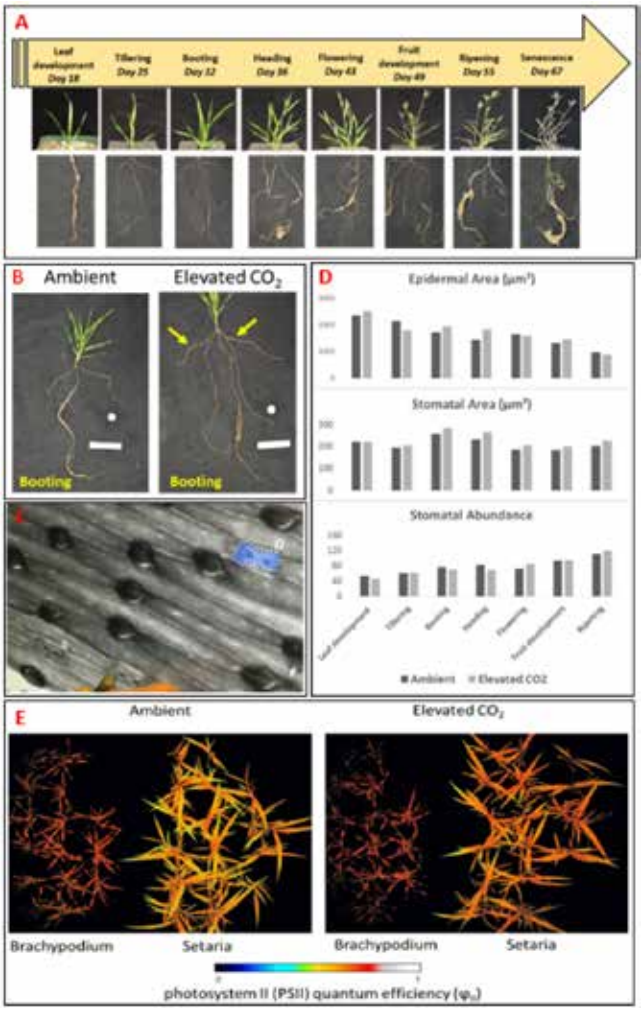
PhenoAccess: Physiological Phenotyping of *Brachypodium* Accessions

Amir H. Ahkami

This research project aims to obtain a better understanding of available phenotypic diversity of model grasses for bioenergy crops in a fluctuating environment.

The main goal of this work is phenotypic analysis of the genetic variability in model grasses by systematic quantification of phenotypic traits. PhenoAccess aims to physiological characterization of natural accessions, inbred lines, and selected mutants of the model grasses *Brachypodium* and *Setaria* under different climate scenarios by extracting an extensive list of phenotypic traits. In collaborating with other projects, the plant phenotypes associated with a cellular structure, plant gene, or a biochemical marker will be identified in a fluctuating environment (e.g., drought, heat wave, elevated CO_2 levels), which eventually allow us to take steps towards bridging the genotype-environment-phenotype gap.

In this study, accessions of *Brachypodium* distachyon Bd21 (C3 model grass) and *Setaria viridis* A10.1 (C4 model grass) were grown under current and elevated CO_2 levels in growth chambers. Detailed growth-stage-based phenotypic analysis revealed different above and below-ground morphological and physiological responses of C3 and C4 grasses to enhanced CO_2 levels condition. According to our results and by screening values of total biomass, water use efficiency, root to shoot ratio, Root System Architecture parameters, and net assimilation rates, we postulated a three-phase physiological mechanism (i.e. RootPlus, BiomassPlus and YieldPlus phases) for grass growth under elevated CO_2 condition. In collaboration with Michigan State University and to characterize additional physiological changes, we used novel, non-invasive, image-based Dynamic Environmental Photosynthesis Imager (DEPI) chambers capable of revealing new transient or environment-specific phenotypes. The analysis of generated images and revealed photosynthesis-specific parameters, including photosystem II (PSII) quantum efficiency (φII), light-driven linear electron flow, and dissipative non-photochemical quenching of absorbed light energy, is in progress. Moreover, these comprehensive sets of morphological and process-based observations are currently in use to develop, test, and calibrate biophysical whole-plant models and, in particular, to simulate leaf-level photosynthesis at various developmental stages of *Brachypodium* growth using the model BioCro (in collaboration with modeling team). Also, to further link the observed phenotypic traits in the organismal level to tissue and molecular levels, the whole-plant phenotypic observations were complemented with the stomatal density, stomatal area, epidermal cell area, as well as gene expression profiles. Using an RNA-Seq. approach, on average, 38 million passing-filter reads with a length of 2x50 bp were generated per sample; 93.6% reads have a quality score higher than Q30, which is far better than Illumina's specification of 85%. Analysis of up/down regulated genes is in progress.



A) Principal growth stages in *Brachypodium distachyon*. **B)** Leaf Node Root (LNR) formation in early stages of *B. distachyon* under ambient and elevated CO_2 levels. High CO_2 condition expedites LNR initiation in the tillering and booting stages. Yellow arrows indicate LNRs. **C)** Stomatal abundance and area measurements using confocal microscopy. **D)** Growth-stage specific stomatal and epidermal abundance/area in *B. distachyon* under ambient and elevated CO_2 levels. E) Fluorescence imaging of *Brachypodium distachyon* and *Setaria viridis* 18 days after planting in DEPI chambers.

PN16107/2884

Probing Complex Microbiomes Using Mass Spectrometry and Sequencing Capabilities to Understand How Microbiomes are Influenced by their Environment

Stephen J. Callister

We are developing and applying a suite of omics measurement and bioinformatics capabilities to better understand how microbiomes present in the soil are influenced by their environment. Soil microbiomes are important contributors to the biotic cycling of nutrients required for all living things.

The goal of this project is to develop next-generation sequencing bioinformatics tools and mass spectrometry capabilities to address key knowledge gaps in our understanding of how complex microbiomes function within the soil environment. Soil microbiomes are responsible for key biogeochemical processes such as carbon and nutrient cycling. Predictive models concerning these nutrient cycles can be improved through better functional characterization of microbiomes at a molecular level. This molecular investigation can only be achieved through the development of sophisticated workflows for measurement of the metaproteome.

Unfortunately, empirical measurement of the soil metaproteome by current analytical workflows results in a low depth of coverage (i.e., a relatively small number of proteins are measured compared to what is predicted by metagenomic sequencing). To address this challenge, we are evaluating and improving metaproteomics workflows combining both demonstrated and emerging bioinformatics and empirical capabilities. These workflows are being systematically evaluated, first on simplified constructed microbiomes and then on complex natural microbiomes such as the soil microbiome found in the Konza Prairie Biological Station (KPBS); a Long-Term Ecological Research site.

For FY 2017, we tackled a fundamental question regarding how different liquid chromatography approaches coupled to mass spectrometry instrumentation (LC-MS/MS) impact the measurement of soil metaproteome depth of coverage. Conventional workflows use a one-dimensional (1D) liquid chromatographic separation to reduce the complexity of peptides from proteins extracted from the soil microbiome (soil metaproteome). The proteins originally extracted must be enzymatically digested into peptides for analysis by the mass spectrometer.

The utilization of a less conventional two-dimensional chromatographic separation approach (2D LC-MS/MS) can also improve depth of coverage by employing orthogonal separations based upon different chemical and physical properties of the peptides. Here, a 1D LC-MS/MS approach and two different 2D LC-MS/MS approaches were evaluated in their ability to increase soil metaproteome depth of coverage. The first 2D approach evaluated utilizes both charge (1D) and physical interactions of the peptides with the solid separation matrix (2D) to separate the complex mixture of peptides directly to the MS instrument ("online" approach).

The second 2D approach performs the first dimension (1D) of fractionation "offline" the MS instrument. While each 2D LC-MS/MS approach is commonly utilized for proteomics research in general, few demonstrations exist for metaproteomics conducted on the soil microbiome. All approaches were applied to: 1) a constructed metaproteome, consisting of a mixture of peptides generated from 47 different microorganisms, and 2) to the soil metaproteome extracted from the Kansas native prairie soil microbiome. Application of each separation approach to both sample types was performed to evaluate whether findings were consistent for samples of differing proteome complexity and to provide insight into other factors that may contribute to a low metaproteome depth of coverage.

From this evaluation, we observed that the "offline" approach resulted in greater metaproteome depth of coverage; however, there are advantages to using the "online" approach when a soil metaproteomics study is limited by peptide mass.

Surprisingly, all approaches resulted in a larger number of protein identifications from the constructed metaproteome than from the soil metaproteome, even though the soil metaproteome had approximately 8 times the number of predicted proteins compared to the constructed metaproteome. This observation and others pointed to the finding that, while implementation of either 2D LC-MS/MS approach will greatly improve soil metaproteomics depth of coverage, additional optimization of these approaches along with greater mass spectrometry dynamic range sensitivity will be required to further increase soil metaproteome depth of coverage.

PN16031/2808

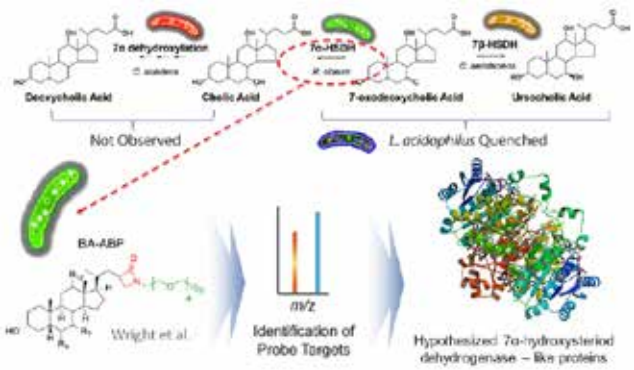
Probiotics and Secondary Bile Acids as Regulators of the Gut Microbe Interactome

Hans C. Bernstein

We have designed a unique benchtop microbiome model that enables us to learn how microbes in our gut collaborate in a way that affects human health. We are using this tool to discover the mechanisms by which microbes modify human-derived bile acids to act as signaling molecules. Ultimately, this understanding will help us model more complex environmental microbiomes, in addition to potentially having useful impacts on human health.

The gut microbiome has been conceptualized as a "virtual organ" that has co-evolved with the host to provide a number of functional feedbacks onto the host's health, including the production of metabolites that regulate bacteria. Bile components represent a highly abundant pool of host-derived extracellular metabolites and are major regulators of the gastrointestinal-tract microbiome. Primary bile acids (e.g., cholic acid) are excreted into the small intestine (duodenum) to promote solubilization and absorption of nutrients; however, approximately 400–800 mg of bile components escape enterohepatic circulation daily and become substrate for microbial biotransformation in the large bowel. We have developed and implemented a new capability that integrates a host-independent human gut microbial model with metabolomics and transcriptomics capabilities to elucidate multi-species chemical synthesis pathways. The capability will first be established by building upon a putative multi-species bile acid transformation process, identified via preliminary experiments. This capability has been applied as a tool to test the hypothesis that secondary bile acids (e.g., ursocholic acid) and related intermediates regulate microbial community interactions in the human gut microbiome.

Model "bench-top-microbiomes" were built from microbes isolated from the human gut. The isolates were chosen based on three criteria: 1) their ability to be cultured in isolation and consortia, 2) the availability of a sequenced/assembled genome, and 3) *a priori* knowledge or predictions on bile-acid-responsive enzymes. Each organism in the model is able to perform one-of-three important metabolic reactions involving human bile acids: 1) 7α-dehydroxylation, 2) oxidation (bile acid hydroxysteroid dehydrogenase [HSDS]), or 3) hydrolysis (bile salt hydrolase [BSH]).



We used this model microbiome to discover new multi-species collaborations that result in the transformation of human-derived cholic acid into the secondary bile acid product, ursocholic acid (UCA). Furthermore, we have experimental evidence showing that the probiotic *Lactobacillus acidophilus* disrupts the accumulation of UCA. UCA is present in high abundance within the human gut, but little is known about its effects on human health or microbial physiology. Preliminary data suggest that there is a strong negative correlation between the successful outcomes in gastral bypass surgery and the abundance of UCA.

The next phase of development will be applied to a more complex, host-independent microbial gut community and tested and mechanistically validated using activity-based probes that will aid the discovery of new bile-acid-responsive enzymes that transform or support the transformation of bile acids to secondary products.

PN17028/2918

Process-Based Understanding of Perturbation Impacts in Tidally Influenced Nearshore Terrestrial-Aquatic Interfaces

James C. Stegen

This project will reveal processes governing the response of tidally influenced nearshore systems to salinity-based perturbations, with a focus on salt-water-intrusion-induced increases in salinity. This knowledge will be obtained, in part, by developing a novel belowground smart sampling system (BGSSS) that can be used in the lab or field to collect minimally invasive samples for high-resolution characterization from locations targeted based on non-destructive spatially resolved biogeochemical and rhizosphere monitoring. Knowledge enabled by the BGSSS will be used to design ecosystem management strategies that enable the consistent provisioning of ecosystem services to human society in the face of perturbation.

Nearshore tidally influenced terrestrial-aquatic interfaces (TAIs) have high ecological/economic value and are vulnerable to perturbation. These TAIs have a relatively small footprint and play a disproportionally large role in local-to-global biogeochemical cycles, yet are not adequately accounted for in earth system models. These systems also provide ecosystem services related to carbon storage, energy production, infrastructure protection, water quality, food production, and ecological habitats. We currently lack the process-based understanding needed to predict how ongoing sea level rise will govern the future state of TAIs and their ability to provision services to human society. Enhanced predictive ability will improve decision-making related to mitigation and adaptation strategies.

Nearshore TAIs are experiencing a broad range of perturbations, but sea level rise is globally distributed and may, therefore, have some of the largest impacts. A focus on impacts of sea level rise is particularly pressing given the increasing rate of sea level rise. The impacts from sea level rise have global, as well as local, implications. Shifting biogeochemical and vegetation dynamics in nearshore TAIs may alter global biogeochemical cycles, and these changes may also impact local features such as the ability of TAIs to provide clean water, fish spawning habitat, and stabilize the physical system to minimize impacts of storms (e.g., erosion, infrastructure integrity).

In the future, we will perform model-informed laboratory experiments to reveal processes governing the impacts of sea level rise on hydrobiogeochemical function and plant physiology. To enable these experiments, we have made significant progress during the short existence of this project. We designed and built a new mesocosm capability. This new capability has enormous potential to inform mechanistic models being developed.

PN17106/2996

Response of *Brachypodium* Associated Microbiomes to Drought and/or Elevated Carbon Dioxide Illuminated by Biogeochemical and Molecular Measurements

Ljiljana Paša Tolić

We are applying a suite of omics and biogeochemical measurement capabilities to better understanding how microbiomes associated with* Brachypodium *influence its response to drought and elevated CO_2.* Brachypodium *serves as a model grass for bioenergy crops.

The goal of this project is to combine biogeochemical and molecular profiling approaches to compare the rhizosphere and phyllosphere microbiomes associated with drought tolerant and drought resistant *Brachypodium* distachyon variants grown under current and elevated CO_2 conditions.

The rhizosphere microbiome is composed of microorganisms that interact with the below ground portion of the plant, while the phyllosphere consists of microorganisms inhabiting the above ground portion of the plant. It is generally understood that the microbiomes habiting these regions are dynamically responding to and modulating the physiology of the plant host as it responds to the environment. Because of these interactions, we are testing the following hypotheses: 1) rhizosphere, phyllosphere, and surrounding soil microbiomes associated with distinct genetic *Brachypodium* variants will differ structurally and functionally in their response to drought and/or elevated CO_2, and 2) the exchange of root exudates between plant-microbe-soil are highly sensitive to changes in atmospheric pCO_2. Results from these hypotheses will contribute to development of a Virtual Plant-Atmosphere-Soil System (vPASS) model, which enables a predictive understanding of how plant microbiomes are shaped as a function of plant genotype and its ecosystem environment.

Brachypodium is an ideal model grass for addressing the question of how a plant's associated microbiome helps it to ameliorate the negative effects of environmental perturbation. Characteristics that make it an ideal grass model include 1) a small diploid genome, 2) a completed genome sequence, 3) a large collection of difference accessions (> 50) suitable for experimentation, and 4) a relatively short growth cycle (about 12 to 16 weeks to senescence). Its above ground biomass, short stature, and environmental tolerance make are important traits for bioenergy associated crops.

To complete the goal of this project, advanced multi-omics capabilities are being applied to the rhizosphere and phyllosphere of different *Brachypodium* accessions, ranging from high to low tolerance to drought. The capabilities being applied are generating metagenomics sequence, bacterial and fungal amplicon sequence, metaproteomics, metatranscriptomics, and dissolved organic matter data streams. These data are being used to develop a baseline understanding of microbial taxon and their phenotypes within the rhizosphere and phyllosphere. In addition, advanced isotope ratio mass spectrometry (IRMS) is being used to track carbon uptake into the plant, its transport, and its stimulatory effect on microbial populations through photosynthate exchange in the rhizosphere and phyllosphere.

Two of three *Brachypodium* accessions, available after successful bulking, were selected for a water limitation experiment. Only two were selected at this time because of limited space within growth chambers. Accession Bd1-1 is relatively drought tolerant, while Bd21 is a relatively less-drought-tolerant accession. Having completed the cold forcing stage for Bd1-1 (8 weeks), both germinated Bd21 and Bd1-1 were placed within a small growth chamber in preparation for the perturbation, which began 6 to 8 weeks after germination. The experimental design consisted of twenty pots utilizing a Prescott soil for planted and unplanted controls. Prior to harvesting, both watered and water-deprived plants were subjected to an atmosphere containing $^{13}CO_2$, allowing for the production of labeled photosynthate derived metabolites to be tracked from leaf to rhizosphere using IRMS.

As expected, plant phenotypes from the two accessions exhibited differences in tolerance to water-deprived conditions. At harvest, different fractions of soil representing rhizosphere and soil outside of the plant's influence were collected along with the phyllosphere. Extractions for ribonucleic acid, deoxyribonucleic acid, and dissolved organic matter have been completed.

For FY 2018, multi-omics data generation will be completed. Our aim is also to extend our research to additional *Brachypodium* accessions grown in the laboratory and in garden plot experiments, as well as to keystone species identified from sensitive terrestrial aquatic ecosystems.

PN17010/2900

Retro-Fitting Non-Traditional Microbes with State-of-the-Art Synthetic Biology Tools: Towards the Next Generation of Engineered Microbial Biosensors

Hans C. Bernstein

We are re-engineering microorganisms to act as autonomous, microscopic sensors using state-of-the-art synthetic biology tools. Ultimately, this will lead to new biotechnologies that help us monitor and protect our nation's critical water and agricultural infrastructure from contamination and disease.

New technologies made possible by modern synthetic biology are strengthening our nation's ability to continue addressing priority focus areas in energy and security. Rapid breakthroughs resulting from new developments in genome editing and gene synthesis now make it possible to build and program functions into living cells. This renaissance in molecular biology and genomic sciences has enabled us to construct genetic parts with circuit-like connectivity that have the ability to process logical operations, sense chemical threats, and store encrypted memory. However, almost all successful examples of these capabilities to date have been demonstrated using well-characterized "laboratory" organisms.

While useful for the development and demonstration of capabilities under stable conditions, these species do not survive well in most "real-world" chemical sensing applications. These traditional organisms are also limited in their metabolic activity, preferring substrates, such as sugars, that are typically not available in settings relevant to chemical threat detection such as surface water or soil.

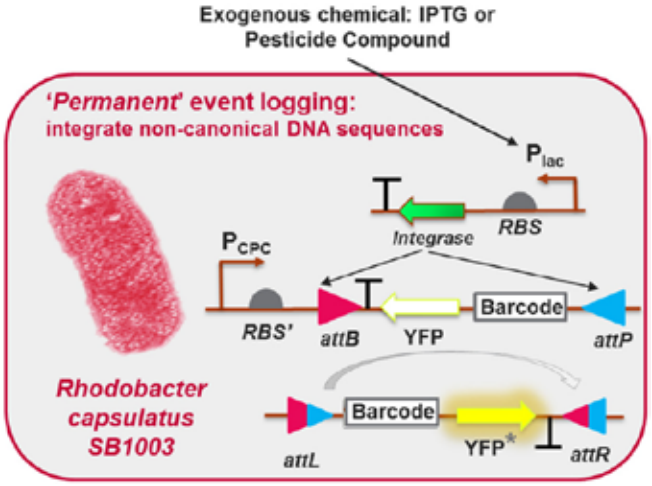
A key challenge that applied researchers currently face is to be able to manipulate the best suited organisms for the particular environment or purpose. To effectively harness new developments in programmable genetic circuits for national security applications, such as environmental sensing of chemical warfare agents, we must expand our collection of programmable host chassis beyond the limited set of species that constrain the current state-of-the-art.

We have advanced the state of this technology by designing, building, and testing genetically regulated event detectors in non-traditional microbial host chassis, which provides an expanded framework for programing genetic circuits into non-traditional chassis species that are already adapted to operationally relevant environments.

Our team has developed integrase-based event detectors that are capable of sensing exogenous chemicals in the environment and storing permanent, digital records of exposure in custom deoxyribonucleic acid sequences that are non-canonical to the host's genome.

The proof-of-concept has been prototyped with traditional inducible promoters, while new regulatory systems are also being engineered to respond to environmental contaminants such as pesticides. We have chosen and prioritized new host chassis that are capable of persisting autonomously in a complex, dynamic environment due to native biological properties that are not currently represented by the microbes traditionally employed by synthetic biologists.

Our new devices are specifically designed to operate within metabolically versatile soil microbes, *Rhodobacter capsulatus* and *Cupriavidus necator*. These hosts can grow and persist in a wide range of environments, which include dynamic soil and agricultural habitats, making them ideal sensors for a future capability to protect some of our nation's most critical infrastructure.



The next phase of this research will be focused on applying these living, event loggers to sense and respond to real-world chemical and biological threats. This research is providing an important roadmap for how to apply the rapidly emerging synthetic biology capabilities for building living, autonomous chemical sensors that are robust and operationally relevant to complex environments.

PN17001/2891

Signatures of Warfighter Response to Pathogen-, Toxin-, and Activity-Induced Stress

Justin G. Teeguarden

Exposure to pathogens, extreme physical stress, and chemicals reduce the performance and health of civilians and the armed forces. We are developing the next generation of computational and experimental technologies for detecting exposures so they can be quickly identified, eliminated, and effects treated.

Individual and combined exposures to toxins, pathogens, and physical stresses affect human health. Current technologies allow detection of exposure to some pathogens and toxins, but the signatures of exposure to most remain hidden or difficult to find. Conventional technologies detect up to a few thousand chemicals and far fewer signs of physiological stress. Global proteomic signatures of the adverse effects have low sensitivity and specificity. We sought to overcome these limitations by creating the capability to identify the presence of thousands of unknown toxins using advanced computational chemistry enabled by high-performance computing. We will also develop a complimentary targeted proteomic method for detecting the response of humans to toxins, pathogens, and other stresses that will increase sensitivity up to 100 times higher than current methods.

Pioneering new capabilities to more rapidly detect harmful exposures and mitigate and treat affected populations to improve health and performance is our major objective. We will accomplish this through major advances in computational chemistry and proteomics, applied to a subset of our population often uniquely exposed and uniquely positioned to rapidly deploy countermeasures or mitigate exposures. The research, however, will benefit the general population.

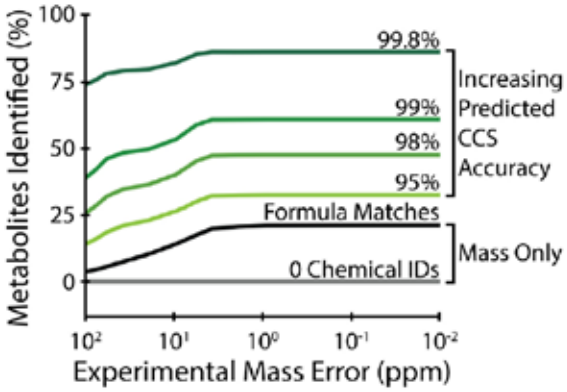
Many naturally occurring chemical toxins cannot be identified in the environment or in humans because the "signature" of the toxin on analytical instruments cannot be determined in the absence of running a chemical standard that reveals that unique signature.

We assembled a team of analytical chemists, instrument engineers, chemical engineers, quantum chemical experts, and computer scientists and developed an approach to identify members of a group of common marine toxins without chemical standards. "Signatures" are calculated from initial 3D geometries by first predicting arrays of ionized chemical structures using density functional theory, and then chemical ion structures are constructed using quantum molecular dynamics on our NWChem platform using the Cascade Supercomputer. The final "signature" for each toxin takes the form of a collisional cross section (CCS).

The initial calculations were initiated and will be followed with more accurate approaches in FY 2018. Generally, we showed that the method has an accuracy exceeding 95% for most chemicals. We also showed how accuracy related to chemical identification.

Plausible complementary peptide signatures of fatigue, toxin, and pathogen exposure amenable to high sensitivity targeted proteomics were identified using human proteome databases. Protein entries in the PeptideAtlas database were cross referenced with published findings from exposure studies to select a subset of proteins present in human biofluids associated with thee exposures.

In the next FY, our main objectives will be to complete the design of these protein signatures and use them to measure the level and nature of stress in humans or laboratory test systems. In parallel, we will use the computationally derived "signatures" of several marine toxins to identify the compounds in unknown mixtures of toxins. These complementary methods will next be used to identify both unknown signatures of exposure and signatures of human response to toxins and other forms of stress, enabling more rapid, directed interventions and improved health.



The molecular formula of 25% of chemicals in the Human Metabolome Database can be can be established based on mass only (solid dark black line), but no compound identifications can be made (light black line) without standards (mass only). Adding calculated CCS as the orthogonal feature, between 30% and 85% of chemicals can be uniquely identified depending on the accuracy of the calculated CCS. Extended to a much larger chemical space (DSSTox), even the lowest accuracy CCS would lead to the ability to identify > 150,000 compounds without the need for analytical standards.

PN17103/2993

Spatially Resolved Quantitative Gene Expression Analyses Applied to Transitioning Mouse Gut and Soil Microbiomes

Galya Orr

Quantitative gene expression analyses of complex microbial communities have been performed exclusively by population measurements, which lump the behavior of an entire population into a single average value, ignoring the impact of the diverse local environments on the responses of individual organisms and the complex spatial relationships between them. To address this gap, we will establish quantitative gene expression analysis approaches in individual and small clusters of bacterial cells, while preserving their spatial context within their native environments.

The function of individual organisms within a microbial community is governed by the local environment and interactions with neighboring organisms. In complex spatial organizations of communities that consist of hundreds or thousands of different species, individual cells of a given clonal population of organisms can experience vastly different local environments and are likely to express very different genes in response. The spatial relationships between individual organisms within the community are, therefore, key to understanding the role of the individual organism and ultimately to deciphering how the community functions as a whole.

Meta-transcriptomic analyses have gained important insights into processes that take place by the community as a whole and suggested functional roles for different species. However, this averaged cell population approach lumps the behavior of an entire population into a single average value, while ignoring the impact of the diverse local environments on the responses of individual organisms and the complex spatial relationships between them. Adding a spatial context to gene expression analyses will be vital to understanding fundamental rules by which complex communities function and evolve under changing environments.

Recent advances in single-cell analysis have enabled quantitative gene expression analyses in individual cells, while preserving their spatial context within the tissue. To date, these approaches have been demonstrated only in eukaryotic cells.

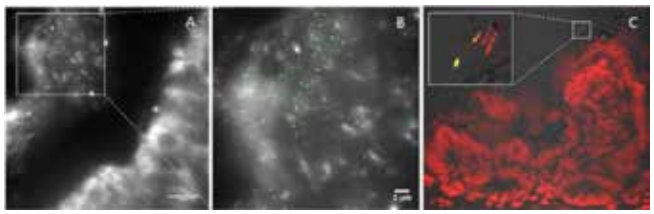
The translation of these techniques to bacterial cells presents multiple challenges due to their small size and content, as well as their complex environments within microbiomes. Building on our experience in single-cell ribonucleic acid sequencing (RNA-Seq) in mammalian cells, we will apply methods and reagents developed specifically for small amounts of starting RNA to small clusters of bacterial cells, isolated by laser capture microdissection from mouse gut cryo-sections. We will also apply quantitative fluorescence *in situ* hybridization (FISH) to enable multi-gene expression analysis in single cells, while preserving the spatial context of the individual cells. When applied using single molecule fluorescence imaging techniques, FISH enables the quantification of single transcripts in intact cells with high accuracy. Further, by combinatorial barcoding of the FISH probes, it will be possible to significantly increase the number of genes that can be quantified within a cell simultaneously.

Our project is designed to establish, for the first time, two challenging techniques for quantitative multi-gene expression analysis in individual or small, spatially defined clusters of bacterial cells, while preserving the context of their local environments and spatial relationships within the microbiome. Our collaborators at Lawrence Berkeley National Laboratory have found that spleen tyrosine kinase (SYK) plays an important role in controlling the gut microbial composition and behavior. To quantify gene expression in wild-type and SYK-/- mice, we have established quantitative FISH targeting SYK mRNA copies in mouse gut cryo-sections, while also establishing the use of antibodies to identify epithelial cells and dendritic cells, two cell types where SYK is expected to be expressed.

PN16110/2887

In parallel, we have established the use of FISH to identify the presence of bacterial cells and their species in these cryo-sections. By combining the two approaches, it will be possible to quantify and spatially correlate the expression of SYK with the presence of specific bacteria associated with the gut epithelial cells. We also have established the isolation of small, spatially defined microbial clusters from the gut cryo-sections, followed by RNA extraction and library preparation for RNA-Seq.

Ultimately, this effort will identify fundamental rules and mechanisms by which microbiomes operate with details that cannot be achieved by averaged cell population measurements that ignore the local environments and spatial organization among the community members and the host.



Identify bacteria in mouse gut cryo-sections using super resolution FISH. A. Stochastic optical reconstruction microscopy image (green) of a section treated with the universal bacterial FISH probe (EUB338), which recognizes most bacterial 16S rRNAs, overlaid on a wide-field image. B. A larger view of the area marked by the white square in A. C. Confocal fluorescence microscopy image of a section treated with FISH probes targeting 16S rRNA unique to *Lactobacillus reuteri* (green/yellow). Nuclei and bacterial DNA are stained using DAPI (red).

Statistical Integration of Omics Data from Microbiomes

Joseph M. Brown

We will develop an infrastructure and computational tools designed to process, analyze, and visualize microbial community sequence data to address the challenge of computational analysis.

Microbes exist and function in communities. Studies of natural environmental or human-associated sites often reveal hundreds to thousands of microbial species co-existing (i.e., a microbiome). The interactions between microbes within such a microbiome are often complex, as is the impact the microbiome as a whole has on its environment. Deciphering microbial activities in their natural environments is a prerequisite to understanding their functional role in environmental processes. Research efforts to understand the vital roles these communities of microbes play in human health and environmental sustainability have, thus, turned to high-throughput technologies to measure the various biomolecules extracted from a sample (i.e., nucleic acids, protein, and metabolites) to infer the functional capability and emergent properties of the microbiome.

One of the most readily available data types for microbiome research is nucleic acid sequence data. It is feasible to sequence the DNA extracted from natural microbial communities consisting of many species. However, such DNA samples consist of a composite mixture of the genetic material of all of the community members (i.e., a metagenome), and inferring the functional capability or assigning taxonomy from complex mixtures of genetic material has proven challenging. Further, sequencing technologies are capable of generating approximately a terabyte of data in a week, requiring the need to develop novel analytical tools. Robust analysis of complex data from microbial communities in human hosts and in the environment will be key to understanding their vulnerability to change.

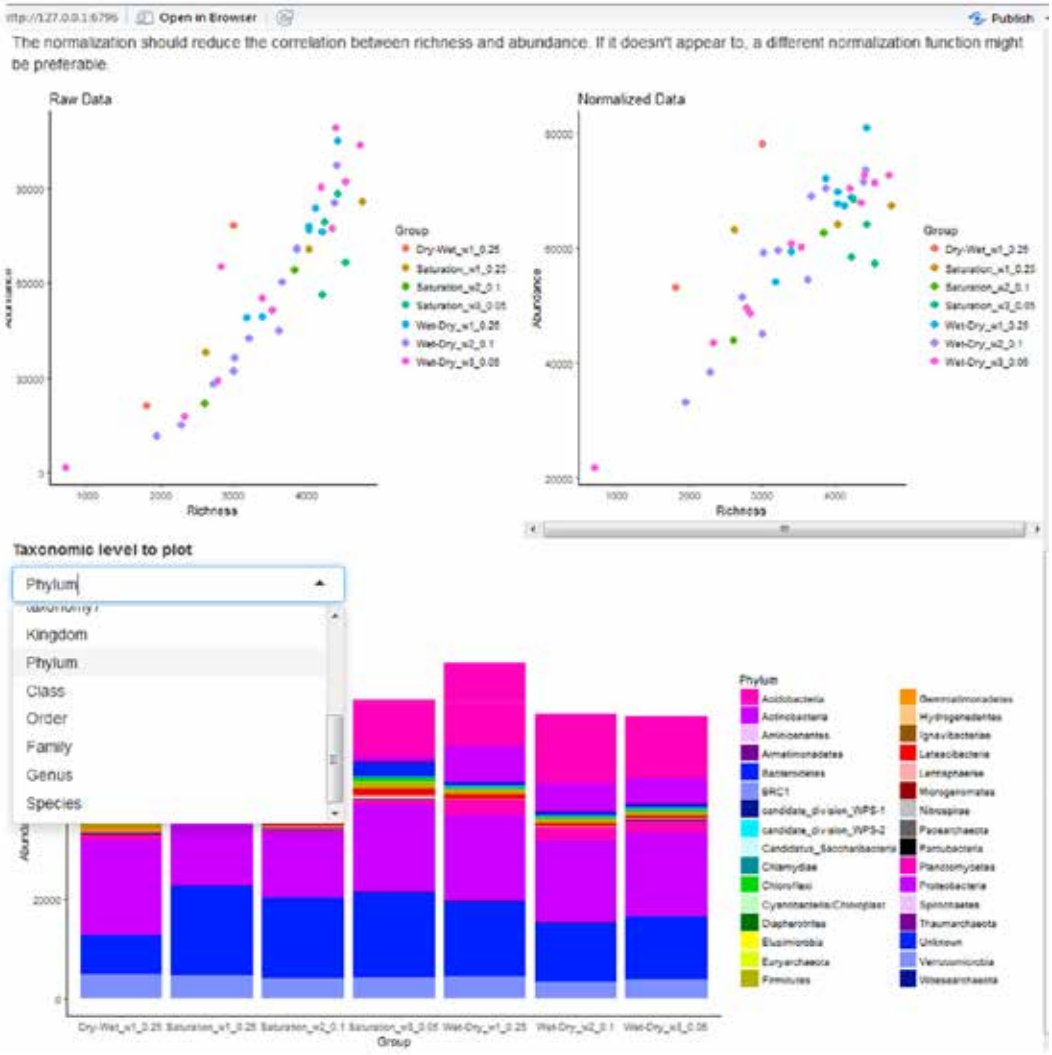
We have built a sequencing capability and data delivery process at PNNL around industry best practices. Data obtained from our instrument go through rigorous quality scrutiny, facilitated by an interactive and extensible dashboard that we have developed and released to the open source community. Our automated protocol allows us to utilize institutional computing for archiving, processing, and data delivery to accelerate the pace of data distribution to researchers.

Sequencing fragments of an rRNA marker gene as a means of capturing community structure and dynamics is standard practice in microbial ecology. We developed a protocol to complement our sequencing capability that enables rapid quantification and classification of the marker gene amplicon data. An average run of the full protocol takes less than 30 minutes and includes common output file formats embedded into a processing report to be delivered to collaborators.

To further advance this capability, we have developed an interactive analysis dashboard to pre-process, quality check, and perform community composition abundance testing across sample groups. The dashboard is compatible with industry standard input files, making it the next logical step for analysis. We have integrated best practices that were developed by PNNL researchers to perform analyses and track provenance of the underlying data. The dashboard also acts as a means of data processing standardization across the initiative and facilitates the sharing of results through its data export functionality.

Future work for this project includes expanding metagenomic annotation capabilities to better align with existing projects and to include more comprehensive metabolic pathway connections. More visualization tools are needed to facilitate the exploration of metagenomic and metatranscriptomic data and drive hypothesis confirmation and discovery. Some of this work has been prototyped using the Trelliscope framework and is planned to be completed in the next fiscal year.

PN15077/2752



Screenshot of Metagenomic Exploration and Analysis Dashboard. Plots of pre- and post-normalization of count data and its relative effect on community composition is seen here and represents on one of many pages of data processing and analysis options available to the user.

Understanding Fundamental Design Principles Underlying How Biological Systems Adapt to Engineered Functions

Scott E. Baker

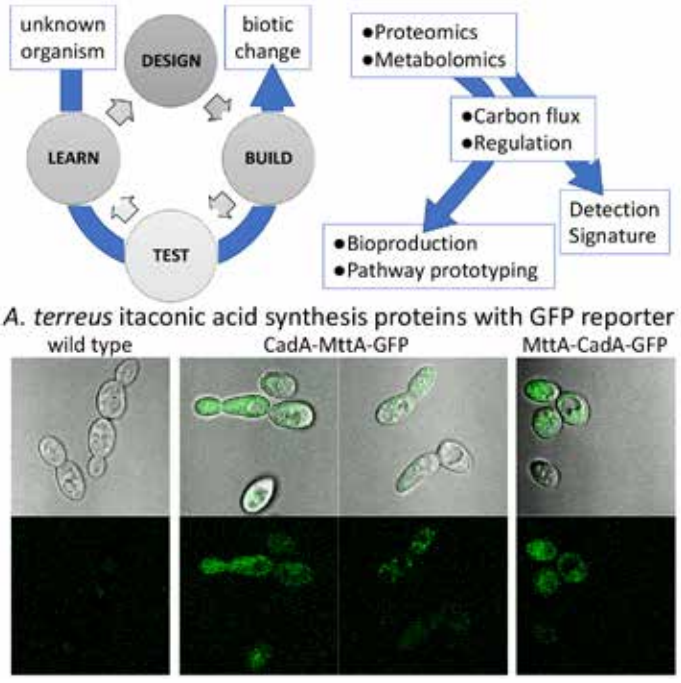
Synthetic biology is a critical element of the U.S. bioeconomy.* Yarrowia lipolytica *is a powerful genetically tractable yeast species, tolerant to many stresses and, therefore, a potential workhorse organism for synthetic biology applications.

We are exploring the use of *Yarrowia lipolytica* for production of lipids for biodiesel production and as a synthetic biology chassis. We are currently working to better understand the central carbon metabolism of this yeast by tracking metabolite and protein levels under different conditions. By learning about the regulation and response of this yeast to growth on a variety of carbon sources, we will be better able to channel carbon flux into small molecule production. Though laborious detection of engineered organisms can be done using genome sequencing, we are using a reversal of the iterative "design-build-test-learn" process to detect signs of genetic engineering. With mass spectrometry proteomics capabilities, we have identified expression of non-host proteins and modified cellular signatures, from protein to product, which are affected by the process of genome engineering.

This year, we used genome modification, microscopy, proteomics, metabolomics, and other molecular biological techniques to make yeast strains that produce the commodity chemical itaconic acid. We have implemented a system for genomic incorporation and balanced protein production (co-translational cleavage with a 2A viral peptide) that can be used for a diverse array of biosynthetic pathways clusters and performed analysis to detect altered expression in order to define signatures of genome engineering.

We have also gained biological insight into co-translational cleavage: results show that intron absence is associated with reporter green fluorescent protein production. Additionally, gene order in the P2A system affected cleavage efficiencies.

During the next fiscal year, we will incorporate novel pathways into *Yarrowia lipolytica*, including multiple terpene biosynthetic pathways from plants and xylose usage genes from an archaean, and a separate fungal species. We expect the system we have generated to be used for prototyping of biosynthetic pathways in a tractable host yeast.



PN17099/2989

Understanding How Saltwater Intrusion Interacts with Plant Communities to Affect Soil Greenhouse Gas Fluxes in Coastal Forests

Ben Bond-Lamberty

The uncertainties surrounding greenhouse gas (GHG) fluxes are particularly high at the terrestrial-aquatic interface (TAI). This project will examine whether forest TAIs soil gas fluxes are suppressed by salt water intrusion, in ways mediated by access to fresh water and vegetation community.

Coastal forests and TAIs are sentinel systems at the boundaries of surface water bodies and terrestrial ecosystems typically defined by classic biome type (e.g., temperate rainforest, prairie, desert, etc.), and we do not have either the required process understanding, data, nor models to correctly understand the effects of perturbations on TAI stability and ability to recover from perturbation. Disturbances and nutrient loading may also increase soil respiration and organic carbon turnover in both soils and marine receiving waters. We will calculate annual mean fluxes, explore possible spatial aggregation of environmental variables, and compare these results with laboratory-oriented manipulative mesocosm experiments and tree-ring analyses along the sampling transects from other projects.

During FY 2017, we have researched sea-level rise and inundation of coastal carbon stocks for use in developing the field work component. We have also been investigating the regional sea-level rise projections within Hector-BRICK to potentially study the impacts of future sea-level rise projections on coastal wetlands.

We explored the effects of land use policies on water availability and water demand and, thus, on global water scarcity, necessary for linking our field-scale results with regional- to global-scale processes. We also collected and analyzed global net primary productive products in preparation for a grassland productivity study on both upland and lowland.

In addition, we conducted new data analyses for a study on U.S. West Coast drought, which included remote sensing (e.g., radar, vegetation indices, and solar-induced fluorescence) and meteorological data (e.g., Palmer Drought Severity Index, North American Land Data Assimilation System).

Most importantly, we completed a site visit to the Smithsonian Environmental Research Center as part of an extended site evaluation effort being conducted.

During FY 2018, we will begin with a West Coast site visit, concentrating on the Olympic Peninsula in western Washington, and including Sequim. We will deal with site permitting and final experimental design issues and install measuring equipment and take first samples for lab analyses.

PN17107/2997

Using Modified Proteins for Forensic Deconvolution of Xenobiotic Dose Quantitation and Timing

Jordan N. Smith

A novel biomonitoring approach utilizing modified-protein fingerprints to reconstruct dose and timing of dose is being developed. To support this approach, we developed a multipronged technique to identify protein modifications caused by environmental chemical exposures and measured global protein turnover rates in rats. This novel biomonitoring approach could allow for dose reconstruction with greater accuracy and precision than current methods.

Humans are exposed to a broad mixture of chemicals from occupational settings, consumer products, accidental releases, and the environment. Many chemicals are known to cause or have been linked to adverse health effects. To better understand the potential risks associated with these exposures, efforts are underway to biomonitor targeted human populations for chemical exposures, as highlighted by the National Research Council of the National Academies report, *Exposure Science in the 21st Century: A Vision and a Strategy*.

Current state-of-the-art methods to predict how much of a chemical or when exposure occurred (i.e., dose reconstruction) require *a priori* knowledge of one of two factors: dose or timing of dose. Blood or urine samples are analyzed for biomarkers of exposure, usually the chemical of interest or a metabolite of that chemical. Computational models are used to reconstruct a dose that may have caused a measured biomarker concentration. Biomarker concentrations are highly dependent on time and dose, and without some information about one of these factors, the most sophisticated dose reconstruction strategies can only conclude the dose occurred. For compounds that are rapidly cleared from the body, even robust *a priori* information may not be sufficient to reconstruct doses without a high level of uncertainty, since biomarkers can only provide adequate information over a few biological half-lives.

We are developing a new approach to reconstruct dose and timing of dose using chemical adducts on proteins. Many highly reactive chemicals (which are typically toxic chemicals and are difficult to biomonitor) will react with proteins in blood to form chemical-protein adducts. Since there are many proteins in blood, exposure to a chemical will result in the formation of chemical adducts on many proteins. Different proteins have different natural turnover rates, which means each chemical exposure will create a modified-protein fingerprint that is unique by dose and time after the dose occurred. The goal of this project is to identify, measure, and utilize modified-protein fingerprints for dose reconstruction.

In our first aim, we developed a multipronged approach to identify protein adducts. First, global proteomic measurements were used to search for protein adducts in purified proteins, plasma, and whole blood treated with reactive chemicals. Custom bioinformatics software was developed to aid in verification of protein adducts from mass spectrometry data. Second, activity-based protein profiling was used to search for protein adducts. This technique utilizes chemical probes designed to enrich specific protein targets, which we used as an inverse indicator of targets for protein adduct formation. Used in tandem, global mass spectrometry and activity-based protein profiling were used to identify and confirm protein adducts caused by organophosphate pesticides, herbicides, fumigants, and other environmental chemicals.

As a second aim, we measured turnover rates of plasma proteins in rats. Rats were fed isotopically labeled lysine to label proteins. Then feed was changed to an unlabeled equivalent feed, and blood was repeatedly sampled from these rats. Global proteomics and activity-based protein profiling were used to measure changes in isotopic ratios of proteins. Using this approach, we were able to identify 273 proteins and measure turnover rates of 174 proteins. This study is the first to measure global plasma protein turnover rates in rats, measure variability of protein turnover rates in any animal model, and utilize activity-based protein profiling for enhancing turnover measurements of targeted, low-abundant proteins. Together with our ability to identify protein adducts (Aim 1), these protein turnover measurements allow us to utilize protein adducts for advanced dose reconstruction.

In the upcoming year, we plan to extend our work in plasma to tissues. This is important, as identification of protein adducts in tissues will allow us to study novel mechanisms of toxicity. Protein adducts will be identified in whole tissues using recently developed global proteomics and activity-based protein profiling. Second, a novel nano-proteomics platform (SNaPP) will be applied to map protein adducts on tissue slices, effectively imaging adducts co-localized with various heterogeneous features of tissues. The ability to identify and map protein adducts in tissues will allow for discovery of novel toxicity targets and provide insight on spatial-specific toxicity and disease progression in tissues.

PN16023/2800

Virtual Plant-Atmosphere-Soil-System (vPASS) 1.0: Quantifying Signatures of Phenomic Expression of a *Brachypodium* Ecosystem as a Function of Genomic and Environmental Variables

Daniel C. Fortin

We are faced globally with the challenge of increasing essential plant ecosystem services; critical to this challenge is understanding how the plant phenotype is shaped as a function of its genotype and environmental interactions. This project will perform statistical analyses of experimental data generated by plant experiments and develop empirical models relating* Brachypodium *ecosystem phenome expression as a function of controlled environmental conditions and variations in the plant genome.

The ability to discover statistically significant correlations between variations in the plant genome and *Brachypodium* ecosystem phenome expression requires a robust exploratory data analysis (EDA) framework that can inform the development of appropriate statistical models. EDA is of paramount importance for the purposes of assessing data quality, determining relationships among explanatory variables, assessing relationships between explanatory and outcome variables, and making a preliminary selection of appropriate models. The multi-omics measurements that will serve as predictive variables for plant phenomic expression require novel data integration and EDA approaches.

This research project will lead to the development of a statistical framework that can perform EDA, data integration, discover biologically meaningful relationships, and explore the appropriateness of candidate statistical models on data to inform downstream plant models. This research will also result in empirical models of *Brachypodium* phenome expression that will identify meaningful correlations between phenome expression and variation in the plant genome.

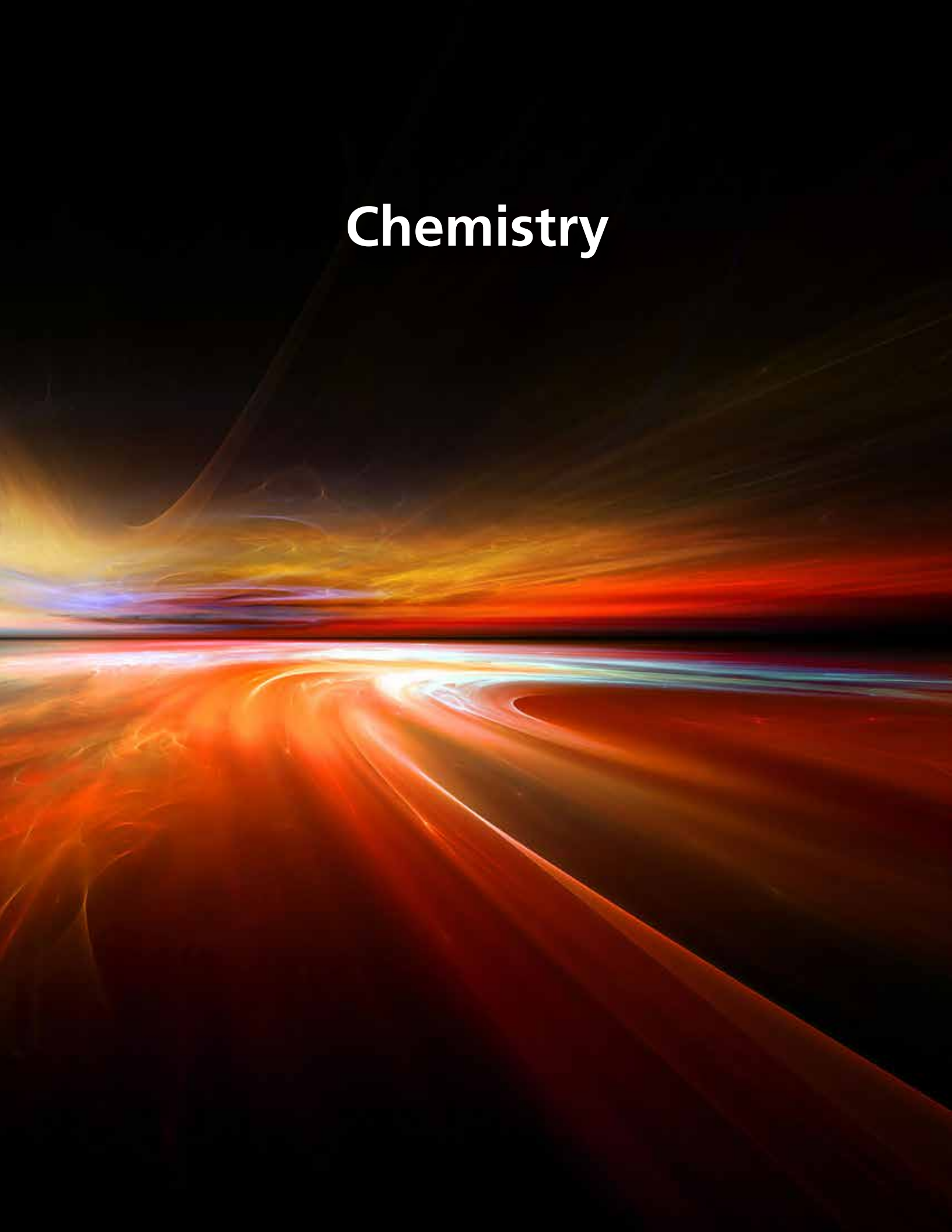
In FY 2017, we finalized the data processing framework (i.e., normalization, quality control, and EDA). This framework leveraged data normalization methods, algorithms for multidimensional outlier detection, and filtering methods for various data (e.g., metabolites) based on objective measures.

Our main effort was developing statistical models to analyze experimental data from initial plant experiments where control and drought conditions were imposed on plants in the lab and measurements were collected on below- and above-ground biomass, as well as leaf and root metabolites. Through this analysis we were able to connect metabolite profiles to biomass accumulation and its response to drought conditions. Our analysis demonstrated the predictive power of molecular information on high-order plant phenotypes, as well as provided some insight into the role of metabolites and metabolic pathways associated with drought response. This multi-source modeling framework is useful, in general, to investigate the covariance between plant genotype, environmental conditions, and higher-order plant phenotype.

We developed a workflow for data integration, visualization, exploration, and analysis for plant phenotypes and metabolites. Empirical predictive models were developed connecting metabolite abundance to drought response in *Brachypodium*. We also performed a root-shoot allometry analysis to investigate how root-shoot ratios correlate with metabolic profiles.

PN16112/2889

Chemistry



Acid-Base Catalysis for Converting Electrocatalytic Hydrogenation Intermediates

Huamin Wang

We are working to understand fundamental aspects of condensed-phase acid-base chemistry for conversion of mixed oxygenates to produce bio-fuel. We will learn and transfer principles inspired by nature, such as enzymes, to construct new catalysts with enhanced activity and robustness.

Biomass liquefaction and produced bio-oil/bio-crude upgrading is a very promising approach to produce renewable fuels; however, it is still facing significant challenges in catalytic upgrading for carbon efficiency and economics. Electrocatalytic hydrogenation (ECH) provides an alternative approach for bio-oil/bio-crude upgrading. However, oxygen removal and carbon backbone rearrangement are still required to produce finished hydrocarbon fuels with maximized carbon efficiency. Acid/base-catalyzed reactions will serve in this role through dehydration of alcohols and C–C bond formation. Significantly higher reaction rates based on the current state-of-the-art is, therefore, required for a streamlined and integrated process with ECH. Such a rate enhancement will need to be achieved for both oxygen elimination and C–C bond formation reactions, with complex ECH intermediates from bio-crude, in the presence of aqueous reaction media, and at low temperatures.

To address these challenges, we propose to abstract and transfer principles/properties inspired by enzymes to construct new solid catalysts with enhanced activity and robustness. We will provide fundamental insight on the impact of water on the acid-base sites and reactions, impact of reactant complexity on the reaction rates, and impact of nano-environment of acid-base sites on their activity. We will develop fundamental understanding of atomic-level structure-activity relationships to suggest active site and its microenvironment requirements and identify catalysts and reaction parameters that are practical for integration with ECH.

In the first year of the project, we have identified zeolite catalysts with much improved stability for dehydration of alcohols with existence of condensed water, investigated the impact of water content on alcohol dehydration rate at conditions relevant to the proposed bio-oil/bio-crude intermediates, and constructed and commissioned a trickle bed reactor for acid-base catalyzed dehydration and C-C formation reactions.

To improve the stability of zeolite catalysts used for the proposed application, the impact of major factors, such as Si/Al and synthesis media, on zeolite stability has been studied and the correlation of surface hydrophilicity with stability was established. Zeolites with improved stability for application in conversion of bio-oil/bio-crude intermediates are being developed.

The impact of water on cyclohexanol dehydration, a representative reaction in bio-oil/bio-crude conversion, has been investigated. Water in the condensed-phase negatively impacts reaction rate via changing reaction mechanism and active site properties, suggesting that reducing water content near active sites could potentially enhance reaction rate for bio-oil/bio-crude conversion. Thus, the surface hydrophilicity/hydrophobicity of zeolite was tuned by using core-shell structure to control water content near active site, and this effort is ongoing.

In FY 2018, we will focus on dehydration reactions for oxygen elimination on Brønsted acidic zeolites under conditions relevant to the overall process. Specifically, we will 1) determine the structure and properties of active sites in zeolite confines under the relevant conditions, 2) establish the effect of molecular size and structure of alcohols on their ability to interact with active sites in zeolite confines and the mutual interaction of reactants, and 3) further understand factors that impact stability of zeolite and identify approaches to enhance zeolite lifetime.

PN17021/2911

Chemical Bonding in Uranium Oxides Studied by Uranium-233 and Uranium-235 Nuclear Quadrupole Resonance Spectroscopy

Herman Cho

The aim of this project is to develop a novel method to study the unique and distinctive way the actinide elements, which include uranium, neptunium, and plutonium, combine with other atoms. The discovery of an explanation for this behavior would be a major contribution to the effort to elucidate the complex chemistry and physical properties of these elements, with significant implications for national security, nuclear energy, and environmental remediation.

Soon after the synthesis of the first transuranic elements, Seaborg proposed that the actinide elements were analogs of the lanthanide series and recommended the placement of the actinides below the lanthanide row in the Periodic Table. In the 70 years of actinides research since then, however, a more complex picture has emerged, with the elements below plutonium (Z = 94) being often found to behave like transition metals rather than lanthanides. The validity of the lanthanide analogy continues to be a ubiquitous theme in actinide science, with far-reaching relevance. The purpose of this project is to construct a first-of-its-kind pulsed nuclear quadrupole resonance (NQR) spectrometer designed to probe actinide elements for insights on the similarities and differences of these elements with lanthanides and transition elements. The feasibility of actinide NQR spectroscopy will be evaluated with experiments on uranyl salts. The goal is to achieve the first ever detection of NQR signals from an actinide isotope, which would open a new avenue for investigating electronic structure in actinide complexes. Uranium isotopes are the focus of this project, but almost every actinide element up to curium has a long-lived isotope that is theoretically detectable by NQR methods.

NQR spectrometer design and construction. Instrumentation has been assembled and tested for frequency-swept time domain NQR experiments in the requisite spectral range (0.1–3.0 GHz). The main components include a pulsed console and a broadband cryogenic probe with a 5-mm resonator. Custom sample capsules were fabricated for containment of radioactive material.

Computational analysis of candidate systems. With theoretical electric field gradient calculations provided by Prof. Jochen Autschbach's group at the University of Buffalo, a range of different uranyl compounds were surveyed as candidates for an initial demonstration of this approach. Their results made possible calculations of NQR transition energies that guided the choice of $Cs_2UO_2Cl_4$ as the most promising first candidate due to its favorable NQR transition energies and the potential for measuring spectra of neighboring ^{35,37}Cl- and ^{133}Cs-occupied sites.

Sample preparation. Large crystals of $Cs_2UO_2Cl_4$ in quantities sufficient for NQR experiments have been synthesized, and the product has been verified by x-ray crystallography. Samples have been made with both depleted and 95% ^{235}U-enriched uranium.

FY2018. With samples in hand and a functioning spectrometer, experiments have been initiated to seek the ^{235}U NQR signals of $Cs_2UO_2Cl_4$, which is predicted to consist of three lines near 461.2, 922.5, and 1383.7 MHz. Concurrent with the NQR experiments, ^{35,37}Cl and ^{133}Cs nuclear magnetic resonance spectra will be acquired to determine electric field gradients of the Cl and Cs sites, which will enable us to develop a comprehensive picture of the electric fields surrounding the actinide site. Cryogenic accessories will also be procured to allow measurements down to 4.2 K to be made. Lower temperatures will enhance the sensitivity of the measurements, as well as suppress lifetime (T1) broadening of the ^{235}U NQR lines.

PN17029/2919

Select Sorbent Physical Characteristics

Isotope	Nuclear ground state spin *I*	Quadrupole moment ($\times 10^{24}$ cm^2)	Half-life (ψ)
^{227}Ac	3/2	1.72	21.8
^{229}Th	5/2	4.3(9)	7340
^{231}Pa	3/2	−1.72(5)	3.30×10^{4}
^{233}U	5/2	3.663(8)	1.60×10^{5}
^{235}U	7/2	4.55 - 6.0	7.00×10^{8}
^{237}Np	5/2	3.886(6)	2.10×10^{6}
^{239}Pu	1/2	0	2.41×10^{4}
^{241}Am	5/2	4.34(5)	432
^{243}Am	5/2	4.34(5)	7370
^{247}Cm	9/2	unknown	1.56×10^{7}

Left: Uranium-235 enriched $Cs_2UO_2Cl_4$ crystals; Middle: cryogenic broadband NQR probe, with a meter stick for scale; Right: resonator section of the NQR probe, revealing the 5-mm sample tube in the solenoidal coil detector.

Development of Integrated Framework for High-Accuracy Excited-State Simulations of Dynamical Processes

Karol Kowalski

This project will enable highly accurate excited-state simulations capable of utilizing existing peta- and soon-to-be exa-scale computer architectures. New excited-state capabilities will significantly advance our abilities to describe processes related to energy conversion, photo-catalysis, light harvesting systems, and various spectroscopies. Additionally, new functionalities will be used to interpret data produced by DOE light sources.

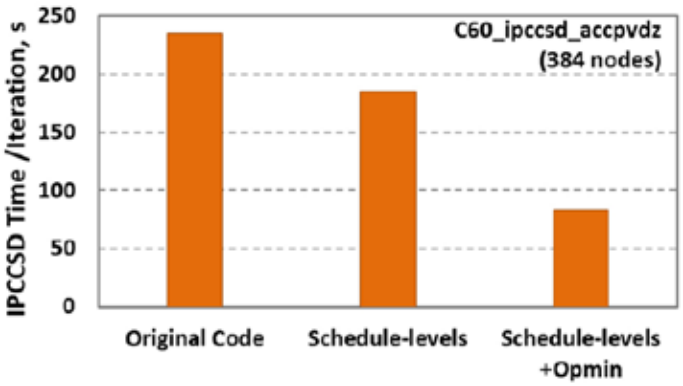
One of the prominent frontiers in computational chemistry is the study of excited-state-driven chemical transformations. The main challenge to be addressed is the inclusion of the appropriate level of accuracy to describe complicated structure of excited-state potential energy surfaces and interactions between them, including avoided crossings, barriers, and conical intersections. This can only be achieved be employing excited-state wave-function-based approaches that can properly capture a vast array of excited-state correlation effects. Although these methodologies are known to be accurate, the main challenge is posed by their steep numerical overhead. To address this issue, we plan to integrate the state-of-the-art, theoretical Equation-of-Motion Coupled-Cluster (EOMCC) formulations with emerging programing models that will enable efficient utilization of high-end peta- and near-exa-scale computational platforms.

We are hoping to achieve a significant progress in the development of parallel EOMCC formulations to enable them for systems composed of hundreds of atoms. Proposed development will also significantly extend the area of application of the EOMCC formalism. We are planning to develop an integrated environment to tackle excitations in various energy regimes for closed- and open-shell systems. We also are hoping that the proposed development will pave the way for the future extension of the EOMCC formalism to linear response coupled cluster (LR-CC) and closely related CC Green's function formulations.

The above-mentioned goals will be achieved through the integration of the theoretical developments with state-of-the-art programming models. In particular, the availability of specialized tensor libraries plays a critical role in this project. For this reason, the effort associated with providing certain class of capabilities to Tensor Algebra for Many-Body Methods (TAMM) is an inextricable part of this project. TAMM is a framework designed to support quick development of performance portable CC methods. This effort will include: efficient implementations for tensor data structures, element-wise tensor operations, and optimized operations of blocks of tensors.

In FY 2017, we 1) developed IP/EA-EOMCC TAMM approaches, improved scaling and performance of electron-affinity/ionization potential EOMCC methods, performed the largest up-to-date simulations of EA/IP with EOMCC methods, began a large CC-scale calculation with 1,658 basis set functions and full utilization of the Cascade supercomputer system in (T) calculations (3.4 PF, in progress), and had one paper published, with another paper in preparation.

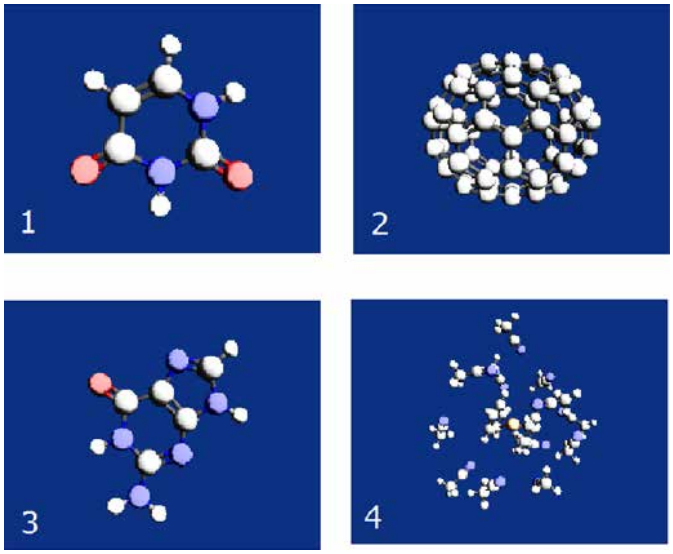
In FY 2018, we will further extend the EOMCC capabilities by focusing on the following topics: 1) enabling TAMM implementation of the EOMCCSD approach for calculating vertical excitation energies; 2) developing a new class of parallel eigensolvers for EOMCC theories to tackle excited states in arbitrary energy regime, including valence excited states and core-level excited states; 3) further enhancing the performance of TAMM-generated EOMCC implementations by integrating



Speed-up of the TAMM-generated IP-EOMCCSD code compared to the existing (NWChem) implementation.

PN17019/2909

EOMCC implementations with highly efficient parallel integral library (Libint) and utilizing novel runtimes to enhance the parallel performance of the EOMCC implementations.



Very accurate IP-EOMCCSD calculations for ferrocene in acetonitrile solutions are the largest IP-EOMCCSD simulations performed up-to-date (approximately 1,300 orbitals).
1) Graphical representation of the uracil molecule.
2) Graphical representation of the C70 fullerene.
3) Graphical representation of the guanine molecule.
4) Schematic representation of the ferrocene molecule ($Fe(C_5H_5)_2$) surrounded by twelve acetonitrile molecules (CH_3CN).

Electrocatalytic Reduction of Phenols and Ethers

Donald M. Camaioni

In this project, we examined the pathways and challenges of catalyzing low-temperature electrochemical processes for electrical energy storage in the form of hydrocarbons suitable for transportation fuels derived from the lignin components of biomass.

Biomass is the oldest renewable energy supply; however, many biomass and bio-oil feedstocks are not used for modern transportation due to high levels of oxygenation. A wealth of catalysis research exists to thermally convert aqueous bio-oils into hydrodeoxygenated liquid hydrocarbons through reaction with hydrogen. However, for locations where high temperatures and high-pressure hydrogen are not available, the best route to convert biomass to hydrocarbons is unclear.

Previous work has shown that model biomass compounds, such as phenol, can be electrochemically converted to products like cyclohexanone, cyclohexanol, and cyclohexane using metal electrocatalysts. Our project was initiated to provide understanding of the mechanism, rate-limiting steps, and energies required for electrocatalytic reductions in comparison to thermal hydrogenation.

We gained insight into the potential of electrocatalytic processes to produce transportation fuels using renewable feedstocks at a wide scale of plant sizes, even in decentralized locations. It is anticipated that this work will contribute to energy security by developing processes for producing liquid hydrocarbon transportation fuels on demand, in a renewable manner, without greenhouse gas emissions, and also absent the risks associated with using a dwindling fossil fuel source with extremely volatile availability.

Our investigation has shown that the mechanism for the electrocatalytic hydrogenation (ECH) of phenol and aryl ethers on platinum group metals is the same as that for thermal hydrogenation (TH), where the adsorbed phenol is reduced via sequential addition of adsorbed hydrogen atoms. The difference between ECH and TH is the origin of the hydrogen reduction equivalents (H_2 or water) and the hydrogen coverage on the catalyst surface, which is controlled by H_2 pressure in the thermal process and by the electrochemical potential in the electrocatalytic process. We determined that Rh/C and Pt/C are more active for electrocatalytic reduction of phenol than Pd/C, with Rh/C being the most active at two to three times higher activity than Pt/C. Methyl or methoxy substituents led to lower rates compared to the unsubstituted compounds due to repulsion with the metal surface. In addition to hydrogenation of the phenyl ring, ethers underwent C-O bond cleavage reactions via hydrogenolysis and hydrolysis.

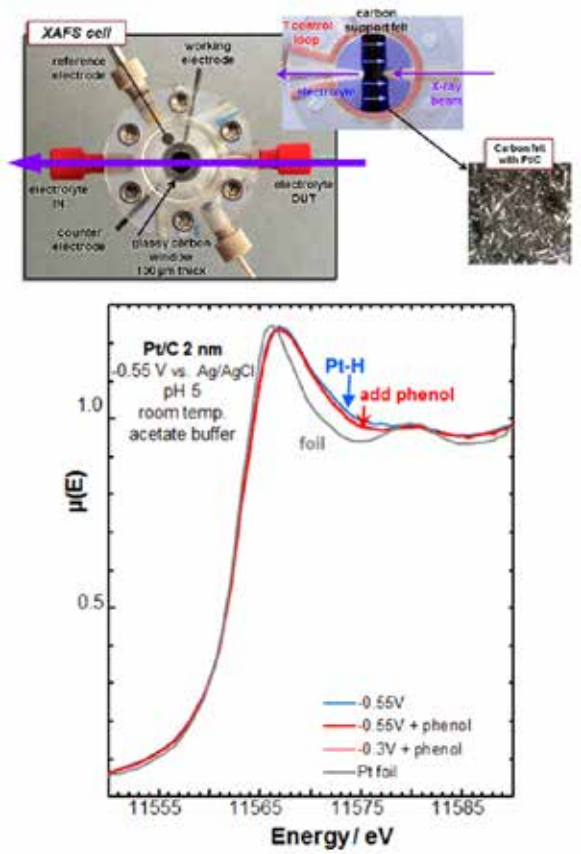
We explored the effects of increasing temperature and pressure on phenol conversion and determined that, at ambient pressure, there was deactivation of the catalyst at 60° to 80°C for both TH and ECH. We conducted TH measurements at 20 bar H_2, where the deactivation was not observed, even at 100°C. A pressure of just 6 bar H_2 at 80°C was sufficient to prevent catalyst deactivation during TH. We hypothesize that deactivation is due to insufficient hydrogen at the catalyst surface, which allows dehydrogenation of adsorbed organic species (e.g., phenol and/or partially reduced intermediates), which fouls the metal surface. The observation that operating at higher H_2 pressure during TH prevented deactivation of the catalysts suggests that operating electrocatalytic reductions under several bar H_2 may well be a remedy for the deactivation of electrocatalysts.

Scanning Tunneling Electron Microscopy (STEM) was employed to investigate the nanoscale morphologies of the catalysts. In the case of carbon-supported Pt catalysts, atomic resolution High-Angle Annular Dark Field (HAADF) STEM images showed faceted crystalline Pt nanoparticles coexist with both amorphous and atomically dispersed Pt species on the carbon. Approximately 5 to 20% of the total Pt existed as small clusters of 2–4 Pt atoms and isolated single Pt atom sites.

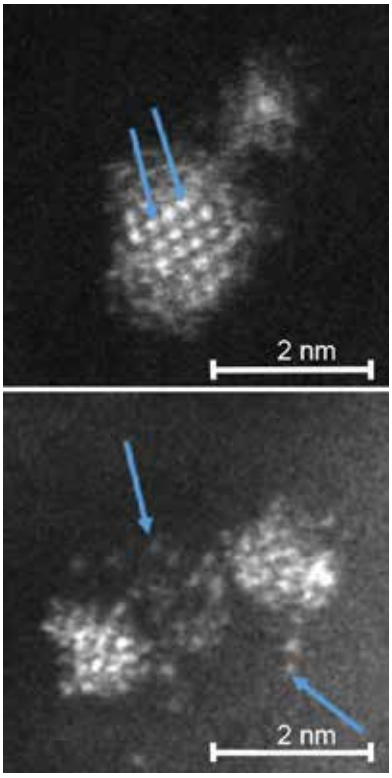
Additionally, we developed an electrochemical cell for *in situ* X-ray absorption spectroscopy measurements to probe the surface state of Pt-on-carbon electrocatalysts. Using this cell, we monitored the change in hydrogen adsorption on Pt with applied potential. At more reducing potentials, more hydrogen was adsorbed onto Pt, while in the presence of phenol, the hydrogen adsorption was decreased. We also detected a Pt-C interaction at a distance that matches what is expected for phenol adsorption on platinum, indicating X-ray absorption spectroscopy can be used to measure coverages of hydrogen and phenol during catalysis.

PN15099/2774

In summary, this project has advanced Laboratory capabilities to understand and control thermal and electrocatalytic hydrogen addition reactions, developing and applying techniques based on microscopy, *in situ* spectroscopy, and detailed kinetic measurements.



(Top) Photograph and schematic details of the electrochemical cell developed for *in situ* X-ray Absorption Fine Structure spectroscopy of electrocatalytic hydrogen addition to phenol. (Bottom) X-Ray Absorption Near-Edge spectroscopy spectra (Pt L3 edge) showing the formation of Pt-H surface species during electrochemical charging of a Pt/C catalyst and the subsequent decrease when phenol is added to the electrolyte solution. The Pt foil spectrum is shown for reference.



HAADF STEM images of Pt/C catalysts showing crystalline Pt nanoparticles coexist with atomically dispersed Pt species. Arrows highlight single Pt atoms both in the nanocrystal and on the carbon support surface.

Ethanol Conversion to Fuels and Co-Products via Ketones with Five or More Carbons in the Chain (C_5+ Ketones)

Karthi Ramasamy

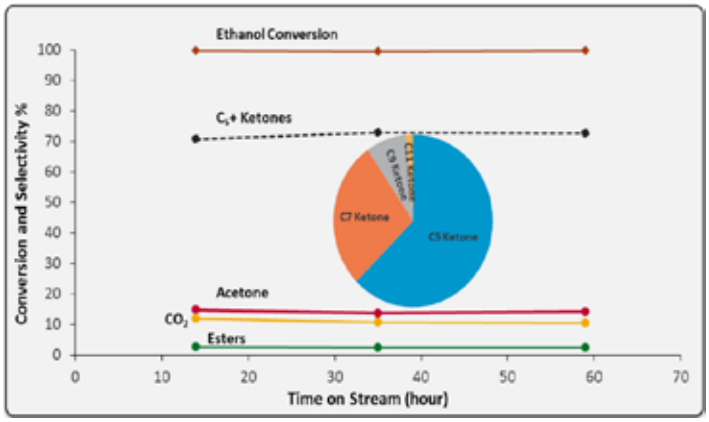
Ethanol is commercially produced today with domestic production near 15 billion gallons. The ethanol markets are limited by a "blend wall" (e.g., production is 10% of current gasoline use). Recent work conducted at PNNL has demonstrated a novel pathway to selectively convert ethanol to high value and market chemical intermediates (C_5+ ketones) with minimal loss of carbon over a low-cost, multi-functional, mixed oxide catalyst.

The C_5+ ketones have several uses or can be further converted to higher-value compounds. For example: 1) the measured research octane number of the ketone mixture is 99.4, and this is the only thermochemical-derived candidate in the co-optima program as a fuel additive component list; 2) the C_5+ ketones can be easily hydrogenated to corresponding alcohols, which are similar to plasticizer alcohols, such as Shell's Linevol 79 process, containing C7-C9 alcohol mixture; and 3) The C_5+ can be converted to detergent alcohols and lubricants by aldol condensation-hydrogenation and this higher alcohols can be simply dehydrated-hydrogenated to make jet and diesel fuel. The conversion of ethanol to C_5+ ketones goes through a sequence of several steps.

First, ethanol dehydrogenates to form acetaldehyde, then two molecules of acetaldehyde condense to form hydroxybutanal, which then isomerizes (intermolecular hydride shift) to form hydroxybutanone, followed by retro-aldol reaction to form acetone and formaldehyde. Once acetone is formed, it continuously condenses with acetaldehyde to generate C_5+ ketones such as pentanone and heptanone. The cascading nature of the chemistry in converting the ethanol to higher ketones requires a multi-functional catalytic system. The catalysts of interest are those containing both acidic and basic properties. For example, mixed oxide materials, such as MgO-ZrO_2, CaO-P, MgO-Al_2O_3, MgO-SiO_2, and ZnO-ZrO_2, along with promoter materials, such as CuO, PdO, and NiO, are envisioned as the catalysts of interest. The specific focus of this project was to 1) improve product yield, 2) get better control over selectivity between C5 and C9 ketones, and 3) further information about catalyst activity, stability, and lifetime.

Initial experiments were conducted on copper (Cu) promoted mixed oxide catalyst for the conversion of ethanol to higher ketones. The conversion of ethanol was 100%, and the selectivity to the higher ketones was in the approximately 75% range for the first 4 hours of time-on-stream. Beyond the 4 hours of time-on-stream, the ethanol conversion maintained at the 100% level, but the selectivity to higher alcohol dropped at a faster rate and reached below 50% levels at around 20 hours of time-on-stream. At later stages of the experiment, the main products generated were the esters formed via the chemical reaction between aldehydes (Tishchenko Reaction) that are generated from the ethanol dehydrogenation. Cu is highly dispersed in the freshly synthesized catalyst and, at the ethanol conversion reaction conditions (high temperatures), Cu sinters together and forms bigger clusters, which then promotes the ester formation over the desirable C_5+ ketones.

To improve the lifetime and the stability of this catalyst, palladium- (Pd-) promoted mixed oxide catalyst was synthesized. Over this Pd-promoted catalyst, ethanol conversion was approximately 100 percent, and the selectivity to C_5+ ketones was approximately 75 percent. With the Pd catalyst, the undesired by-product esters formation was maintained below 3%. The selectivity to CO_2 is at theoretical levels (each mole of ketone generates a mole CO_2 and the longer ketones carbon chain lowers the CO_2 formation). Unlike the Cu catalyst, the Pd-promoted catalyst is demonstrated to be very stable for the ethanol conversion and the selective to C_5+ ketones formation for well beyond the 50 hours of time-on-stream.



Conversion of ethanol and selectivity to the C_5+ ketones and the other by products over a Pd-promoted mixed oxide catalyst.

PN17048/2938

Fundamentals of Electrocatalytic Hydrogen Addition

Oliver Y. Gutierrez-Tinoco

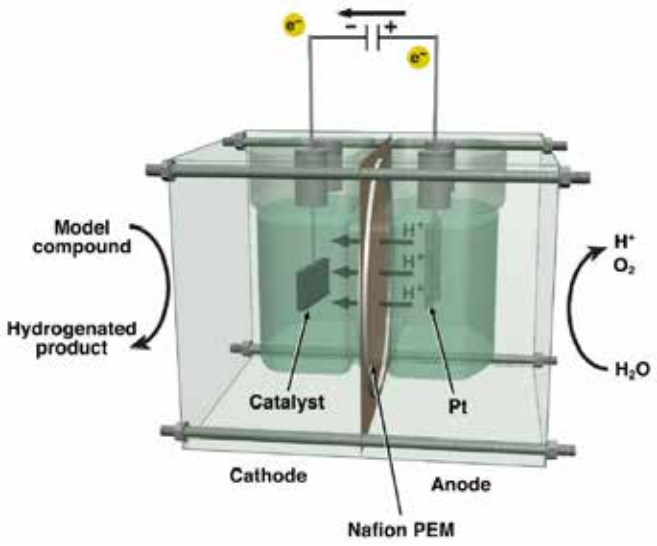
This project aims to overcome the challenges of producing transportation fuels from bio-oil at much milder conditions than those used by current technologies. Electricity is being used as the driving force for that transformation.

Bio-oils are the first products from the conversion of bio-mass into usable fuels and chemicals. The high content of oxygen and reactive compounds in bio-oils hinders their use in combustion engines. Thus, processes are needed for stabilizing bio-oils and for producing drop-in fuels. The last decades have seen a wealth of research aiming to adapt technologies typically used in refining of fossil oils to the conversion of bio-oils. Myriad catalysts have been developed that react H_2 and the compounds in bio-oil to remove oxygen as water at very high pressures and temperatures.

Electrocatalysis, in contrast, offers the unique opportunity of stabilizing reactive compounds and even removing oxygen subjecting bio-oil to electrical cathodic potentials at near ambient temperatures and pressures. Previous work has shown that the compounds present in bio-oil can be converted applying such approach. Exploratory research focused on the conversion of bio-oil compounds, such as phenols and ethers, has shown complex and unexpected dependences of the conversion rates on the experimental conditions. In other words, the chemistry that operates during electrochemical reactions of oxygen-containing compounds has been poorly charted. This lack of understanding prompted our group to initiate this project and to perform fundamental research targeting the description of catalytic phenomena at the molecular level.

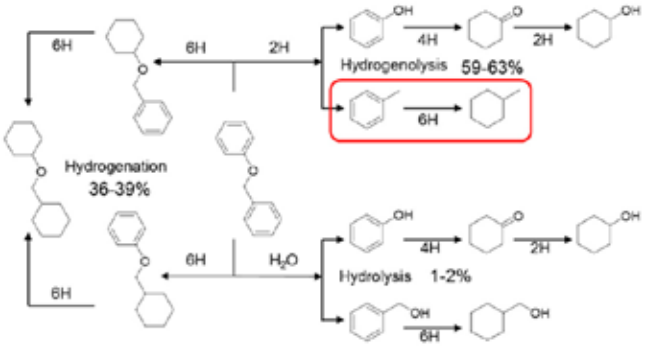
The main result of this project will be the understanding of chemical principles that dictate desired transformations. In turn, this understanding will bridge from the concept of combining water electrolysis with bio-oil conversion to novel technological development contributing to energy security.

During FY 2017, we developed capabilities needed for accomplishing the objectives and milestones of the project. That included the design and construction of electrochemical cells that allow performing experiments under well-controlled conditions of temperature and half-cell potentials, while minimizing the negative effects of resistances to the flow of ions between electrodes.



Schematic showing the concept for converting bio-oil compounds via electrochemistry, as well as the actual configuration of the experimental reactor.

One of the main scientific finding of the work during FY 2017 was the description of the reaction mechanisms for hydrogenation of increasingly complex model molecules ranging from phenol to substituted aryl ethers over noble metals at room temperature and atmospheric pressure. At such mild conditions, the aromatic rings of phenolic compounds and carbonyl groups in ketones and aldehydes can be fully hydrogenated to aliphatic alcohols. Moreover, carbon-oxygen bonds can be cleaved in ethers, effectively producing hydrocarbons from some of the explored compounds.



Reaction network and catalytic pathways for the conversion of benzyl phenyl ether under electrocatalytic hydrogenation. The red square highlights the hydrocarbons produced in the reaction.

PN17022/2912

Another main finding was that the conversion rates depend on the concentration of hydrogen adsorbed on the metal surface. This concentration can be manipulated by varying hydrogen pressure or cathodic potential. More important, catalyst deactivation observed during the conversion of particular compounds can be counteracted by increasing the magnitude of those two parameters. An important aspect of this work is that these experiments probe the intrinsic activity of the metals and the effective potential at their surface. Thus, parameters such as conductivity of the media and catalyst preparation do not affect the interpretation of the results.

As our investigations also cover the synthesis and characterization of catalytic materials, we have designed and constructed electrochemical cells that allow the spectroscopic characterization of the catalysts by X-ray absorption under reaction conditions. With these measurements, we have confirmed the correlation between hydrogen adsorbed at the metal surface and conversion rates. An outstanding finding of the spectroscopic studies is that the nature of the reactant being converted not only influences the concentration of species at the surface but the bulk composition of the catalyst itself.

The implication of our findings is that reaction rates and selectivities can be predicted and tuned, within the thermodynamic and kinetic boundaries, by rational catalyst design and choice of the reaction conditions.

Next year, we plan to expand our investigation to other functionalities and compounds relevant for bio-oil conversion, as well as to other pH ranges.

Preliminary results indicate that the rates of hydrogen addition are promoted and even C-O bonds in alcohols are cleaved under electrochemical conditions at acidic conditions. We will continue the spectroscopic characterization of catalytic materials under reaction conditions to quantitatively link the reaction environment with changes in structure and performance of metal catalysts and in concentration of reactants adsorbed to the metal catalysts. This research is expected to reveal unexplored catalytic routes for the conversion of bio-oil and to enable novel strategies to prepare highly active, selective, and stable catalysts.

Fundamental Understanding of Nucleation Processes to Assess Solution Stability and Phase Growth and Genesis

Gregg J. Lumetta

We are developing a physicochemical framework for predicting and manipulating precipitation reactions relevant to nuclear materials processing and nuclear forensics.

Nucleation and crystal growth are fundamentally important to processing nuclear materials and radioactive wastes, but information is lacking regarding the basic science underlying these transformations. Many chemical processes used in the nuclear industry—both for recovering special nuclear materials and for disposing of radioactive wastes—involve manipulating highly concentrated salt solutions. Precipitation of solids from such solutions is often a concern. In this context, precipitation might be desired (e.g., of plutonium oxalate during plutonium oxide synthesis) or not (e.g., during ion exchange or solvent extraction processes).

In this project, we are developing a framework based on recent developments in crystal growth that have been providing insight into two distinct pathways for nucleation and growth. In the "classical" pathway, nucleation occurs through ion-by-ion addition to a growing cluster with increasing free energy of formation until a critical size is reached. At this point, the free energy change for further growth becomes negative and occurs via ion-by-ion addition to kink sites at atomic steps on the crystal surface. The dependence of crystal morphology and growth rate on applied supersaturation is well understood for both nucleation and growth. These dependencies are distinct from those of crystals that form through the second, "non-classical" pathway broadly described as "particle-mediated growth." In the latter scenario, nucleation and growth occur via the addition of higher order species ranging from oligomeric complexes to liquid droplets to amorphous or crystalline nanometer-scale particles. Detailing the morphologies, internal structures, and dependencies of growth rates on supersaturation allows the formation mechanism to be determined and related to process conditions. Moreover, rate measurements versus supersaturation and temperature will allow us to determine the activation barriers that control the nucleation and growth processes.

Our research is focused on the precipitation of oxalate salts. In particular, we are investigating nucleation and crystal growth of Pu(III) oxalate, $Pu_2(C_2O_4)_3 \cdot 10H_2O$, because the process through which this compound forms is of interest to the nuclear forensics community. Due to the challenges of working with the highly radiotoxic plutonium, initial investigations have used Ce(III) and Eu(III) as surrogates for Pu(III), which allows the development of the experimental techniques to examine the plutonium system safely.

A combination of theoretical and experimental techniques is being used to develop a framework for explaining nucleation and the subsequent crystal growth of the metal oxalate systems from the atomistic to hundreds of micrometers scale. Both classical and *ab initio* models are being used to provide direct but complementary information on the mechanisms underlying phase transformations.

This project has pushed the limits of the experimental tools available for investigating nucleation phenomena. In particular, measurement of the evolution of the particle size distribution (PSD) at scales < 50 nm has proved to be very challenging. Much of the work conducted in FY 2017 focused on establishing reliable PSD in this regime, focusing on high-resolution ultrasonic spectroscopy (HRUS) and dynamic light scattering (DLS). However, reliable results have proved elusive for the HRUS and DLS equipment currently available at PNNL. Further investment in this area will be needed beyond the end of this project.

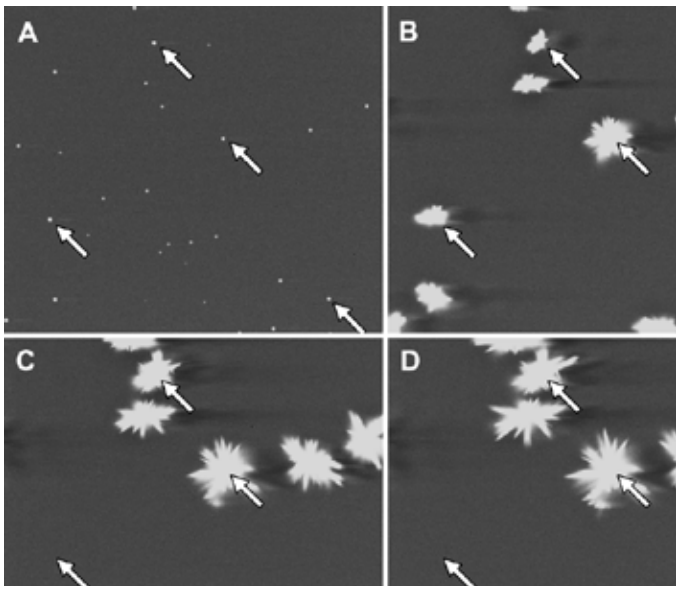
In this project, a variety of experimental and computational techniques have been employed to gain insight into the crystal growth mechanism of $Eu_2(C_2O_4)_3 \cdot 10H_2O$. This report highlights the electron microscopy examinations of this system that have been performed. A combination of high-resolution transmission electron microscopy (TEM), cryogenic TEM, and liquid cell scanning electron microscopy (SEM), has been used to monitor the $Eu_2(C_2O_4)_3 \cdot 10H_2O$ growth over the range of nanometers to tens of micrometers to identify potential crystal growth pathways.

The synthesis of $Eu_2(C_2O_4)_3 \cdot 10H_2O$ was explored under static and stirred conditions. The synthesis was conducted by mixing a solution of $Eu(NO_3)_3$ with a solution of $H_2C_2O_4$. When

PN15085/2760

synthesized without stirring, the $Eu_2(C_2O_4)_3 \cdot 10H_2O$ crystals were composed of multiple rod-like subunits. Examination of these crystals by atomic-resolution TEM showed that these large branched structures are polycrystalline. In contrast, when $Eu_2(C_2O_4)_3 \cdot 10H_2O$ was prepared under stirred conditions, the product was a mixture of individual rods and branched structures.

Cryo-TEM images were obtained for the formation of $Eu_2(C_2O_4)_3 \cdot 10H_2O$ quenched at different points along the reaction timeline. Again, both static and stirred systems were investigated. In the absence of stirring, small, rod-shaped crystals were observed during the first 10 minutes of reaction. Crystal size and branching complexity increased as the reaction proceeded. The first evidence for a branch crystal was observed in a specimen vitrified 12 minutes after initial mixing. A morphology consisting of a cluster of rod-like crystals on the order of approximately 200 to 500 nm was observed after a 20-minute reaction time. This very much resembled the morphology observed in the larger scale SEM images, which showed rods approximately an order-of-magnitude larger.



Backscattered electron SEM images of europium oxalate crystals grown *in situ* in a WetSEM cell. Frames A–D were captured approximately every 2 minutes, beginning 4 minutes after the initial mixing of reagents. Arrows identify the same location in each frame and are used to highlight some of the small crystals that grew between frames A and B. The crystal identified with the lower left arrow disappeared between frames B and C and is thought to have settled at the bottom of the SEM cell, outside the range of the BSE interaction volume. Scale is the same in all images.

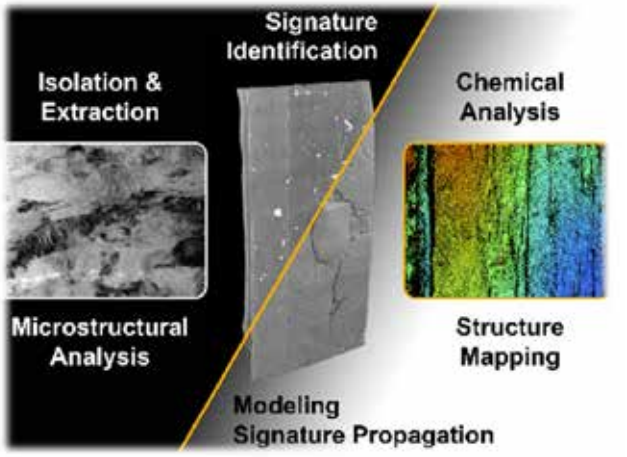
Under stirred conditions, the cryo-TEM investigation revealed extensive networks of tens to hundreds of 2–5 nm particles at 6–7 minutes after mixing. Small rods were also occasionally present and always observed in close proximity to the networks of round particles. Ultimately, complex branched structures were produced.

In situ WetSEM imaging was performed on $Eu_2(C_2O_4)_3 \cdot 10H_2O$ grown under static conditions. The first visible crystals, seen at 4 minutes after mixing the reactants, were about 2 µm in diameter. A number of these grew into branched structures > 10 µm across over a period of 2 minutes. The final crystal morphology varied somewhat when replicates of the same synthesis were performed.

Investigation of the Signatures of Additively Manufactured Objects Using Advanced Chemistry and Materials Science Techniques to Identify Counterfeits

Chris Barrett

The purpose of this project is to develop a focused capability to address the need for additively manufactured device forensics, identify signature rich aspects of different build processes that can be used to distinguish counterfeits, and to determine provenance and predict performance.



The rise of additive manufacturing (AM) in recent years is, in part, due to the open-sourced nature of the printing processes and reduced cost and capital barriers relative to traditional manufacturing. The advent of "desktop" 3D printers has, quite literally, brought manufacturing to the masses. Thus, it is reasonable to assume that AM will lead to an increase in objects of importance to national security, which could be fabricated at a point of use, and to global markets, where an influx of parts and products of unknown origins could lead to rapid destabilization. This democratization of manufacturing has spurred a growing demand from producers and end users to verify the authenticity and quality of individual parts.

To this end, we are developing an anti-counterfeiting approach that utilizes both intrinsic and extrinsic properties of a three-dimensional- (3D-) printed part that can be non-destructively interrogated to validate authenticity. Our parallel efforts to identify and study signatures of AM objects are being directly leveraged in generating part-specific chemical signature data that can be linked to a securitized, distributed, and time-stamped blockchain ledger entry. Validation is facilitated by a QR code, printed directly to the part, which provides a searchable reference to an Ethereum-based blockchain entry using a standard smart phone.

Efforts for the initial phase of this project have been split between identification and attribution of intrinsic signatures of objects fabricated by AM and the incorporation and measurement of additives with unique extrinsic signature profiles. The merits of the first task not only support the anti-counterfeiting objective, but also offer a mechanism by which counterfeits can be identified and provenance determined. Using techniques such as computed tomography, coupled with focused ion beam milling, and electron microscopy, particulates inherently incorporated during (Fused-Deposition-Modeling-) FDM-style printing were studied as a means of correlating parts to specific printers. Pyrolysis gas chromatography mass spectroscopy was also evaluated as an ideal technique for not only polymer identification in SLS prints, but also detection of trace additives and impurities that can be sourced to specific printer feedstock and their associated supplier.

Chart showing the change in RON achieved by sample in comparison to the RON of the original surrogate fuel. Changes in RON are also compared to changes in peak area of a product peak observed in the respective GC chromatogram.

PN17036/2926

In the second case, we recently published a new, low-cost, anti-counterfeiting method for 3D-printed parts by use of a blockchain platform, where inclusion of measured fluorescence data with a digital thread, enhanced part security. Nanocrystals were uniformly doped into polylactic acid (PLA) and printed in selected areas of an object. Example objects were designed to contain QR codes with raised features printed with the PLA composite material. While the measured emission color data in itself is not particularly complex, it becomes significantly more powerful when archived as a component of the digital twin on a blockchain ledger.

Incorporating chemical signature data from extrinsic tags, or intrinsic signatures of a printed part, to complement other digital part metadata on a blockchain database may provide end users, intermediaries, and manufacturers with an easy and low-cost verification tool. As such, the project will seek to explore this approach further, expanding to new methods of handheld interrogation, direct attribution of certain signatures, and production of AM standards for model validation.

Low-Scaling Electronic Structure Methods for Accurate Modeling Chemical Transformations in Complex Environments

Bo Peng

This research is focused on developing low-scaling approximations and parallel algorithms for the expensive, yet accurate, coupled-cluster (CC) method to achieve the microscopic, molecular level characterization and understanding of chemical transformations during the process of photocatalysis, photovoltaics, and metabolism.

Among all available theoretical methods, the CC model, as one of the wave-function-based *ab initio* methods, is capable of providing very accurate results by proper inclusion of many-body effects. In particular, the CC model with single and double excitations corrected by perturbative triples, also known as CCSD(T), has been recognized as the "gold standard" for high-accuracy computational chemistry. However, this method always suffers from very high computational cost due to expensive tensor contractions (e.g., canonical formulation of the CCSD[T] approach scales as $O[N^7]$ with N representing system size). The expertise of PNNL in the development of quantum chemical methodologies and the cutting-edge computational resources provide a unique platform to tackle this challenge. Employing the local character of the correlation effects and applying recent advances in applied math with efficient parallelization strategies can significantly shift the system-size limit tractable by accurate CC methods.

We envision that our CC simulations will be enabled for systems consisting of hundreds/thousands of light atoms and lay the foundations for extension of local CC approaches beyond the ground-state applications, including description of electronic excited states and properties, using linear-response CC formulations.

The CC module in the parallel infrastructure of NWChem open-source software, when running on the Cascade computing cluster, has shown an almost linear parallel scaling for the most computationally expensive N^7 part of CCSD(T) calculations.

Recently, we applied the extensions of the equation-of-motion CC module to accurately calculate the ionization potential (IP) and electron affinity (EA) of a series of carbon nanotubes consisting of hundreds of light atoms. Based on the accurate IP and EA values, we demonstrated that the electronegativity of the more relativistic carbon nanotube systems remains, to a large extent, independent of nanotube length.

Secondly, by rewriting the CC equations in terms of low-rank tensors generated by employing pivoted Cholesky decomposition (CD) together with reverse Cuthill-McKee technique on integral tensors, we successfully obtained a new working equation of CC with doubles (CCD). This new working equation shows a natural compatibility, in comparison with the conventional equation, with the high parallelism and the low storage requirement, which will facilitate an efficient parallel coding in the near future.

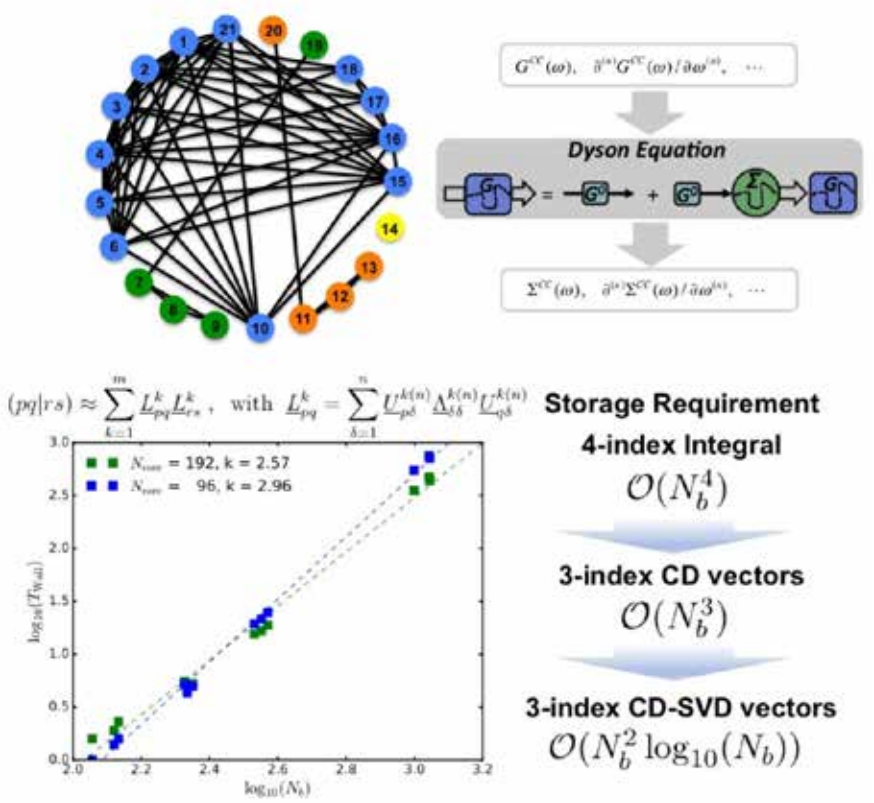
To further reduce the storage requirement, we then combined the pivoted CD with truncated singular value decomposition (SVD) to approximately represent the two-electron integral tensor in terms of low-rank CD-SVD vectors. A comprehensive test has been performed for a series of one-dimensional (1D), 2D, and 3D carbon-hydrogen systems, with the basis sets ranging from approximately 100 up to 2,000. The observed numerical scaling of our implementation shows $O[N^{2.5\text{-}3}]$ versus $O[N^{3\text{-}4}]$ cost of performing single CD on the two-electron integral tensor in most of the other implementations. More importantly, this decomposition strategy can significantly reduce the storage requirement of the atomic orbital two-electron integral tensor from $O[N^4]$ to $O[N^2 \log N]$ with moderate decomposition thresholds.

These results demonstrate high efficiency and scalability of the compound two-step decomposition of the two-electron integral tensor in our implementation. The accuracy tests have been performed using ground- and excited-state formulations of coupled-cluster formalism employing single and double excitations (CCSD) on several benchmark systems, including the C_{60} molecule described by nearly 1,400 basis functions. The results show that the decomposition thresholds can be generally set to 10^{-4} to 10^{-3} to give acceptable compromise between efficiency and accuracy.

PN16102/2879

Finally, we successfully derived the CC formulation of Dyson equation, or Green Function CC, as well as the analytic form of its derivatives, with respect to the frequency. Not only can this formulation accurately describe the propagation of a "particle" added to the ground state of the molecular system, but it also provides an efficient way to accurately compute the self-energy of the "particle" and its derivatives, with respect to the frequency, which is essentially related to the coupling between the "particle" and the system.

In the next year, we are going to combine our developed tensor decomposition technique with the tensor algebra library for many-body problem (TAMM) with the expectation of being able to perform equation-of-motion and linear-response CC calculations to compute excited states and properties of quantum chemical systems described by 5,000–10,000 basis sets. In this level, we will then be able to calibrate other, more ambitious approximations. In the meantime, applications of this new implementation for the ground state and excited state of large molecules will be carried out to obtain some meaningful scientific results.



Making, Measuring, and Modeling Materials for Quantum Computing

Nathan A. Baker

This project aims to develop a basic science understanding of materials pertinent to quantum information technologies and the tools and techniques that can be applied to many of the most promising qubit architectures (e.g., quantum dots, defects in solids, and superconducting Josephson junctions).

Materials science has a well-established approach for designing new and improving existing materials, represented as the "make, measure, model" mnemonic.

Make – to "make," existing PNNL materials synthesis capabilities will demonstrate materials pertinent to quantum information technologies preparation by assembling nanostructured quantum computing (QC) materials using focused-ion beam (FIB) methods and improving QC device fabrication through post-processing to remove photoresist residue from delicate device features.

Measure – to "measure," existing PNNL chemical imaging capabilities will demonstrate characterization of QC materials by developing novel sample preparation methods that enable analysis of unique and delicate structures for materials of interest to the QC field. This also will entail developing new methods in z-contrast scanning transmission electron microscopy (STEM) to characterize structures of interest to QC and using atom-probe tomography to understand the structure and chemical composition of QC structures.

Model – to "model," existing PNNL mesoscale materials modeling and mathematics capabilities will augment QC materials simulations by evaluating the impact of materials defects on the performance of topological QC paradigms.

Part of this project focuses on discreet methods of hierarchical self-assembly to arrange nanoscale building blocks into useful structures necessary to exploit quantum properties of vacancies and defects. Specifically, methods now exist that can functionalize nitrogen-vacancy-containing nanodiamonds to immobilize them onto substrates of interest. This immobilization allows for rational and selective arrangements of single or near-single vacancy-containing diamonds onto architectures useful for quantum information science.

Many relevant parameters for QC occur in materials on the atomic to nano scales. As such, it is important to develop characterization methods on the same scales. STEM is necessary to image materials down to the atomic scale. Yet, the materials initially need to be thinned down to approximately 100-nm thickness. This is most easily controlled using an FIB. However, for some materials, the FIB creates unwanted damage or implantation of gallium (Ga) arising from the Ga source used as the "knife." In exploring FIB specimen preparation and hand-polished specimens, it was noted that, while FIB works best for control and uniformity of sample thickness (even across multilayer systems), hand-polished samples work best for Ga-beam sensitive materials, such as aluminum (Al).

Once in the electron microscope, several parameters have been tested to determine optimal imaging conditions. First, the electron beam's accelerating voltage is varied, and images are taken at 80 and 200 kV accelerating voltages. Al is used as a test material. While Al damages more easily at 200 kV, the lower spatial resolution at 80 kV negates the benefits of reducing the voltage. This will be equally important in other quantum materials with small atomic spacings. Next, the camera length is varied. This value determines how coherent the image is based on the Rutherford scattering of the electrons off the specimen and the angles where they are collected. Notably, while higher angle (less coherent) imaging is desirable, it also reduces the signal-to-noise ratio (SNR) of the images. In terms of balancing SNR to obtain spatial resolution versus removing the unwanted strain contrast, it becomes more complex at interfaces, dislocations, and other defect sites. For Al interfaced with a substrate, this balance has been achieved.

Lastly, the ability to characterize the QC materials spectroscopically has been studied. Electron energy loss spectroscopy (EELS) is especially useful to compare to X-ray absorption f ine structure and density functional theory calculations to more deeply understand a material's electronic structure. However, much like the previous discussion about imaging, the information obtained depends on the material's ability to withstand the electron dose without breaking bonds or rearranging the lattice. The fastest acquisition times have been explored. Still, Al and its supports were unable to withstand the electron dose. Any information obtained will closely

PN15080/2755

depend on the materials being analyzed. For materials that may have trace elements of unwanted contaminants (e.g., precursors such as chlorine) or unwanted dopants, it is important to use high-resolution energy dispersive X-ray mapping to assess all of the elements present before delving into a more detailed analysis of bonding using EELS.

This effort also has included work to develop new bounds for the Egyptian fraction problem—an important mathematical problem dating back thousands of years with applications in electrical engineering, geometry, and topological phases of matter—with few prime factors. The previous best-known bounds were exponential, whereas ours are linear. This dramatic improvement should afford the ability to partially classify integral modular categories through Rank 15 and completely characterize weakly integral modular categories through Rank 7. Both results are of broad interest to the mathematical community and vital to studying topological quantum computation.

This work also has led to the generalization of metaplectic categories, an important class of categories yielding universal computing after incorporating measurement. The underlying phases for these categories are of particular interest to the topological qubit community. The classification of metaplectic categories has been completed, as well as all categories that have the same scaling laws. To aid this work, a Groups, Algorithms, Programming software package has been developed to compute various data associated with these categories and currently is available on GitHub. These classifications are necessary for understanding the properties of near-term topological devices.

A collaboration has been initiated to understand the effects of non-standard lattices in models for topological QC. Preliminary studies show these non-standard lattices can improve system robustness in the presence of noise and defects. The lattices also may have implications in waveguides and metamaterials. Work to investigate how the mathematics used to study topological phases of matter applies to nucleotide pseudo-knots is ongoing.

External researchers also have contributed to understanding how fermions can be incorporated into certain models for topological phases of matter. Specifically, the team has demonstrated half of a conjecture known as the "16-fold way" that outlines how to simplify fermionic models to bosonic models without losing generality. In related work, they have classified all fermionic topological phases of matter with no more than eight particle species.

In addition, new combinatorial approaches to optimal quantum measurements, including investigating the relationship between symmetric informationally complete positive operator-valued measures, mutually unbiased bases, and mutually unbiased subspaces have been explored. Combinatorial techniques have been brought to a classical algebraic/geometric problem, while another collaboration aims to discern the algebraic building blocks of topological phases of matter that succeeded in understanding the algebraic properties of loop braid groups with applications in (3 + 1)-dimensional topological quantum field theory. A related collaboration has succeeded in classifying all topological phases of matter that have, at most, bosonic five particle species.

The team also has classified all multiplicity free fusion categories through Rank 9, and new collaborations have started to study topological phases with simple scaling laws for their Hilbet spaces, so-called *low diversity modular categories*. These categories correspond to topological phases, with the simplest particle collision combinatorics. In some sense, these are the most likely to be realized in near-term hardware.

Another theoretical approach has been implemented to decompose the electron density of a solid over localized directed orbitals, which provides an intuitively convenient representation of bond orbitals between neighboring atoms. This approach has been implemented as an add-on code using the results of NWChem calculations and applied to reveal the character of bonding in PbSe and PbTe, as well as their solid solutions $PbSe_xTe_{1-x}$.

Mimicking the Function of the Enzyme Scaffold

Molly J. O'Hagan

Natural enzymes catalyze many reactions that can be used for sustainable energy applications, such as the interconversions of H_2 and H^+, O_2 and H_2O, or CO_2 and CO, using earth abundant metals and mild conditions, with performance unmatched by synthetic catalysts. We aim to create synthetic catalysts with performance that can rival that of the natural systems by using hybrid systems (i.e. synthetic catalysts attached to protein scaffolds) to identify design principles by which the protein matrix can be used to control reactivity and increase performance.

The chemistry that enzymes catalyze occurs at a specific active site within the protein matrix, but that active site, in the absence of the protein scaffold, cannot do the same chemistry. The protein scaffold has been shown to control important properties such as stabilizing the active site geometry, transport of products and substrates, controlling acid/base chemistry, creating the environment around the active site, and altering redox properties.

Many studies on molecular functional mimics of enzymes have been focused on the role of the first and second coordination spheres, showing a tremendous impact in moving protons, stabilizing structures, and delivering substrates. However, there is still a significant performance gap between the function of these complexes and enzymes performing the same reaction. By incorporating existing molecular catalysts into protein scaffolds, we can make systematic changes that enable the identification of design principles by which the protein scaffold influences efficiencies. The design principles can then be directly translated to create new, efficient molecular catalysts and, ultimately, heterogeneous catalysts.

We are using two approaches to create the synthetic catalyst-protein hybrid systems. One approach is to covalently link a synthetic catalyst to a well-defined protein using bioconjugation techniques, and the second approach is to develop phage display techniques as a high throughput method to quickly screen proteins suitable for binding different synthetic catalysts.

Our strategy for covalent linkage uses a maleimide functional group that reacts with cysteine side chain thiols within the protein to attach the catalyst. We have successfully synthesized a nickel catalyst that has the bioconjugation linker within the ligand backbone characterized by nuclear magnetic resonance (NMR) spectroscopy. We have also successfully expressed and purified three protein targets with cysteine mutation that maintain their native structures.

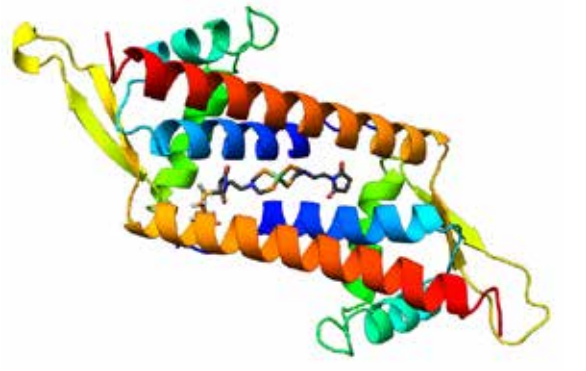
We have successfully covalently coupled the nickel catalyst to the first target, Lactococcal multidrug resistance regulator (LmrR) protein. We have also successfully expressed ^{15}N labeled LmrR and are using NMR structural characterization to analyze the interaction between the protein scaffold and the synthetic catalyst in solution. The parent molecular nickel complex is a catalyst for H_2 oxidation, but when the catalyst is incorporated into the protein matrix, only H_2 production catalysis is observed. The protein matrix enables unprecedented H_2 production activity for this system. Interactions between the protein matrix and the molecular catalyst enable catalysis of the reverse process (i.e., H_2 production) for the hybrid system compared to the molecular catalyst alone (i.e., H_2 oxidation). Continuing structural characterization studies and systematic mutations to the protein matrix will enable us to identify how this change in reactivity is accomplished.

Our phage display approach uses phage viruses that have a coat-protein present on the viral surface where the peptide sequences have been randomized to generate a library of millions of different surface exposed proteins (i.e. an extensive number of potential protein scaffolds). We can then screen this library against different molecular catalysts, enabling a high throughput method to identify protein scaffolds that can accommodate molecular catalysts. We have synthesized three molecular catalysts with biotinylated ligands, which are necessary for binging assays. We used high throughput phage display screening to identify three proteins that bind one of the targets. Phage display screening has never before been used to identify metal catalyst binding motifs in proteins.

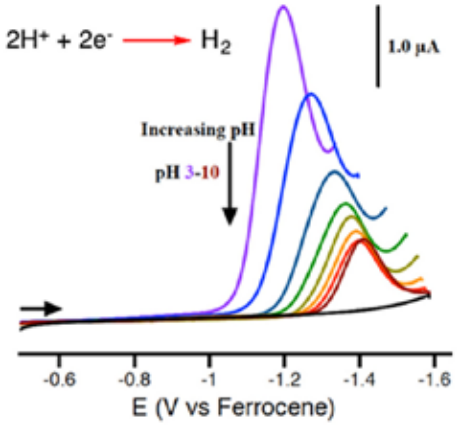
With the hybrid systems in hand, the continuing work will focus on structural characterization of the hybrid system and selective mutations to identify how the protein matrix

PN17061/2951

controls the reactivity of the molecular catalyst. Particularly, we seek to understand how to control substrate delivery, dynamics, pK_a, and catalytic redox potentials using the protein scaffold. Identification of these design principles will enable the design of synthetic homogeneous and heterogeneous catalysts with performance that can match or exceed that of native enzymes.



Hybrid catalyst created by incorporating the $[Ni(P^{et2}N^{et\text{-}maleimide}P^{et2})_2]^{2+}$ within the LmrR protein.



Cyclic voltamograms of H_2 production catalyzed by the hybrid system at various pH. Increasing current, µA, indications increased H_2 produced.

Multimodal Approach for Rapid, Robust, Reliable and Economic Environmental Monitoring

Dev Chatterjee

Availability of a universal, field-deployable sensor platform for rapid monitoring of multiple subsurface analytes will streamline DOE's environmental remediation and clean-up efforts. This project utilizes multimodal analytical approach to develop a single, integrated sensor platform for continuous in situ monitoring of multiple contaminants. The proposed approach is based on electrochemical and optical luminescence techniques, enabling measurements of the analytes of dissimilar nature in wide concentration ranges.

Simultaneous rapid detection and quantification of multiple analytes remain a great challenge in environmental sensing. Detection of an analyte is achieved via recognition by a sensor of its unique property (i.e., intrinsic or acquired), producing a selective response. As the distinguishable properties of dissimilar analytes vary, finding a single sensor module for the detection of different classes of analytes (e.g., chemical versus biological, organic versus inorganic compounds, ions versus neutral species) is challenging.

One approach to designing a sensor that can detect multiple analytes having diverse characteristics is to translate the different responses into a common form of expression, irrespective of the properties of the analytes. To achieve this, the target analytes can be selectively captured by specific probes immobilized on a single electrode platform, resulting in a marked increase of its electrochemical impedance proportional to the analyte concentration. The sensitivity and detection limit of the electrochemical response can be enhanced through a secondary optical luminescence detection mode, which monitors the luminescence intensity during a simultaneous redox conversion of a non-luminescent molecule (***NL***) to its luminescent form (***L***) at the electrode; analyte capture inhibited this conversion through electrode passivation, proportionally lowering luminescence intensity.

A key feature in this approach is that the capture can take place through multiple different mechanisms, so long as the end result of the capture is a change in impedance or luminescence. Therefore, such detection is not limited to any analyte-specific properties and can be applied to a wide array of analytes. The sensor platform universally translates these diverse responses to a common output signal (i.e., impedance and/or luminescence), enabling simultaneous quantification of multiple analytes.

FY 2017 research demonstrated a proof-of-concept of the universal sensor approach using three chemically diverse contaminants. Their detection using the same impedance platform methodology illustrated possibility of extension of this approach to other analytes of interest relevant to DOE missions. Three contaminants found at Hanford and many other DOE sites nationwide, including CCl_4, TcO_4^-, and NO_3^- were selected based on their significant chemical diversity.

The following model analytes exemplify key challenges likely to be encountered in the integrated detection of the multiple analytes by a common platform: 1) no reported *in situ* technique for CCl_4 detection and quantification exists; 2) established methods for *in situ* detection of TcO_4^- and NO_3^- are not suitable for sensing using integrated platform; 3) while NO_3^- detection is most studied due to its ubiquitous nature within environmental, food, and industrial systems, no existing sensor for NO_3^- allows its quantification in wide enough range.

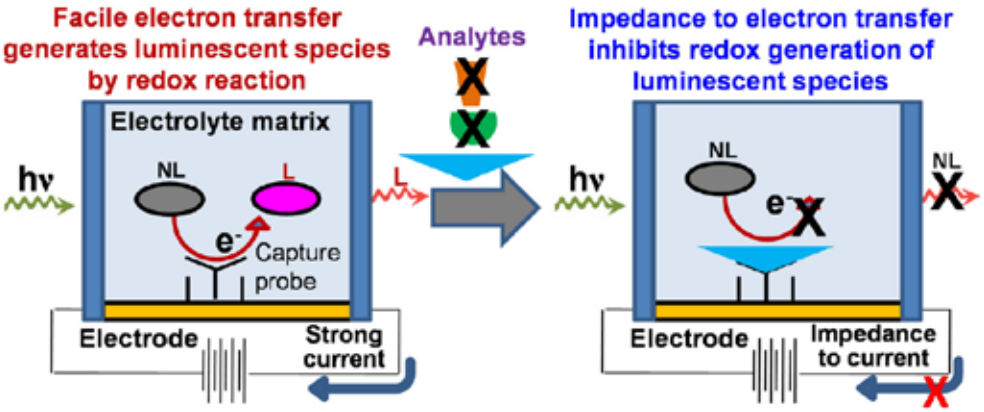
Major project accomplishments include identification of a capture probe for unique recognition of each contaminant and translation of the obtained responses to the common impedance detection. It was discovered that phenylene oxide polymers or cobalt(II) porphyrin complexes incorporated on an electrode surface can serve as selective probes for CCl_4; to the best of our knowledge, these are the first reported examples of the sensor media for CCl_4. Thioalkanes and polypyrrole olymers were identified as selective probes for TcO_4^- and NO_3^-, respectively. Each capture probe recognizes CCl_4, TcO_4^-, or NO_3^- through a different mechanism but generates a common impedance response when incorporated onto the electrode surface, therefore, validating our approach for the universal sensor platform.

PN17051/2941

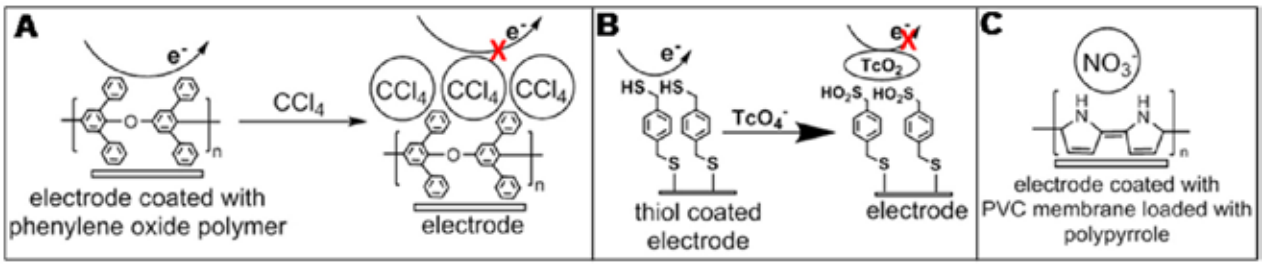
Polymerized phenylene oxide probe is used for unique CCl_4 recognition by dipole-induced/dipole interaction, specific enough to distinguish it from other chlorinated organics such as $CHCl_3$ and dichloroethane. Thioalkane capture probe recognizes TcO_4^- through reductive precipitation based on their mutual redox affinity, resulting in an increase in impedance proportional to TcO_4^- concentration. It was established that other redox active contaminants co-mingled with TcO_4^-, such as IO_3^- and CrO_4^{2-}, do not interfere with this process. Recognition of NO_3^- was achieved using polypyrrole-based ion-selective capture probe designed to have cavities matching size, sterics, and electronics of NO_3^- via template effect and screening out interferences such as Cl^-, SO_4^{2-}, PO_4^{3-}, etc.

We also demonstrated that the impedance-based systems for detection of CCl_4 and TcO_4^- are compatible with potentiometric modality for NO_3^- detection. The proposed approach fills a key technology gap by allowing integration of newly invented sensor-probes (e.g., for CCl_4) with well-established one (e.g., electrochemical NO_3^- detection) on one platform. Further, the potentiometric approach is able to detect concentrations as high as 1 M. Complementary luminescence mode of detection was incorporated in the sensor platform to significantly lower limit of detection. The well-established redox active molecule $[Ru(bpy)_3]^{3+}$ (***NL***) → $[Ru(bpy)_3]^{2+}$ (***L***) was used to serve as the luminescent read-out for highly sensitive analyte detection. The luminescent mode lowered the limit of detection by several orders of magnitude compared to the electrochemical impedance, reaching as low as 10^{-12} M. As a result, the electrochemical/luminescence sensor platform may enable unprecedented detection range for analyte concentrations from 10^{-12} M to 1 M, providing deployment and operational flexibility.

Overall, the investigated approach to the universal sensor platform addresses DOE needs for monitoring of relevant contaminants within the dynamic, multicomponent matrices of Hanford vadose zone and groundwater. It opens possibility of detection and quantification of multiple contaminants and analytes of interest on one platform through the ability to translate dissimilar stimuli to a common response (i.e., electrical or optical). The approach is expandable to other contaminants, as well as for detection of and predictive monitoring of co-mingled contaminants.



Schematic representation of biomodal analyte detection.



Representative capture probes for A) CCl_4, B) TcO_4^-, and C) NO_3^-.

Production of Para-xylene Enabling 100% Renewable Polyethylene Terephthalate

Karl O. Albrecht

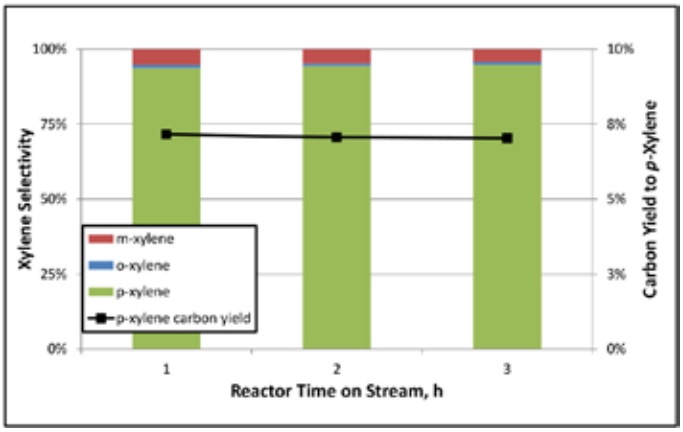
About 56 million tons of polyethylene terephthalate (PET) plastic is generated annually using petroleum-derived para-xylene (pX) for products such as pop bottles and clothing fibers. We have demonstrated a new process to selectively produce pX derived from renewable C_4 alcohols.

2,3-butanediol (BDO) is a renewable chemical produced biologically from the fermentation of sugars or CO-rich waste gas. Previous work at PNNL developed methods to convert BDO to several different C_4 derivatives, including 1,3-butadiene, methyl ethyl ketone (MEK), acetoin, methyl vinyl carbinol (MVC), 2,3-butanedione, and 2-butanol.

PNNL is building a portfolio to convert renewable intermediates, such as BDO or its derivatives, into high volume chemicals such as pX. pX is currently produced from petroleum-derived naphtha. A mixed stream of aromatics and paraffins is processed to polymer-grade pX through an energy-intensive, capital-intensive aromatics separation plant. To reduce separation costs, it is desirable to produce renewable pX at above 85% selectivity relative to other xylene isomers. Highly selective methanol conversion to pX over modified ZSM-5 catalysts is known in the literature, but no accounts of BDO or its derivatives as feedstocks were identified, suggesting a substantial opportunity. We have investigated two routes in parallel to selectively produce pX: 1) conversion of BDO and its derivatives over modified ZSM-5 zeolite catalysts; 2) directed aldol coupling reactions of BDO derivatives.

Our first approach to converting BDO to pX involved reaction over modified ZSM-5 zeolite catalysts. The first catalyst tested, a P/Zn/Si-modified ZSM-5, generated an equilibrium concentration of pX at about 24% selectivity versus the *ortho*- and *meta*-xylene isomers. The second catalyst tested, a Mg/Zn/Si-modified ZSM-5 catalyst, demonstrated 88% selectivity to pX versus the other xylene isomers. BDO produced pX for only less than one hour, which was attributed to substantial catalyst coking. MEK was observed as the primary product when feeding BDO both during the first hour of reaction when pX formation was active, as well as during the later portions of the test. These initial runs served to demonstrate that we could indeed selectively produce the *para* isomer of xylene.

We continued our investigations with other BDO derivatives in order to understand if alternative starting compounds would provide greater pX yields using the same Mg/Zn/Si-modified ZSM-5 catalyst. MEK, acetoin, MVC, and butanedione all produced pX at high xylene selectivity during the first hour of reaction. However, in each case, no pX was observed during the second hour of reaction, again likely due to substantial catalyst coking. Surprisingly, 2-butanol produced pX at 93–95% xylene selectivity over the course of 3 hours, which is the best result obtained during the zeolite portion of our investigation.



2-butanol produced pX at a consistent yield over the duration of a 3-hour test at greater than 93% xylene selectivity.

Our second approach to selectively generate pX was through an aldol process, whereby butanedione is selectively dimerized to produce 2,5-dimethyl quinone, which may then be deoxygenated and aromatized to produce pX. Butanedione was found to readily dimerize over both acidic and basic catalysts. Proline catalyzed the condensation at room temperature to give a nearly quantitative yield of open-chain dimer. Lewis and Brønsted acid catalysts such as Fe(III) and sulfuric acid also formed the dimer in lower yields, probably limited by equilibrium. Products of cyclization, via internal aldol condensation, and dehydration to the quinone, such as by heating in dilute sulfuric acid, were not identified. Such cyclizations are known in the literature, and it is possible that steric hindrance makes the cyclization difficult. Higher temperatures or a higher acid concentration might be required.

PN17083/2973

More investigation into the cyclization step is warranted to determine the viability of the ring closing step.

We also investigated the reduction and deoxygenation of the 2,5-dimethyl quinone to ensure the final process step is feasible. 2,5-dimethyl quinone is a commercial product. The quinone was successfully hydrogenated over various Pt catalysts under a variety of conditions. Using THF as the solvent, a 10% feed concentration, 150°C, and 500 psig H_2 pressure, 88.7% conversion of the quinone was observed with 97.1% selectivity to the desired 2,5-dimethyl-1,4-benzenediol. With a high yield of the desired reduction product, and assuming that aromatization will be successful, the viability of the quinone as a target intermediate is supported.

Quickstarter – A Crowd-Sourced Selection of Novel Ideas for DOE Missions

Kelly O. Sullivan

In this project, PNNL staff members submit creative small proposals that align with Department of Energy/Department of Homeland Security mission space and then collaborate with their colleagues to determine which projects move forward.

Crowdsourcing is a funding model originally developed on the internet, where individuals solicit contributions from internet users. One of the best-known examples of crowd-funding is Kickstarter (www.kickstarter.com). In this project, we employ crowdsourcing in the form of an innovation contest to provide a way for PNNL staff members to explore new research areas.

Each proposal is required to convey the significance of the research and describe the importance and relevance of the proposed project. After all proposals are submitted, each staff member participates in selecting which projects move forward via "pledging" of virtual dollars. Those proposals that achieve their goal are then funded as a task under this project.

In winter of FY 2017, a competition round was held and ten proposals were selected by the "crowd" to move forward. More than 40% of the Laboratory's staff members participated in making the selections. The selected proposals included a variety of research areas, ranging from catalyst development to enhanced oil recovery to X-ray diffraction characterization of transuranic crystals. Outcomes from these small efforts have provided data that are being used to develop proposals and as part of publications in peer-reviewed journals.

At the end of FY 2017, another competition was held, and a total of fifteen proposals were selected by the crowd to start work in FY 2018. This time, more than 45% of the Laboratory's staff members participated in selecting the winning proposals. Again, there was wide variety in the topic areas, from oil spill cleanup technology to three-dimensional printing to hydrogen storage technologies. These fifteen tasks will complete by the end of FY 2018. In addition, another round of competition is envisioned during the winter of FY 2018.

A significant goal of this project is to engage Laboratory staff members in collaboration and innovation working together in a transparent way to allocate resources. The participation of such a large percentage of the Laboratory's staff members indicates success. We are now working to help other national laboratories implement this type of program at their sites.

PN17073/2963

Scalable Synthesis of Spinel Stabilized Metal Catalysts

Yong Wang

We are identifying, investigating, and developing new kinds of stable, active, supported metal catalysts that could be used by various industrial applications such as selective oxidation reactions and vehicle emission catalysts components.

Recently, we reported that $MgAl_2O_4$ support with controlled {111} facet can stabilize platinum catalysts under harsh reaction conditions. Conventional synthesis approaches via mechanical mixing or aqueous co-precipitation of precursors typically lead to significant inhomogeneity in the amorphous mixture of magnesia and alumina and, consequently, usually produce the spinel phase with inhomogeneity in local composition, structure, and morphology. We previously showed that $MgAl_2O_4$, synthesized via a novel non-aqueous approach, is capable of stabilizing platinum particles during severe aging at 800°C. We revealed that the superior stability is due to strong interactions between spinel surface oxygen and epitaxial platinum {111} facets. However, the relationship between synthesis methods and surface properties of $MgAl_2O_4$ spinel need more detailed systematic investigation and understanding. In particular, the effects of solvents, precursors, and Mg to Al ratio on the surface chemistry of the produced spinel, as well as their relationships, need to be developed before scaling up the synthesis technology.

The goal of this project is to identify and develop scalable synthesis approaches for producing spinel $MgAl_2O_4$ with well-controlled surface properties and to understand the fundamental synthesis-structure-properties relationships. Spinel $MgAl_2O_4$ is a good model support material and can provide a base for the general spinel family for providing ultra-stable metal catalysts. These catalysts have attractive hydrothermal stability and have potential applications in several important industrial reactions such as biomass conversion and vehicle emission abatement. Ultimately, we expect to identify and develop a general scalable and low-cost synthesis approach that can produce homogenous and hydrothermal stable $MgAl_2O_4$ catalyst support capable of stabilizing metal catalysts under reaction conditions.

In the past two years, we have identified that stabilization of precious metals on spinel $MgAl_2O_4$ is due to lattice matching that brings strong epitaxial interaction between oxygen terminated {111} facet of $MgAl_2O_4$ and Pt {111} facet. This is further supported by observations of similar stability of Pt nanocrystals supported on the spinel of $ZnAl_2O_4$ that has similar lattice parameters with $MgAl_2O_4$. These suggest the proposed lattice matching stabilization mechanism is a general phenomenon.

We further identified the surface properties of spinel crystals, such as acidity and basicity, surface charge, and surface composition, depend strongly on the synthesis methods. Since these factors are believed to determine the initial dispersion and the stability of the supported metal particles, we employed a series of well-controlled synthesis approaches that aim at understanding roles of solvents, molar ratios of Mg and Al precursor, and cation doping in order to correlate the synthesis-structure-properties relationships. The surface properties of the synthesized $MgAl_2O_4$ support, particularly with respect to its ability to disperse and stabilize Pt nanocrystals, are systematically characterized with nuclear magnetic resonance, Brunauer–Emmett–Teller, X-ray diffraction (XRD), NH_3 and CO_2 temperature-programmed desorption (TPD), zeta-potential, thermogravimetric analysis, and density functional theory (DFT) methods. The stability and activity of the catalyst are examined with the model reaction of CO oxidation, which has important implications in vehicle emission abatement, since stability under high temperature is required.

The type of solvent used in the synthesis was found to have critical influences on the properties of $MgAl_2O_4$. We previously established the use of ethanol instead of water produce mixture of precursors with much better homogeneity due to improved control on precipitation. Here we further extended to the use of isopropyl alcohol (IPA) as the solvent. The use of IPA provides better solubility of the Al precursor (aluminum isopropoxide) and improved control on its hydrolysis and condensation and, therefore, produces more homogeneous $MgAl_2O_4$. Experimental results show that IPA-$MgAl_2O_4$ has the highest preferred orientation toward spinel {111} facet and best capability on stabilizing Pt nanocrystals compared with those synthesized in ethanol and water. These results agree well with our original hypothesis proposed in this project.

The role of molar ratios between Mg and Al on the surface properties of $MgAl_2O_4$ was also studied. Since the inhomogeneity of $MgAl_2O_4$ synthesized with conventional methods naturally suggests occurrence of local Mg to Al ratios away from the stoichiometric, it is meaningful to correlate how variations of this ratio affect the surface properties of $MgAl_2O_4$ and its ability on stabilizing Pt nanocrystals.

To do this, a series of well-defined spinel supports were synthesized via the new IPA method with controlled molar ratio of Mg and Al precursors (from 1:1 to 1:4), and their surface

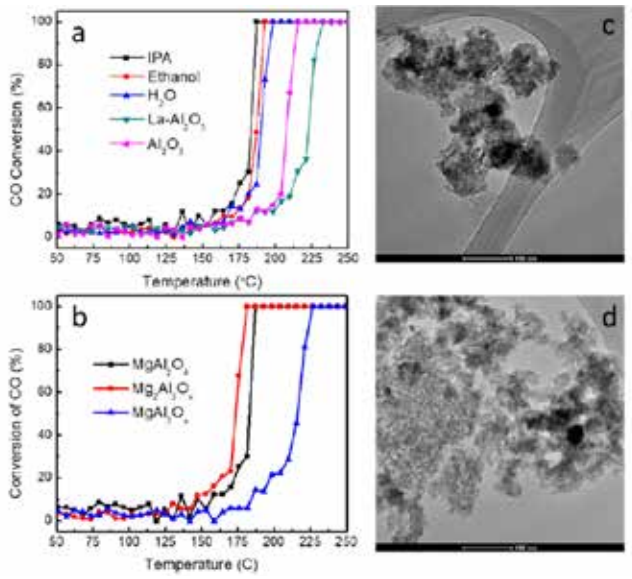
PN15020/2695

properties were carefully analyzed. Preferred orientation studies with both experimental XRD and theoretical DFT methods both suggest increased preferred orientation toward the {111} facet as the ratio of Al was increased. In addition, the increases in Al ratio also result in increases in surface acidity and changes in O 1s binding environment, as observed with CO_2-NH_3 TPD and X-ray photoelectron spectroscopy analysis, respectively. Catalysis studies show that Pt nanocrystal has substantially reduced stability when supported on supports with increased Al ratios, and significant Pt sintering was observed particularly with the case of the 1:4 ratio. These results suggest, in addition to lattice parameters and facet type, the surface composition also play critical roles in stabilizing Pt nanocrystals and the surface should optimized toward homogeneous composition with less Al than stoichiometric ratios for better Pt nanocrystal stability.

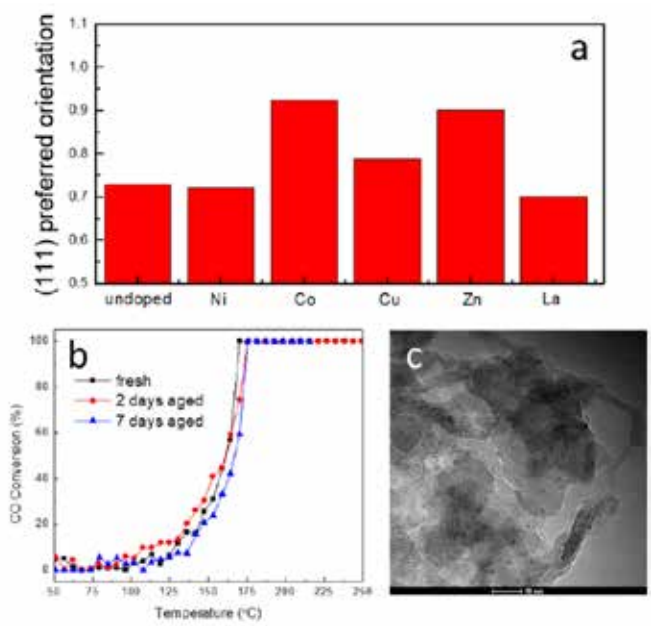
Another new approach that has been introduced into this project to better control the surface properties of $MgAl_2O_4$ is the use of cation doping. The unique crystal structure of $MgAl_2O_4$ allows for doping with divalent and trivalent cations with substantial levels without compromise its crystal structure. Supports with doping levels up to 10% with several representative cations have been successfully synthesized by far and their properties have been carefully examined. Our results show the introduction of dopants provides new opportunities on systematic manipulation of the surface chemistry of $MgAl_2O_4$. The preferred orientation of the {111} facet, for example, shows levels of control unachievable with either control of solvent type or Mg:Al ratios, as discussed above. Preferred orientation close to 1.0 was obtained with the use of Co^{2+} and Zn^{2+}. More importantly, the binding between Pt nanocrystals and spinel facets could be further strengthened with this new doping approach. In fact, significantly improved stability of Pt nanocrystals was observed to the level that the catalytic activity had almost no decay after severe decay from aging in air at 800°C for one week.

In addition to the surface chemistry of $MgAl_2O_4$, we also studied the approach of stabilizing Pt nanocrystals through the use of a secondary metal. Our results show that, when Pt was used together with Ir, the Ir-Pt bimetallic catalyst could undergo surface enrichment of Pt in oxidizing atmosphere. In particular, an up to three time increase in Pt exposure was observed.

In summary, this project was able to establish the relationship between solvent, precursor ratios, cation doping, and the surface chemistry of $MgAl_2O_4$ with respect to stabilize Pt nanocrystals. Substantially improved Pt nanocrystal stability over our original reported $MgAl_2O_4$-EtOH system was demonstrated with the use of IPA as the solvent and increases in Mg molar ratios. More importantly, the best stability was demonstrated with the use of cation doping, and in this case, almost no decay was observed. Even though most of the work was performed with IPA as the solvent, the new insights on $MgAl_2O_4$ surface chemistry developed in this project are potentially applicable to the case of water synthesis and, therefore, could bring new opportunity in scalable and low-cost production of modified $MgAl_2O_4$ for stabilizing precious metal catalysts.



Improved Pt nanocrystal stability on $MgAl_2O_4$ through a) use of different solvents on spinel synthesis. This plot compares activity of Pt nanocrystal toward CO oxidation after aging at 800°C for 2 days; b) for control of Mg to Al molar ratio, the best stability was observed with the ratio of 1:1.5. c) and d) TEM image of Pt-spinel composites after 2 days of aging at 800°C for the ratio of 1:1.5 and 1:3, respectively. Pt particle sintering was observed in the latter case.



New approach of controlling spinel surface properties through cation doping: a) doping with different cations shows unpreceded control toward preferred {111} facet orientation; b) Pt stabilized on Cu-doped spinel shows remarkable stability with almost no decay on catalytic activities after aging at 800°C for 7 days, and c) TEM image of the Pt-Cu-doped spinel after 7 days of aging, and Pt nanocrystals with size ~1 nm were clearly stabilized.

PN15020/2695

Sequence-Defined Polymers Based on a New Backbone Architecture

Jay W. Grate

This project is investigating a novel, bottoms-up, self-assembly approach to achieve controlled design and engineering of polymer-based nanostructures, based on new, sequence-defined polymer architecture.

Sequence control in natural biopolymers leads to tremendous structural diversity and functionality. Realizing the same functionality in stable synthetic polymers will revolutionize materials science. Achieving this challenge will require experimental synthesis of new polymer architectures with sequence control, molecular simulations to predict and understand macromolecular behavior, and investigation of self-assembly of defined polymer structures.

Sequence-defined polymers in nature are epitomized by proteins and DNA, both of which are composed of a multiplicity of monomers, each monomer distinguished from one another by having a different side chain, and the various monomers sequenced into polymers in a predetermined order. In nature, sequence-defined polymers create biomaterials, encode information, perform biocatalysis, participate in molecular recognition, and shuttle species across membranes. More stable sequence-defined polymers prepared by chemical synthesis likewise have the potential to create tremendous structural diversity and functionality based on monomer structure, sequence, and self-assembly.

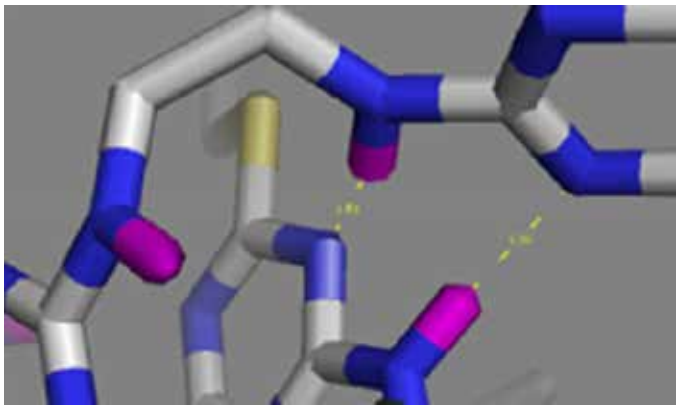
In this project, we have invented, patented, and published a new class of stable, synthetic, sequence-defined polymers. These new macromolecules are based on molecular precursors prepared from cyanuric chloride and assembled into polymers by a solid-phase synthesis method. These new sequence-defined polymers have no precedent in nature and little precedent in chemistry literature; we call these triazine-based polymers (TZPs).

We successfully demonstrated the synthesis of molecular precursors that have a diversity of side chain structures and synthetically sequenced them into macromolecules using a submonomer solid-phase synthesis technique. The side chain diversity typical of peptides and proteins, including ionizable groups, can be achieved in these new synthetic macromolecules. The predetermined sequences range from arbitrary defined sequences to block copolymers, all of defined length. Several methods were used to characterize the new polymers to prove that the sequences had been achieved. Proton nuclear magnetic resonance demonstrated that the side chain content was as planned in the synthesis, and tandem mass spectrometry methods proved that the monomers were, indeed, in the planned sequence.

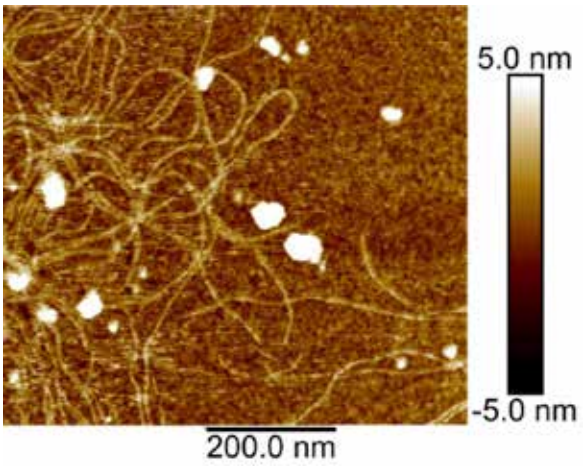
In parallel with the experimental research, molecular dynamics simulations have been carried out that show protein-like patterns of noncovalent interactions between monomer units of the polymers, including paired hydrogen bonds. Nanorod structures found in simulations are stabilized by organized arrangements of hydrogen bonds and pi-pi interactions. These structures bear similarities to peptide beta sheets or helices and are, thus, biomimetic. The interactions found in these simulations suggest that TZPs will lead to self-assembled materials, a prediction that has recently been confirmed experimentally. Using side chain sequences selected to mimic beta-sheet like motifs in peptides, a number of one-dimensional (1D), 2D, and 3D nanostructures have been observed using *in situ* atomic force microscopy and transmission electron microscopy methods.

This work has demonstrated 1) side chain diversity, 2) synthesis of macromolecules with momomer units in defined sequences, 3) noncovalent interactions that can lead to intra- or intermolecular interactions and, hence, conformational behavior or self-assembly, and 4) self-assembled nanostructures.

PN15018/2693



Paired hydrogen bonds between monomers in a triazine polymer conformation.



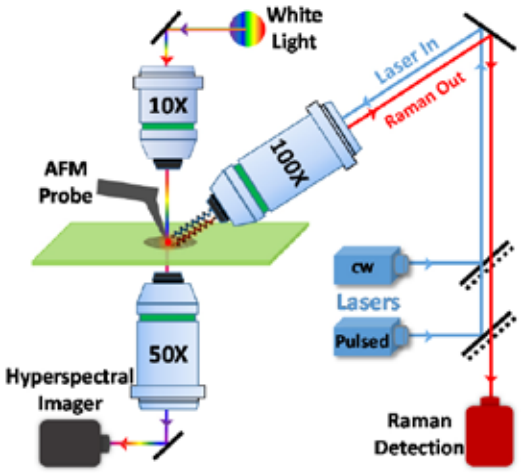
Self-assembled nanofibers composed of sequence-defined triazine-based polymers.

Spectrally Resolved Nanoscale Imaging of Single Molecules, Plasmons, and Their Interaction

Patrick Z. El-Khoury

This project is aimed at the development of ultrasensitive multimodal (bio)chemical imaging technologies that may be used to interrogate single molecules and metallic nanostructures, as well as to advance our understanding of molecule-plasmonic metal interactions. The potential applications of the envisioned technologies are far-reaching and span the realms of several disciplines.

The proposed research efforts are conceived with the hypothesis that a single molecule is the ultimate probe of its local environment. Currently, it is possible to establish the chemical identity of a handful of substances at the ultimate detection limit of a single molecule (i.e., the sensitivity required to identify 1.66 yoctomoles (10^{-24} moles), using surface-enhanced Raman scattering (SERS). Recent work has also demonstrated the possibility of recording chemical images of select classes of individual molecules using tip-enhanced Raman scattering (TERS).



Nonetheless, the mechanisms involved in attaining single molecule detection sensitivity using both techniques are not fully understood; the unique interplay between individual molecules and their distinct local environments complicates their optical signatures and images. In this regard, persistent gaps in knowledge in this field significantly hindered efforts aimed at generalizing and transcending these most promising techniques beyond their proof-of-principle stages.

Our above-stated hypothesis and proposed approach to SERS/TERS are unique. Namely, rather than using these spectroscopic tools to probe molecular properties, we employ the SERS/TERS signatures of a single molecule to probe its local environment and to expose the mechanisms involved in attaining yoctomolar detection sensitivity in Raman spectroscopy. The work will ultimately enable 1) fingerprinting libraries of (e.g., [bio]chemical agents) with single molecule detection sensitivity, 2) monitoring chemical reactions (e.g., bond breaking and making) with single bond detection sensitivity, and 3) probing charge flow across metal nanojunctions, potentially with single electron detection sensitivity.

Several noteworthy technological and conceptual advances were enabled through this project in FY 2016 and FY 2017. We developed and demonstrated a unique approach to hyperspectral optical microscopy, achieved by coupling a hyperspectral imager to various optical microscopes.

Hyperspectral fluorescence micrographs of isolated fluorescent beads were first employed to ensure spectral calibration of our detector and to gauge the attainable spatial resolution of our measurements. Subsequently, different science applications of our instrument were described. Namely, spatially over-sampled absorption spectroscopy of a single lipid (18:1 Liss Rhod PE) layer revealed that optical densities on the order of 10^{-3} can be resolved using our setup by spatially averaging the recorded optical signatures. This was followed by three applications in the general areas of plasmonics and bioimaging. Notably, we deployed hyperspectral absorption microscopy to identify and image pigments within a simple biological system: namely, a single, live *Tisochrysis lutea* cell. Overall, our initial work paves the way for multimodal spectral imaging measurements targeting a broad variety of systems ranging from single cells to plasmonic metal nanostructures.

We coupled the above-described hyperspectral optical microscopy module to a state-of-the-art Raman nanoscope. Using this multimodal chemical and topographic imaging platform, we illustrated how molecular Raman scattering (in the TERS scheme) may be used to visualize several aspects of the local environments in which molecules reside. More specifically, we developed a non-standard application of TERS, whereby various aspects of plasmon-enhanced electric fields localized at nanometric asperities sustained on a sputtered silver substrate were imaged. This unique approach comprises a logical extension to classical TERS spectroscopy and imaging

PN16093/2870

measurements, in which practitioners are primarily focused on nanoscale chemical imaging applications.

We demonstrated ultra-high spatial resolution (sub-nanometer) using atomic-force-microscopy-based TERS electric field imaging measurements performed under ambient laboratory conditions. We also illustrated that optical rectification at plasmonic tip-sample nanojunctions play a significant role in TERS spectroscopy and imaging. Further, our results alluded to the exciting prospect and possibility of imaging the vector components of plasmon-enhanced local electric fields.

On the theoretical/computational front, we explored underappreciated chemical phenomena in ultrasensitive SERS and TERS. We observed a fluctuating excited electronic state manifold, governed by the conformational dynamics of a molecule (4,4'-dimercaptostilbene, DMS) interacting with a metallic cluster (Ag_{20}). This affected the simulated single molecule Raman spectra, as the time trajectories of a molecule interacting with its unique local environment dictates the relative intensities of the observable Raman-active vibrational states. *Ab initio* molecular dynamics (AIMD) of a model Ag_{20}-DMS system were used to illustrate both concepts in light of experimental results. On-going work is aimed at bridging the gap between conventional (statistically sound), AIMD-based Raman spectral simulations aimed at recovering the spectral response of molecular ensembles, on one hand, and non-standard, time-domain simulations that are needed to recover the optical response of a single scatterer. These concepts will be expanded in FY 2018. Furthermore, as similar effects govern ultrasensitive TERS nanoscopy measurements, we will further develop this framework to understand nanoscale chemical images recorded using this powerful technique.

Throughout FY 2018, several coupled technological-conceptual advances are also envisioned. Beyond ultrasensitive nanoscale chemical imaging, we expect our FY 2018 developments to bring us a step closer to following physical and chemical transformations, potentially with yoctomolar detection sensitivity.

Several approaches will be explored (in parallel) to induce and probe chemical transformations in the TERS geometry. The first approach is motivated by a recent realization that light incident onto metallic nanojunctions can be rectified; this leads to molecular charging at plasmonic tip-sample nanojunctions. Using this approach, it is possible to identify the characteristic Raman spectral changes of neutral/charged reactants versus products, but following chemical changes in real space-time (i.e., with joint nanometer-picosecond [nm-ps] resolution) is not feasible.

The second approach is distinct in this regard. It involves coupling ultrafast laser pulses into the existing TERS setup, whereby chemical changes will be tracked with joint nm-ps precision using novel tools of pump-probe nanoscopy. Overall, the proposed FY 2018 developments will enable us not only to fingerprint (bio)chemical targets, but also to follow chemical transformations with joint nm-ps precision, all while operating in the single/few molecule regime, whereby the interaction between analytes and their local environment fully exposes the effect of heterogeneity on chemical properties and reactivity.

Synthesis of Tunable Electro-Catalysts for Biomass Conversion

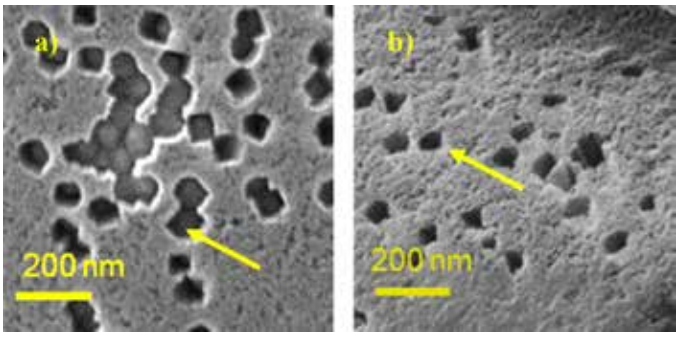
Miroslaw A. Derewinski

The goal of this project is develop a library of robust heterogeneous electro-catalysts composed of metal nanoparticles supported on carbon felts capable of electrocatalytically hydrogenating (ECH) oxygenates such as aromatic aldehydes, ketones, and phenolic compounds primarily present in the lignocellulose and waste oil and fats. The design of active sites and transport pathways is based on the principles deduced from enzymes and theoretical predictions.

ECH is a promising route for low-temperature bio-oil treatment, wherein renewable electricity can be used to store hydrogen in hydrocarbon energy carriers that can be used as transportation fuels. ECH comprises electrocatalytic water splitting, producing *in situ* adsorbed hydrogen on the catalyst surface that reacts with organic substrates. The conversions are ideally achieved under mild conditions, wherein the suitable electro-catalysts minimize the over-potential required for ECH. The feedstock is extremely complex in composition and functionality; additionally, the presence of solvents, such as water, along with other organic solvents, adds an unprecedented level of complexity to the catalyst design and synthesis. Several challenges, such as stability, activity, and reproducibility, limit viability of commercially available catalysts.

To address these challenges, we synthesized a family of catalyst composed of both noble metal (Pt, Pd, Rh) and base metal (Ni, Cu) on the carbon felts. The catalyst synthesis route was composed of two steps—functionalization of carbon felts, followed by wet impregnation technique—to introduce the metal phase to carbon backbone. Since the carbon felts have low surface area (approximately 1 m^2/g), prefunctionalization of the felt is important, since it enhances surface area (increases up to 30 m^2/g) of parent carbon felts and also introduces desirable functional groups, such as -OH, -COOH, etc., that are critical in order to obtain highly disperse metals on carbon felt.

In FY 2017, we have accomplished the synthesis Pt, Pd, and Rh nanoparticles on the carbon felt, which includes the introduction of $Pt(acac)_2$, $Pd(OAc)_2$, and $Rh(OAc)_3$ as respective metal precursor and their subsequent calcination under air at 400°C to realize the metal phase. The thermal activation of the metal precursors also has profound impact on the carbon structure. The carbon felts were found to be significantly roughened compared to their parent structure. Significantly, in case of Pt and Rh catalyst, local oxidation of the carbon structure resulted in embedding the metal particles, which enhances the anchoring the metal particles. Thus, we hypothesized that the extent of oxidation of carbon structure is primarily dependent on the nature of metal. Both metal and metal oxide phases were found to be present in all these systems.



Helium ion microscopy image of a) Rh and b) Pt nanoparticles synthesized on carbon felts.

In FY 2017, we also adapted a diverse synthetic route to prepare size-controlled metal particles and to deposit them on the carbon felt. In one case, chemical reduction method was used to synthesize metal colloids, wherein the size of the particles were controlled (4–10 nm) by tuning the different reaction parameters. The preformed metal nanoparticles were then deposited on the carbon felt, followed by appropriate thermal treatment to remove the capping agent from the metal surface. This method was further developed to prevent aggregation of metal particles.

The synthesized series of metal electro-catalysts showed promising catalytic activity towards hydrogenation of aldehydes and ketones.

In FY 2018, we will develop synthesis of electro-catalysts, as well as modification and characterization of felts for evaluation of catalytic activity. Specifically, we will 1) synthesize particles with specific particle size and incorporate them in the carbon felt, 2) establish the effect of size and morphology of nanoparticles on their ability to catalyze electrochemical reduction of carbonyls, 3) further utilize novel synthetic techniques to modify the carbon felt by developing controlled deposition of nitrogen substrates and porogens to enhance surface area, 4) utilize advanced characterization techniques, such as transmission electron microscopy, helium ion microscopy, and extended X-ray absorption fine structure spectroscopy, for providing structure property relationships.

PN17024/2914

Theoretical Investigation of Low Temperature Electrocatalytic Hydrogen Addition

Roger J. Rousseau

The project is to build new scientific capabilities to model emergent phenomena at the solid liquid interface that will enable us to design new and more efficient catalytic processes.

Most, if not all, theory work in the literature describes catalysis in terms of thermodynamics or simple structural models that lack the necessary physics to describe solid liquid interfaces at finite temperature under an external electrical bias.

Our goal is to build the capabilities to include the explicit description of the solvent, the effect of external bias of structure and bonding at these interfaces, the role of charge transfer (current), and kinetics.

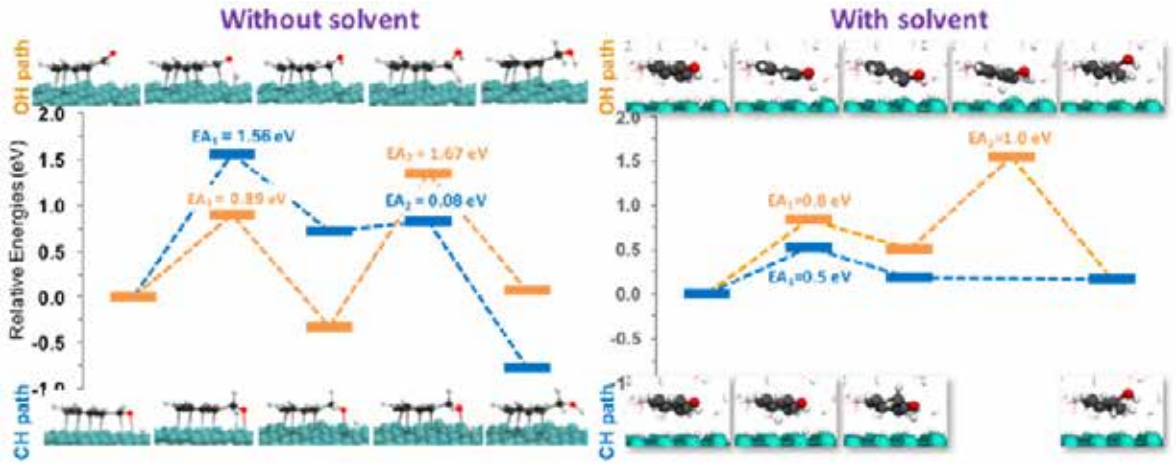
In the past FY, we laid the basis for our integrated tool kit by reproducing the capabilities of others and applying them directly to the chemistry of the electrochemical addition of hydrogen to organic molecules, which are energy poor. We also investigated the role of an explicit description of the solvent inclusion of charge on the electrode surface.

Our major finding was that, although simple models based on primitive thermodynamic reasoning predict trends in electrochemical over-potential (and, therefore, provide some insight into catalyst efficiency), they do not provide insights into the kinetics of a system. Because of this, they cannot predict critical phenomena such as conversion rates or selectivities and faradaic efficiencies.

For instance, the common approach cannot predict the relative energetics of CH versus OH bond formation for hydrogenation of simple probe molecules such as benzaldehyde. This is due to the fact that one needs an explicit representation of solvent molecules to account for the relative energetics of bound species at a surface.

Likewise, predictions of selectivity require intermediate details of how many molecules are bound to an electrode interface, as well as their orientation, in order to predict faradaic efficiencies.

In short, the approach commonly used in the literature is good for quick, qualitative screenings of catalysts but cannot provide the necessary quantitative insights needed to develop truly new catalytic materials pathways. We have taken the first steps in isolating the critical missing elements and will work to develop them into a more robust framework, including current and kinetic effects in the coming year.



Comparison of the relative energetics of catalytic pathways between the standard approach and our approach. Nernstian computational approach commonly used for theoretical electrochemical studies (left) and fully *ab initio* atomistic solvent models (right) for relative energetics of CH vs OH bond formation show completely different energetics, particularly when charged intermediates are involved, as in the OH bond formation pathway.

PN17026/2916

Theoretical Studies of Metal Complex Degradation Products and Their Associated Signatures in the Plutonium Separations Process

Neil J. Henson

Understanding the degradation of chemical species used in the plutonium production processes can be used to better elucidate the dominant chemical pathways present in the processing of nuclear material and enable increased efficiency through improved process design.

A full elucidation of the chemical pathways resulting from breakdown of the coordinating ligands used in the PUREX process has not been performed. The degradation products from these reactions potentially allow a better understanding of why breakdown occurs and the consequences of these species on the efficiency of the overall separations process. This project aims to study computationally possible chemical reaction pathways and compare predicted spectral signatures with available experimental data.

In the first year of this project, we calculated the chemical structure of the tributyl phosphate (TBP) molecule and the chemical breakdown products, dibutyl phosphate (DBP) and monobutyl phosphate (MBP), and their associated vibrational spectral signatures using quantum chemical methods as implemented in the ADF software. We also calculated the vibrational spectra for TBP complexes with nitric acid for comparison with recent Raman data collected on a related project. We focused on the vibrational frequency of the phosphoryl bond stretch (P=O), which is a sensitive measurement of the chemical structure and environment of this class of molecules.

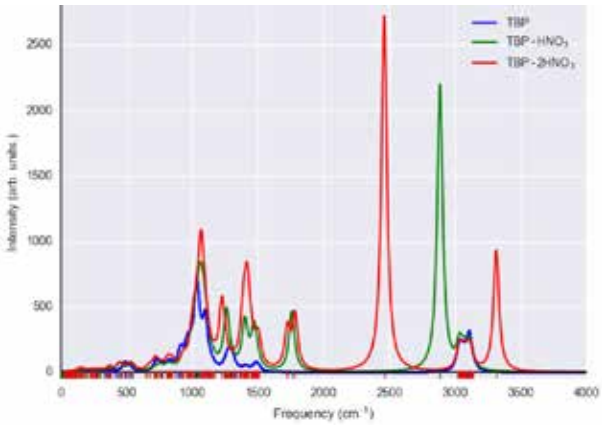
Calculated vibrational frequencies for the P=O stretch.

Molecule	ν(P=O) (cm^{-1})
TBP	1301
DBP	1332
MBP	1342

It has been shown previously that the stretching frequency of molecules containing the phophoryl functional group is proportional to the average electronegativities of the substituents bonded to it. For this series of molecules, the calculated frequencies reflect the systematic substitution of hydroxyl for alkoxy species during the hydrolysis process. For example, the hydrolysis of TBP to DBP:

$$[(BuO)_3]PO + H_2O \longrightarrow BuOH + [(BuO)_2]P(O)OH$$

TBP and nitric acid are known to form adducts in the organic phase of the PUREX process. Nitric acid forms a hydrogen bond to the phosphoryl group of TBP, thus weakening the P=O bond, which is reflected in a spectral shift. The vibrational frequency is shifted from 1301 (TBP) to 1263 (TBP-HNO_3) to 1228 cm^{-1} (TBP-$2HNO_3$). There is insufficient space around the P=O bond to accommodate two nitric acid molecules simultaneously coordinating to the phosphoryl oxygen. Consequently, the second molecule hydrogen bonds to the first nitric acid molecule, resulting in a different spectral shift.



Calculated vibrational spectra for TBP, TBP-HNO_3, and TBP-$2HNO_3$, showing the P=O stretching vibration region.

The work will be extended to compute the activation energies for breakdown of the TBP ligand and to consider more realistic chemical environments.

PN17090/2980

Unveiling the Dynamic Microbial Biofilm and Plant Root Interface Under Extreme Conditions

Xiao-Ying Yu

Biofilm formation on plant roots is known to be associated with biological control and pathogenic response; however, how plants regulate this association is not well known. We propose to enable unique PNNL multimodal chemical imaging capabilities to investigate the dynamic biofilm-plant root interface under environmental conditions to obtain measurements ranging from molecules to cells and associated communities.

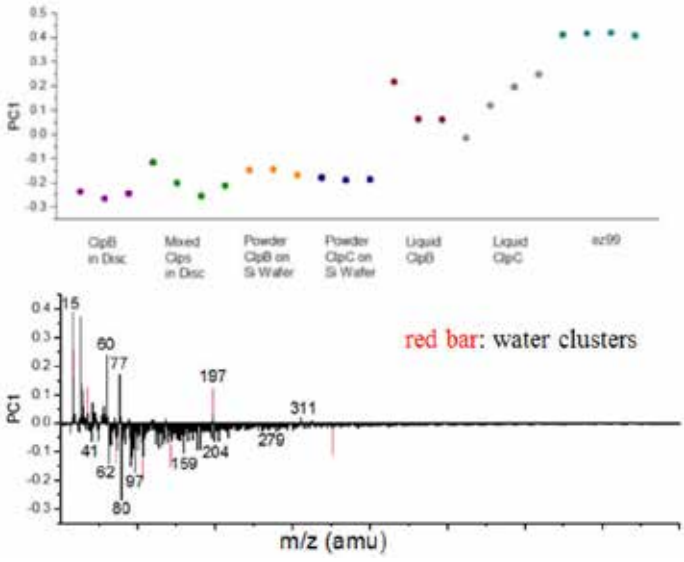
Plant growth promoting bacteria (PGPB) are a prerequisite for biofilm formation and adhesion to plant roots. Single cells play a complex functional role in bacterial biofilm surface components and extracellular factors in plant-bacterial association. Our knowledge of the identification and functions of extracellular proteins, DNA (eDNA), and lipids in the biofilm matrix of plant-associated bacteria remains limited. Cell-cell communication in bacteria-plant systems involves bacterial quorum sensing and the synthesis, release, and recognition of diffusible molecules such as indole. Production, perception, cellular sensitivity, and response, as well as stability of signaling molecule(s), are key criteria to define how these molecules can be monitored dynamically.

This research aims to provide *in situ* dynamic multimodal chemical imaging of the interface between plant roots and microbial biofilms enabled by a versatile microfluidic reactor, namely System for Analysis at the Liquid Vacuum Interface, which has resulted in a number of publications over the last several years. The following three tasks are proposed: 1) study the collagen-like proteins (Clps) and bacteria cell attachment at a surface, 2) obtain molecular distributions of key metabolic species in single bacterial cells at the biofilm-root interface, and 3) determine localized flux and mass transfer in biofilm cell-cell communications at different simulated extreme conditions.

In FY 2017, we optimized the biofilm culture procedure, selected a gram positive (i.e., *Anthrobacter*) and a gram negative (i.e., *Pseudomonas* fluorescens) species that are commonly found in soil as PGPB models for culture, and investigated their biofilm formation and adhesion to model solid surfaces. We hypothesize that the gram positive and gram negative species would have different effects on plant seed growth.

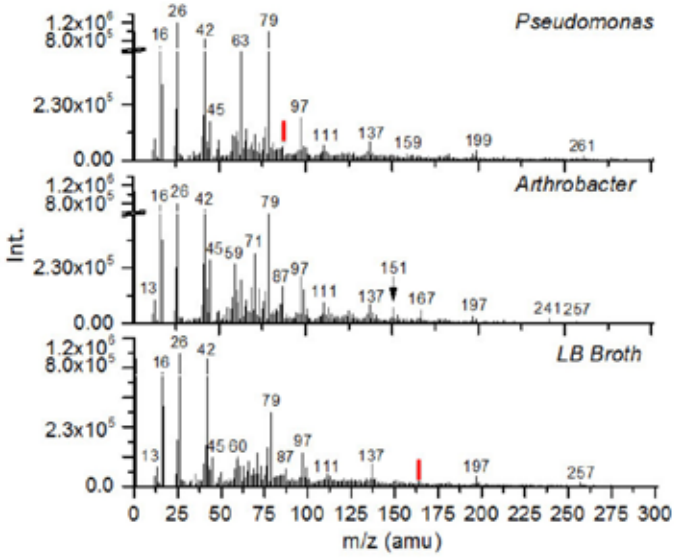
Our liquid secondary ion mass spectrometry (SIMS) spectral principal component analysis (PCA) showed a discrete water cluster environment among different Clps. The liquid SIMS features of cultured *Pseudomonas, Anthrobacter*, and the LB Broth medium solution were determined, providing the technical ground for the seeding/root and biofilm interaction experiments in the next stage.

In FY 2018, we aim to culture grass seeds in the microfluidic channel, followed by *in situ* chemical imaging analysis. *Pseudomonas putida*, a gram negative bacteria, may be used to compare its effect with *B. amyloliquefaciens*, a gram positive bacteria. We will focus on conducting experiments and obtaining molecular-level understanding underpinning the principles of confident redesign of microbes and plants for sustainable biofuel production, improved carbon storage, and/or contaminant bioremediation.



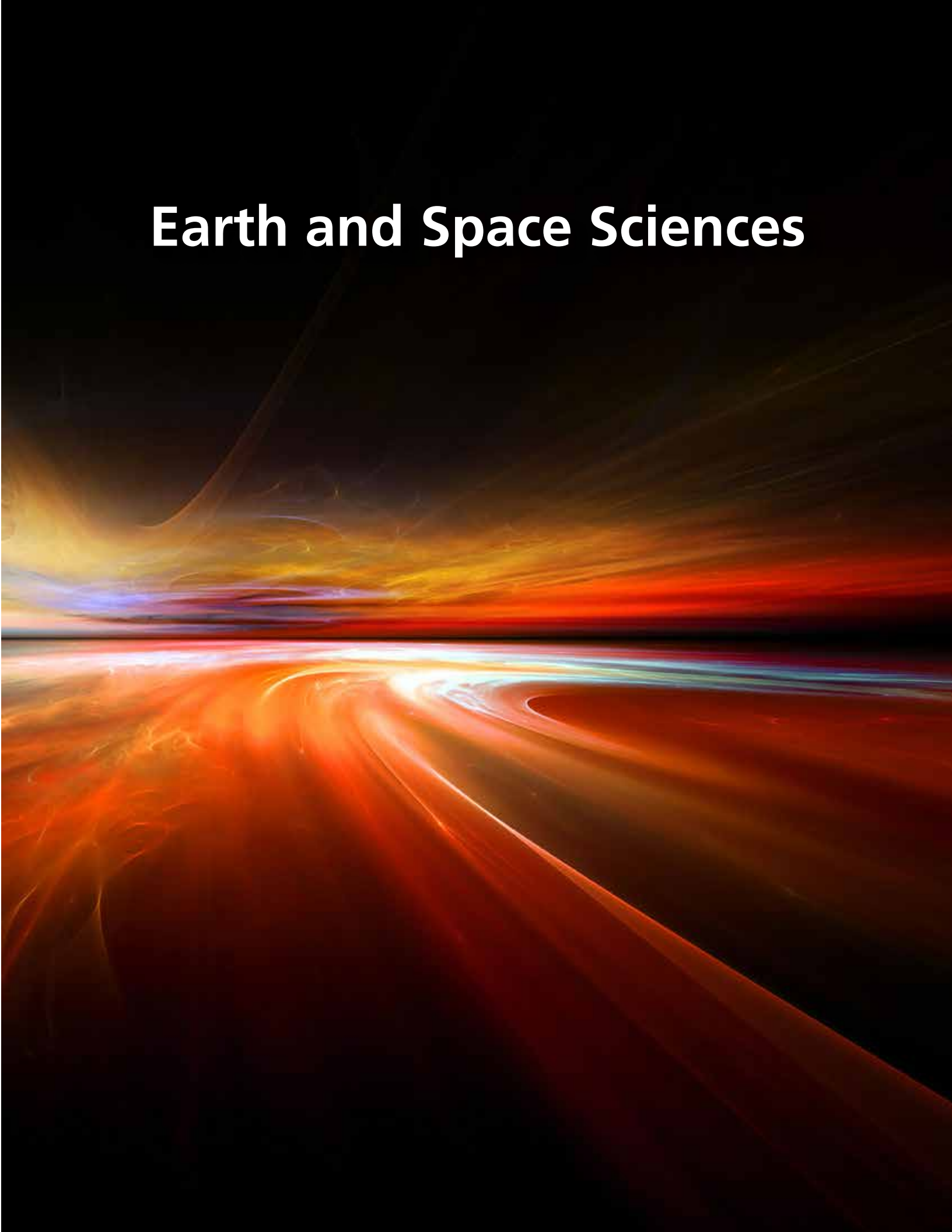
Negative SIMS spectral PCA PC1 score and loading plots showing water clusters affecting the distinction among ClpB dried on Si wafer, ClpB in liquid, ClpC in liquid, and *B. amyloliquefaciens* (ez99) in liquid. Red bars indicate the location of water clusters.

PN16071/2848



Negative liquid SIMS spectral comparison of *Pseudomonas, Anthrobacter*, and the LB Broth.

Earth and Space Sciences



Application of Nalu to Simulate Wind Plant Inflow

Larry K. Berg

The wind energy industry has been plagued by issues associated with premature turbine failure and inefficient wind plant design. Nalu is a new computational fluid dynamics (CFD) exascalable computer code developed for simulating the very fine-scale structure of winds and turbulence near wind turbines. Use of Nalu brings new opportunities to develop more reliable and durable wind plants, as well as new ways to optimize power production from the given wind resource.

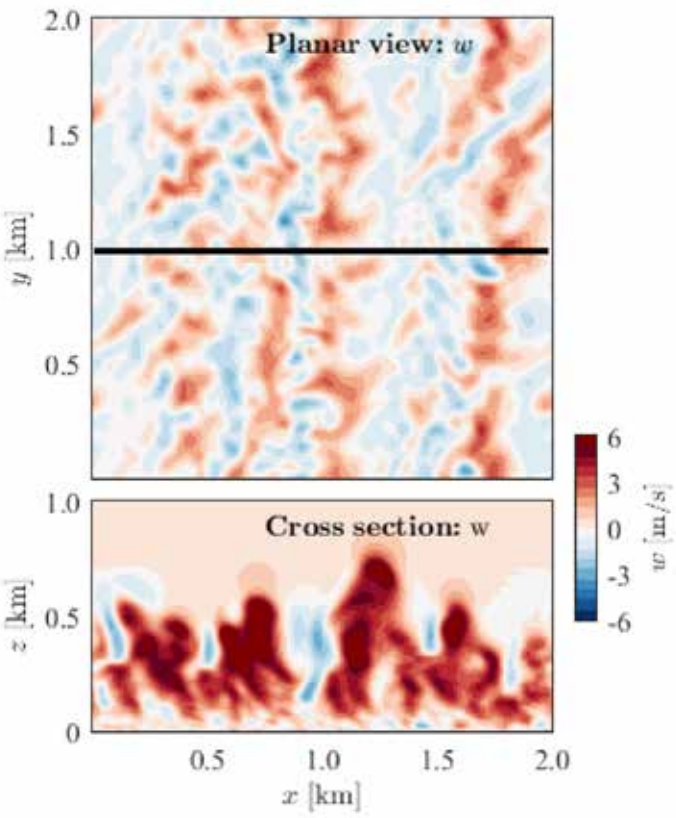
Accurate prediction of wind speed and turbulence flowing into operating wind plants is needed to quantify the wind power, optimize the performance of the wind farm, and improve estimates of the physical loads on the wind turbines. Existing paradigms and tools have focused on the use of a small set of idealized simulations or computationally expensive suites of code to solve the full-physics problem. In contrast to these models, Nalu is highly parallelizable and is optimized for use on DOE's High-Performance Computing platforms. It is a multi-physics solver (code) that uses an unstructured mesh to solve low Mach number fluid flow problems. The increased speed of Nalu will enable the use of larger computational domains and a much larger range of case studies than is possible with existing software tools.

The goal of this project is to gain experience with Nalu for wind energy applications to better position PNNL for future work with DOE's Wind Power Program. Specific tasks included downloading, installing, and using Nalu on the PNNL Institutional Computing (PIC) system to simulate atmospheric boundary-layer flows with varying stabilities such as daytime convective conditions.

Installation and use of an intricate model like Nalu is a complicated process requiring a number of different software libraries, and their dependencies and compiler settings vary between computer systems. At the start of the project, Nalu was successfully installed on the PIC system, and a PNNL-specific user guide was developed to enable others to more easily install and use the code. A number of standard test cases were run to ensure that Nalu was functioning properly.

The Nalu CFD code was run on the PIC system using a computational mesh (built using the Cubit software suite) 2x2x1 km^3 in size, with a grid spacing of 20 m in all directions. The model results were post-processed with Paraview software. Periodic lateral boundary conditions were applied in both horizontal directions. The wall adapting local eddy-viscosity model, available in the Nalu code, was used for modeling sub-grid scale turbulence. The simulation was driven by applying a geostrophic wind of $U_g=V_g=8$ ms^{-1} at the model top and zero wind speed at the surface. To mimic a daytime convective boundary layer, a surface heat flux of 300 Wm^{-2} was applied, which is representative of typical values observed over land.

A test run for a convective case showed that the Nalu code produces realistic eddies near the surface and convective thermals reaching to the top of the boundary layer



Planar view of vertical velocity (top) at 70 m above the surface and vertical cross-section of vertical velocity (bottom) through the middle of the domain, indicated by the black line in the top panel, for convective conditions simulated using Nalu CFD code.

PN17088/2978

(approximately 0.75 km above the surface, in this case), relatively intense updrafts (with up-draft velocities approaching 6 ms^{-1}), and a large area of more gentle downdrafts. In addition, the forcing with moderate geostrophic wind caused the flow structures to be stretched and advected across the domain.

Accurate simulation of coherent structures is essential to account for the energy transport in the lower part of the atmosphere, improving our understanding of boundary-layer turbulence, and for designing more durable wind plants. One primary cause of failure of current wind turbines is fatigue loading caused by unpredicted flow structures, such as these simulated thermals, that are not properly considered during the design applications of the wind turbines. High-resolution simulations, such as those produced in this effort, are also needed to better understand the behavior of turbine wakes and approaches that can be used to steer wakes away from near-by turbines and optimize wind plant performance.

Based on the knowledge and skill gained in this project, our team will work to develop new approaches that can be utilized by DOE and the wind energy community. Future efforts resulting from this project include the use of realistic boundary forcing to drive Nalu simulations, improved treatment in atmospheric surface layer, and inclusion of some of the physics modules (e.g., radiation) to produce accurate simulations of real-world boundary-layers.

Assessing Climate and Human-Exposure Impacts of Polycyclic Aromatic Hydrocarbons and Secondary Organic Aerosol Particles

ManishKumar B. Shrivastava

Despite chemical kinetics measurements showing fast chemical degradation of particle-bound polycyclic aromatic hydrocarbons (PAHs), these PAHs persist in the atmosphere and travel thousands of miles. Based on latest advances in measurements, we developed novel modeling algorithms showing that health-relevant PAHs are shielded in the atmosphere by naturally occurring highly viscous secondary organic aerosols (SOAs).

Despite decades of research, the long-range transport of PAHs is not well understood, mainly due to incomplete knowledge of gas-particle partitioning and chemical loss rates of PAHs.

Previous modeling studies have indicated that particle-bound PAHs like benzo (a) pyrene (BaP) need to undergo much slower heterogeneous loss to explain their observed atmospheric long-range transport. However, a mechanistic understanding of processes that cause this slow heterogeneous loss has been largely missing. Based on measurements that show that SOAs embed and shield PAHs from atmospheric chemical degradation, we developed new modeling paradigms in a global chemistry/atmospheric model, wherein a shell of highly viscous (glassy) SOA forms around BaP that is adsorbed on the surface of soot/black carbon core. This shell effectively shields BaP from heterogeneous oxidation by ozone, at relatively cold/dry conditions.

Using a global three-dimensional atmospheric model, we demonstrate how this new model treatment largely increases BaP concentrations, lifetimes, and their long-range transport, globally.

Simulated BaP concentrations vary widely, with concentrations exceeding the PAH exposure limit of 0.25 ng m^{-3} over major source regions in Asia, Western Europe, Russia, and Africa. At several locations globally, the new shielded treatment predicts an order of magnitude larger BaP concentrations compared to the default unshielded model.

In addition, global lifetime cancer risk due to human exposure to BaP increases by a factor of four in the new shielded treatment, compared to the default unshielded treatment.

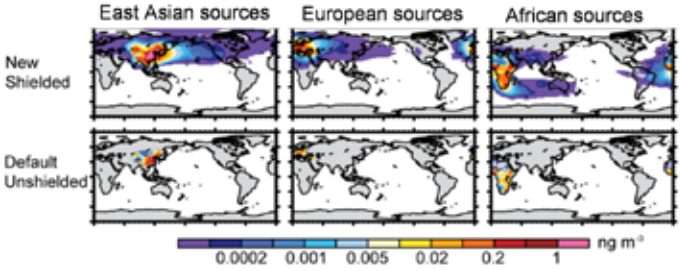
Due to its cross-cutting nature, the proposed research is expected to have multiple impacts affecting exposure-related human health, air quality, and national security, because of transport of air toxics to the United States from emission sources in other countries.

This work provided the first explanation of how PAHs persist in the atmosphere. As a part of this project, we also discovered, for the first time, synergistic interactions between PAHs and viscous organic aerosols.

In the next fiscal year, we will extend our initial modeling studies of toxic PAHs in the atmosphere to address two new and very important aspects of PAHs that were not considered in the first two years of this project.

In the first, conversion of PAHs to more toxic nitro-PAHs, we propose to include new chemistry treatments within a global chemical transport model, wherein PAHs reacts with NO_x species and are converted to nitro-PAHs known for their higher toxicity compared to parent PAHs.

In the second, catalytic effects of PAHs enhancing formation of large number of new particles and SOA, we will develop a new model parameterization that includes, for the first time, a complex chemistry mechanism, whereby PAHs catalyze the formation of multifunctional organics consisting of polymer-like chains and other non-volatile constituents within SOAs, discovered in our recent study. Formation of nanoparticles and their subsequent growth has multiple implications on air quality, lung diseases, the amount of solar radiation reaching the earth, and visibility.



This figure compares the default model treatment (bottom panels), where SOA does not shield BaP, to the new model treatment (top panels), which incorporates the effect of shielding of BaP by viscous organic aerosol coatings. In the new model, BaP, shielded by organic aerosols, travels thousands of miles from its source locations over oceans and continents. But in the default model, BaP is rapidly degraded by ozone and is lost within a couple of hours, so it cannot travel farther.

PN16001/2778

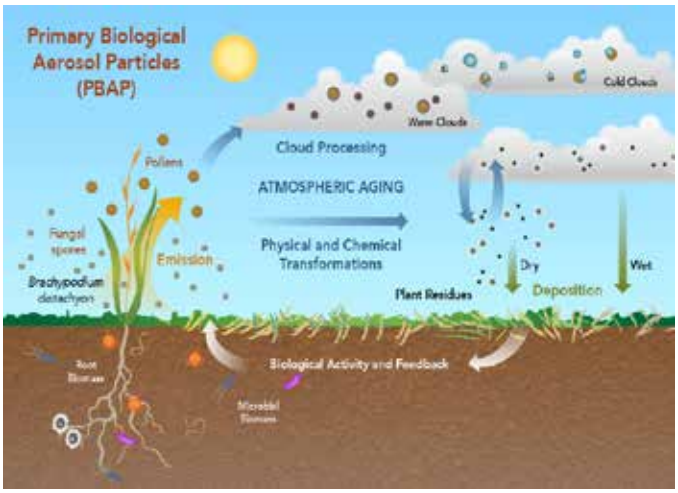
Atmospheric Role of Primary Biological Particles from the Plant-Soil System

Alexander Laskin

Our overarching goal is to develop a capability for measuring emissions of primary biological aerosol particles (PBAPs) in a* Brachypodium *integrated plant-atmosphere-soil system (iPASS), investigate physico-chemical properties of PBAPs controlling their atmospheric impact, and to use these data as the basis for a predictive numerical model describing soil-plant-atmosphere ecosystem behavior.

PBAPs play a vital role in the earth's system through various processes at the atmosphere-biosphere interface. PBAPs are emitted by living organisms directly to the atmosphere and consist of various cellular particles such as pollen, fungal spores, bacteria, viruses, fragments of plants and animals, and debris of dead organisms. After emission, PBAPs undergo various atmospheric aging processes (e.g., gas-particle reactions, photochemistry, coagulation, surface coating, etc.). PBAPs influence cloud microphysical processes and precipitation by serving as cloud condensation (CCN) and ice nuclei (IN), thereby affecting the hydrological cycle and Earth's climate. Removal of PBAPs from the atmosphere includes both dry and wet deposition, which further impacts terrestrial and aquatic ecosystems, triggering feedback on the PBAP reproduction.

Chemical compositions of PBAPs are highly variable and still poorly characterized, even at phenomenological level, because of inherent difficulties of field studies to distinguish between biological and other carbonaceous particles. As a result, contribution of PBAPs to total atmospheric aerosol composition is likely underestimated, and especially, the capability of certain PBAPs to act as effective CCN and IN is not adequately included in atmospheric and climate models at the level sufficient for predictive understanding the impact of PBAPs on climate system. This research project focuses on the determination and representation of actual emission rates and the atmospheric impact of pollen, fungal spores, bacteria and other biological particles from the *Brachypodium* plant-soil-microbe model system, chemical characterization of the emitted particles, and laboratory experiments probing their cloud formation properties.



Cycling and effect of PBAPs in the atmosphere and biosphere.

We studied the PBAPs characteristic of the *Brachypodium* plant-soil system and provided chemical characterization of prominent PBAP types and their variability at different environmental conditions (e.g., plant-soil ecosystem chamber maintained with different level of CO_2). Relative contributions of different PBAP types were determined as a function of plant lifecycle, temperature, and moisture levels. Spectro-microscopy analyses assessed particle external and internal composition variability. *In situ* measurements in a plant emissions chamber provided airborne concentrations and emissions of PBAPs.

We conducted quantitative analysis of hygroscopic properties and particle-cloud interactions of PBAPs, providing fundamental knowledge for cross-correlation between particle chemical composition and physicochemical propensities. Our experimental results will be used as an input for the mechanistic process-oriented modeling framework.

We will expand our studies to investigate effects of atmospheric aging on particle-cloud properties of PBAPs. The evolution of the chemical and physical properties of PBAPs estimated by the process models will serve as an input for evaluation of the impact of different types of PBAPs on the atmospheric environment and climate at regional and global scales in future projects.

PN17008/2898

Breaking the Curse of Dimensionality in Atmosphere Modeling: New Methods for Uncertainty Quantification and Parameter Estimation

Xiu Yang

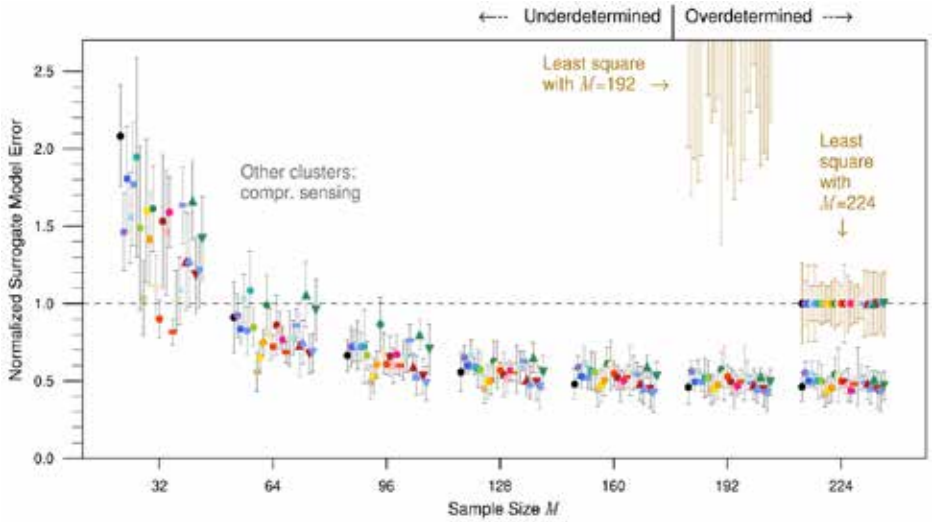
The goal of this project is to develop a new computational framework that uses cutting-edge signal processing techniques to efficiently quantify uncertainties associated with empirical parameters in numerical models designed for weather prediction and Earth system research.

Numerical models that simulate the global-scale dynamical motions of Earth's atmosphere are essential tools for weather forecasting and Earth system research. These models contain a large number of empirical parameters that are inherently uncertain, resulting in large uncertainties in their predictions. Standard uncertainty quantification (UQ) methods suffer from the so-called "curse of dimensionality," (i.e., the required number of simulations increases exponentially with the number of uncertain parameters). The numbers of simulations required by those methods are impractical to achieve due to the very high computational cost of each simulation. This has resulted in a major obstacle to effective application of UQ techniques to atmosphere modeling. The project aims at establishing a new, sparse sampling framework specifically designed for the very complex problem of atmosphere model development to allow for accurate characterization of parameteric uncertainties from an affordable number of simulations.

In FY 2017, we designed a strategy that uses compressive sensing to construct surrogate models describing the relationships between large-scale (global or regional) features of the atmosphere motion and the uncertain parameters in DOE's new global atmosphere model.

Validation results indicate that, compared to the commonly used method of least squares fitting, compressive sensing significantly reduces the error in the surrogate models. In those situations where the number of atmosphere simulations is significantly smaller than the number of degrees of freedom in the surrogate model, compressive sensing effectively plays the roles of both parameter screening and surrogate model construction. This implies that our algorithm is not only more computationally efficient but also more objective, since it requires far less prior knowledge from experts in the specific domain of research. Another highlight of our results is that the excellent performance of compressive sensing is consistently seen in a wide range of physical quantities that are important for atmospheric scientists. This suggests that our results are robust, and the methodology is likely to be applicable to other atmosphere models, as well.

In the next fiscal year, we will extend our method to characterize spatially distributed features of the atmosphere simulations. Since detailed regional features and their spatial correlations are key aspects in the evaluation of an atmosphere model's fidelity, this next step will be crucial for establishing an operationally useful UQ framework. We will also explore the feasibility of using multi-fidelity methods to further reduce the number of expensive simulations needed for UQ. The further reduction will be particularly useful for the high-resolution models that are currently under development at DOE and other leading institutions in the Earth system modeling community.



Compressive sensing is capable of providing more accurate and efficient characterizations of the parametric sensitivities in numerical models designed for global atmosphere simulations.

PN17009/2899

Ecosystem Transitions and Associated Greenhouse Gas Fluxes Following Salt-Water Intrusion from Relative Sea Level Rise

Heida L. Diefenderfer

The overarching goal of this project is to quantify, mechanistically understand, and predict shifts in CH_4, CO_2, and N_2O fluxes from an intact wetland community in response to salt-water intrusion. Understanding the effects of these large-scale disturbances, predicted to occur increasingly in coming decades, on greenhouse gas source and sink function is critical to improving process-rich models of climate change effects on coastal wetlands.

Coastal and fluvial ecosystems bordering the landward limits of salt water are seen as key depositional regions for terrestrial nitrogen (N) and carbon (C) bound for the ocean. These highly productive and globally distributed ecosystems are dominated by herbaceous, shrub, and/or forested vegetation. Long-term trends indicate that the zone of tidal freshwater will migrate substantially in the landward direction in coming decades. Soil chemistry is influenced by shifting hydrologic regimes at the intersection of river basin and coastal nearshore processes. The impacts of legacy material from disturbance-induced rapid succession on C and N cycling depends in large part on microbial community function. Understanding the controlling factors on these complex ecosystem processes is fundamental to DOE Biological and Environmental Research missions related to global elemental and water cycles. This project seeks to discern the sensitivity of soil microbiome molecular processes to altered salinity regimes.

Filling this knowledge gap is critical as it currently limits our ability to robustly predict changes in ecosystem function as a consequence of altered salinity regimes. The project aims to fill this knowledge gap through *in situ* characterization coupled with hypothesis-driven experiments on intact soil cores in the ambient coastal environment. Changes in porewater salinity will overwhelm the salt tolerance of plant and microbial communities. However, the time lags and magnitude of responses by microbial communities to salt-water intrusion, as measured by greenhouse gas flux function, currently remain unknown.

The research focuses on a tidally influenced, forested wetland near the mouth of the Columbia River in Washington State. Climate change is altering the hydrologic regime in this, and similar, tidal freshwater regions, such that historically river-dominated conditions are being increasingly affected by nearshore-ocean conditions, including salt-water intrusion. Aquatic-terrestrial transition zones like wetlands are experiencing the consequences of changes in the magnitude and dynamics of hydrologic flows and associated fluxes at different rates depending on relative sea level worldwide. Coastal wetlands are particularly susceptible to storm events, salt-water intrusion, and increasing water levels.

We hypothesized that sudden changes in salinity regimes will strongly select for particular microbial taxa, likely producing mortality events, and thereby affect community functions, including the emission of greenhouse gases. To address hypotheses like these, however, we needed to improve upon the standard experimental and monitoring systems that have been used for decades, static chambers that are sampled by hand with syringes and then analyzed at a gas chromatograph. The labor involved severely limits the frequency and timing of such sampling and has produced a publication record that appears to be skewed toward summer, daytime, calm weather, and C not N. Thus, we acquired a new, automated system suitable for sampling rapidly changing conditions. We intended this coupled cavity ring-down spectrometer and autochamber system to begin addressing a range of questions that can only be interrogated with high-frequency data. Through 2015, we adapted the autochamber system to sample in tidal areas, and obtained from the manufacturer changes to the software to repair bugs we identified and to enable calibration checks suitable for laboratory requirements.

In this project, we first exhaustively quality-control tested that system for laboratory use in a series of three laboratory experiments and a fourth quantitative experiment, coordinated with the manufacturer. Subsequently, they repaired defects that we identified in certain chambers. In summary, these tests were 1) an injection experiment (30 mL 9570 ppm CO_2, 14.3 ppm CH_4, 384 ppm N_2O into blank core) to assess accuracy and precision of measurement, determine consistency among chambers, and identify chamber leakage from malfunction; 2) continuous measurement, at 10-minute intervals, of fluxes of the three gases from homogenized sediment from the forested tidal freshwater wetland, for 2 weeks, to assess variability among chambers and occurrence of malfunctions; 3) a gap experiment, spanning the

PN16076/2853

range of mechanical malfunctions observed (3 mm, 1.5 mm, 1.0 mm, and 0.5 mm gaps), in order to analyze impacts on data; and 4) development of R-programs to identify malfunction events and to estimate the most precise slopes and, thus, fluxes from the data generated by the manufacturer's software. This satisfied us that we had a full understanding of the system's operations and tendencies to fail, as well as an ability to quantitatively detect any data that must be flagged in flux experiments, because such a system cannot be continuously observed.

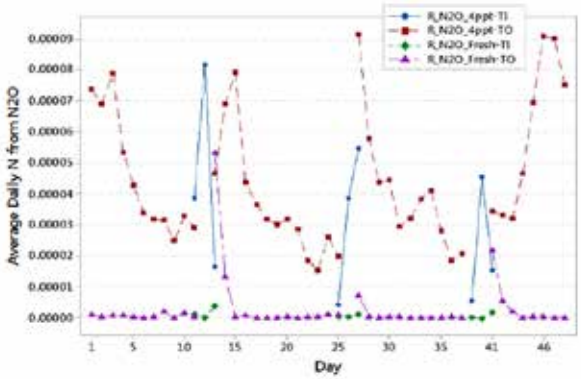
Secondly, we conducted a laboratory greenhouse gas flux experiment in the following three 4-month phases, each including six large-diameter cores from wetland soils: 1) compared freshwater and 4-ppt salinity conditions on cores under saturated conditions; 2) compared saturated and tidal (periodic saturation) conditions on cores at 4-ppt salinity; 3) compared freshwater and 4-ppt salinity conditions on cores under tidal condition.

These experiments were made possible by utilizing the seawater plumbing at the Marine Sciences Laboratory (MSL) and freshwater carried back from the wetland to mix to oligohaline-range diluted seawater similar to expected model-predicted effects of climate change in tidal freshwater regions. Moreover, we conducted the experiments outside, at sea level, a few meters from the Pacific Ocean, so changes in pressure, temperature, atmospheric CO_2, and other environmental factors are not simulated but real; monitoring data from the MSL weather station and other local sources, such as sea level buoys, will be used to analyze covariates with all data.

We also initiated characterization of *in situ* conditions at the study site and a core-scale reactive transport model that incorporates the saturation rates observed in laboratory experiments and analytical results on field-collected soils for total carbon, total organic carbon, total inorganic carbon, bulk density, particle size distribution, and total solids. During the core-incubation study, replicate cores were also sacrificed for porewater analyses (i.e., pH, specific conductance, carbonate, chloride, sulfate, nitrate, total calcium, iron, manganese, potassium, and sodium) to support parameterization of this model. Analysis of data from piezometers deployed at the site to develop a flow field for a model covering a larger areal extent indicates that they became impaired by sediment.

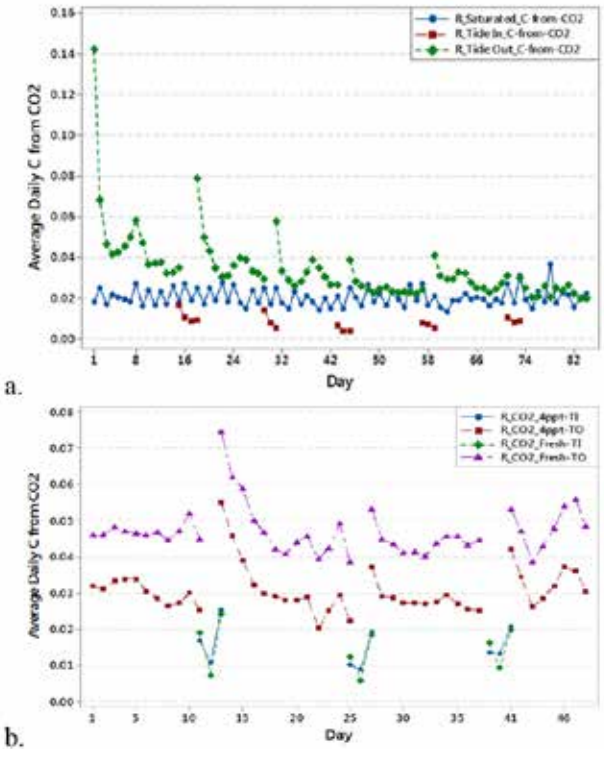
In Phase III, the contribution of N from N_2O was suppressed in the fresh treatment and spiked as the tide was going out and coming in; the spikes were also seen in Phase II. The 4-ppt condition suppresses C as CH_4 compared to freshwater under saturated conditions (Phase I). Under a 4-ppt condition, tidal inundation suppresses C flux as CH_4 compared to saturation (Phase II). Under tidal inundation, both Phases II and III show that both negative and positive fluxes of C as CH_4 occur.

The 4-ppt Phase II C contribution from CO_2 was suppressed by the saturated treatment compared to the tidal treatment, and tide in was suppressed compared to tide out. Phase III confirmed that both the 4-ppt treatment and freshwater treatment were suppressed when the tide was in, and Phase III further showed that when the tide was out, the 4-ppt treatment suppressed the evolution of C compared to freshwater.



N evolved as N_2O in Phase III experiment.

Phase III results indicate that under a sea level rise scenario, salt-water influx at a 4-ppt level would suppress C evolved as CO_2. With salt water, saturation, compared to tidal, also suppresses C evolved as CO_2 but not as CH_4 (Phase II). Phase I showed that, under saturated condition, 4-ppt salt suppressed C as CO_2 and CH_4 compared to freshwater.



C evolved as CO_2 in Phase II (a) and III (b).

Evaluation of Computational Approaches for Delineating Boundaries of Aquifer Exemptions

Inci Demirkanli

This project developed a scientific, data-driven methodology for delineating aquifer exemption (AE) boundaries that is consistent with the Environmental Protection Agency (EPA) regulatory standards and protective of groundwater resources and public health. Resultant delineations encompass suggested portions of drinking water aquifers to be exempted from protection under the Safe Drinking Water Act (SDWA) for operations such as oil or gas development, waste injection, or subsurface mining. The project highlights the importance of science-based decision-making in protection of common resources (particularly in resource-limited regions) that, otherwise, could have a direct impact on the environment and human health.

Groundwater resources are vital inputs to a range of often mutually exclusive uses. A federal policy of exempting some aquifers from drinking water protection rules (under the SDWA) has served as the regulatory mechanism for balancing the competing interests of subsurface activities, such as enhanced oil and gas recovery, subsurface mining, and waste disposal, with the need to protect groundwater resources. This regulatory approach for exemption determinations was structured by a broad definition of underground sources of drinking water (defined by a limit of 10,000 mg/L total dissolved solid [TDS] and of sufficient quantity) coupled with a discretionary exemption mechanism to equitably regulate underground injection on a case-by-case basis. While this approach was intended to provide some flexibility, it also resulted in inconsistent implementation efforts, leading to controversial exemption decisions and concerns about inventory, process, and protection of surrounding groundwater resources. Historically, applications of AE typically rely on a fixed radius approach based on simple structural assumptions and interpretations. In an environment where the effective management of water resources is increasingly important due to competing needs for energy production, waste management, and water supply, plus the compounding effects of increasing droughts, it is important that decisions, such as exempting aquifers, are informed by scientific analyses, which also provide the necessary transparency in decision-making to support the development of much needed energy projects.

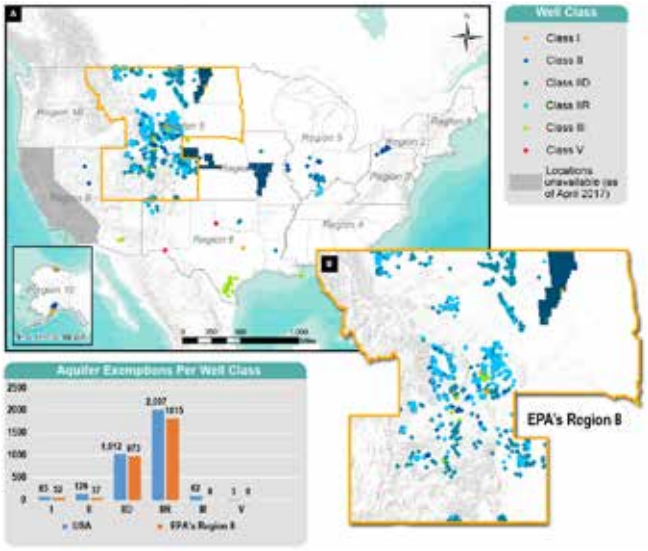
The purpose of this research was to develop a framework that could be used by decision-makers to support planning and/or overseeing of such operations and, simultaneously, ensuring the protection of groundwater resources. Resultant methodologies could guide the decision-makers and/or operators on important subsurface processes and the extent of the area of impact due to a potential subsurface activity. In addition, they have the potential to become the basis of an easily accessible and user-friendly analysis tool.

First, we performed a thorough analysis of existing AE documentation, from site-specific well logs to national-level summaries. We divided data summaries into typical operational characteristics and geologic characteristics to develop representative conceptual models. Determined by the frequency of AEs granted, our conceptual models were based on the disposal operations for oil and gas produced water (Class IID wells under EPA's Underground Injection Control Program) in EPA's Region 8. Next, we incrementally added system complexities (e.g., heterogeneous subsurface property distributions) to our conceptual models. The purpose of this approach was to identify important system characteristics that would be significant in determining the area of impact for the injection activities. Each variation of a conceptual model also required identification of appropriate computational approaches relying on a comparison between expensive numerical models and simpler, analytical solutions, while accounting for appropriate physical processes and significant operational and hydrogeological variations. Overall, this methodology yielded identification of sensitive parameters and processes to these disposal systems through a series of models with incrementally increasing complexities.

The conceptual models were used to set up an ensemble of simulation scenarios, which varied subsurface formation properties (e.g., confinements, thicknesses, depths), hydraulic flow parameters (e.g., porosity, permeability), *in situ* fluid properties (e.g., salinity, regional gradients), and injectate properties (TDS). Numerical simulations were run using the PNNL: STOMP-WS simulator for water-salt systems. Concurrently, a set of analytical solutions was used to evaluate the

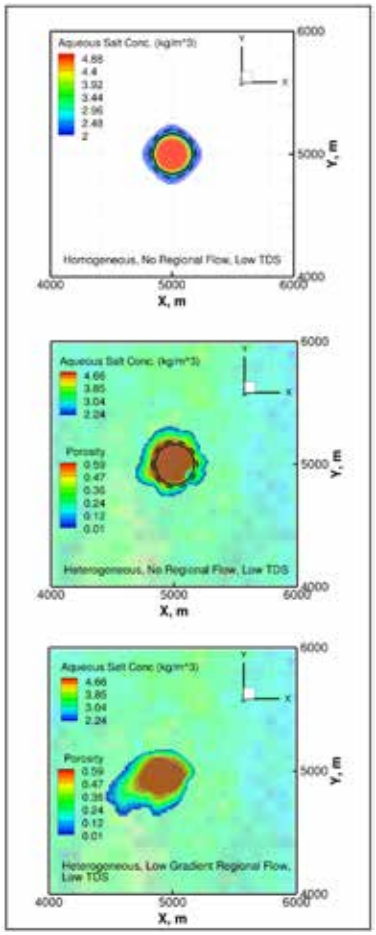
PN17018/2908

same ensemble of scenarios. Analytical solutions varied from simple volumetric approaches to approaches that considered diffusion, regional flow, and density differences between the *in situ* fluid and injectate. Results from numerical and analytical solutions were compared based on orientation, shape, and area of the injectate plume.



Distribution of Aquifer Exemption A) across the United States, and B) in EPA's Region 8.

Results allow us to offer guidance on scenarios where specific analytical solutions may be applicable, as well as scenarios where only a numerical model can represent the subsurface behavior. In general, we found that the analytical solutions are capable of accounting for one additional complexity at a time. For instance, one method evaluated could account for regional gradients but not density driven flow, while another method was complementary, accounting for density driven flow but not regional gradients.



Comparison of predicted extent of injectate based on selected methods to support the delineation of exempted boundaries of the aquifer.

How Do Non-Linear Microbial Processes Lead to Linear Ecosystem Fluxes?

Kathe Todd-Brown

This project incorporates new measurements and understanding into soil carbon models, advancing predictive capabilities of biogeochemical models and bridging the gap between microbial processes and ecosystem observations.

Modeling soil carbon dynamics is critical to predicting how much carbon dioxide will be in the atmosphere to drive future climate change. Soil carbon models have remained essentially unchanged for decades. These models partition soil carbon into one to nine pools; carbon leaves these pools at a rate proportional to the amount of carbon in the pool (i.e., first-order linear decay)—some of the exiting carbon is diverted to other pools, and the rest leaves the system as carbon dioxide. This model type tends to statistically fit observed soil carbon emissions using laboratory incubations. However, despite extensive attempts to characterize the environmental sensitivity, parameters must be re-calibrated for each location, making it difficult to confidently generalize these models to global projections.

In the past 15 years, new microbial and chemical characterization methods have led to a growing interest in the biogeochemistry community in process-rich decomposition models. In these models, soil carbon must first pass through a microbial pool before leaving the system as carbon dioxide. In addition, other explicit processes, such as extracellular enzyme mediated breakdown of organics and mineral-organic interactions, are also considered. These models are appealing, not only because they can describe a more fundamental understanding of soil decomposition, but also for their potential to incorporate new 'omic and high-resolution chemical measurements. However, there is an issue of scale: carbon decomposition occurs at the pore-scale (less than 1 mm^3), and the fluxes of interest occur at the ecosystem-scale (more than 1 km^2). In addition, there are mathematical and computational challenges due to the non-linearity of the new models.

The objectives of this research are 1) to develop better soil carbon models in the short-term via model-data integration of realistic field data into traditional decomposition models and 2) contribute to long-term improvements by capturing process-explicit representations of our evolving understanding of soil decomposition in new models. This research has led to 11 peer-reviewed manuscripts over the past 2 years.

On the traditional modeling front, this research contributed to a *Journal of Mathematical Biology* manuscript with international co-authors from seven institutes. The implications of using a constant input assumption to analyze the non-constant inputs systems were explored in the context of the first-order linear differential equation models traditionally used in soil carbon simulations. In general, given the time scales of interest, this was found to be an acceptable substitution, though care must be taken to distinguish between how carbon dating of the soil versus respired carbon dioxide is incorporated into these models.

In FY 2017, this project provided statistical and modeling support for a study published in *Nature*, evaluating novel ways to examine temperature sensitivity of soil carbon stocks across multiple field-warmed manipulations. We are leading the follow-on research, which strives to place this information in a more traditional modeling context. Soil carbon changes are generally considered to be too noisy to use for temperature sensitivity calculations; previous studies have instead relied on changes in soil respiration over time. We have shown that it is possible to calculate a temperature sensitivity using soil carbon stocks directly, given a large enough sample size, and can, thus, constrain Earth system model results post-hoc to reflect the field data. A presentation at the American Geophysical Union in FY 2018 will contrast these two approaches.

On the process-explicit modeling front, we have developed and parameterized a new model from the literature: JAM (Just Another Microbial) model. This research highlights the need for scale-aware modeling and outlines a general approach to do so. Ongoing research will develop the model codes, parameter sensitivity analyses, and resolve basic spatial heterogeneity effects on emergent system behavior in FY 2018.

PN15094/2769

Non-Intrusive Electrical and Electromagnetic Methods to Detect the Soil Conditions Within and Beneath an Engineered Surface Barrier

Fred Zhang

The objectives of this study are to establish methodologies for using the electrical resistivity tomography (ERT) and the electromagnetic induction (EMI) to detect soil wetness within and below a surface barrier and contaminant plume below a barrier.

The ERT has been used widely to map three-dimensional (3D) soil electric conductivity. The drawbacks of ERT are that a large number of wired electrodes need to be installed in the field, and the number of electrodes is proportional to the aerial size of the engineered surface barrier and inversely proportional to the desired imaging resolution. ERT also has difficulties with detecting the region of and beyond a non-conductive layer, but the problem may be overcome by proper configuration of electrodes.

Unlike ERT, EMI does not need any permanent field installation. A well-configured EMI survey can potentially provide cost-effective monitoring data, especially for a geometry that is thin but large in area. Additionally, the EMI measurement does not need direct contact with the soil, so the instruments can be carried by a person or mounted on a vehicle or a drone. The drawback of EMI is that the theory for 3D inversion is more complex and will likely require substantially greater computing resources than ERT.

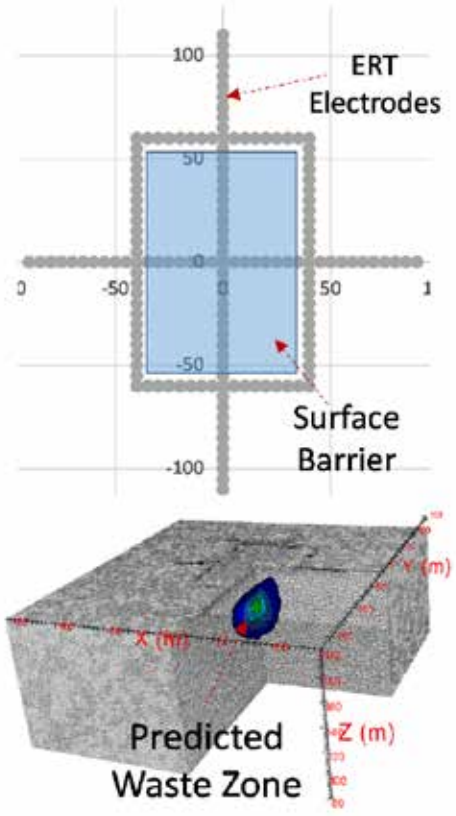
This project investigated the use of a non-intrusive and cost-effective geophysical tools and methods for monitoring the performance of a surface barrier to the hydrology and plume within and beneath the barrier.

The E4D simulator developed at PNNL for inverting the ERT survey data was used to evaluate the configuration of electrodes for optimized ERT monitoring of the performance of a surface barrier. The electrodes may be installed on the ground surface in a lateral configuration. In this case, the electrical current cannot flow through the non-conductive asphalt layer and, hence, has near zero sensitivity in the soil directly below the barrier. To improve the sensitivity, additional electrodes can be installed beneath the surface barrier before or during barrier construction. This improved configuration of electrodes can map the waste zone with higher electrical conductivity compared to the background. Another configuration is to install electrodes in a few vertical boreholes around the barrier boundary. This will also improve the sensitivity of ERT and can well map the waste zone.

These results indicate that the ERT method requires some of the electrodes be installed either beneath the surface barrier or in boreholes around the barrier for high-quality mapping the conditions below the surface barrier.

The RF module for COMSOL Multiphysics was used to evaluate the feasibility of EMI in monitoring subsurface conditions and performance of vadose zone remediation. The results show that both the electrical and magnetic fields are sensitive to the electrical conductivity of the soil. However, any highly conducting and permeable metallic objects at the surface or subsurface have a significant impact on EMI survey.

Either ERT or EMI can potentially be used in combination with other geophysical methods (e.g., gravity survey) for improved monitoring resolution.



The Role of Hurricanes in the Carbon and Oxygen Dynamics of the Coastal Zone and its Global Significance

Karthik Balaguru

This project aims to improve our understanding of the impact of hurricanes on hypoxia, a state of relatively low oxygen, in shallow continental shelf regions, such as the Gulf of Mexico. The results from this study will enhance our knowledge of the global ocean's carbon and oxygen inventory and help improve their representation in climate models.

The current scientific consensus is that, when hurricanes are near shallow continental shelf regions, such as the Gulf of Mexico, the mixing induced by their strong winds erodes the ocean stratification and alleviates hypoxia through ventilation of bottom water. However, the mixing can also cause resuspension of bottom sediments, with subsequent remineralization acting as a sink of dissolved oxygen (DO). This indirect effect on DO levels is typically not considered.

The goal of this project is twofold. First, we aim to demonstrate the process of resuspension of sediments and their subsequent remineralization. Second, we will perform a set of numerical experiments to quantify and parameterize this effect so that it can be applied to climate models.

During the first year of the project (FY 2016), we performed analysis using a suite of observations to demonstrate the process of sediment resuspension by hurricanes and its subsequent impact on water column DO. Initially, we used satellite measurements of suspended matter in the northern Gulf of Mexico to show that there is substantial resuspension of sediments from the bottom of the water column in the aftermath of hurricanes. We then used ship-cruise measurements of DO and the relationship between DO and stratification to infer that the process of resuspension consumes up to 20% of the increase in DO caused by vertical mixing.

As a next step, our goal is to quantify this effect. More specifically, we would like to know how well we can predict changes in DO due to sediment resuspension as a function of the state of the ocean and the storm. To this end, we performed a suite of numerical simulations using the state-of-the-art Regional Ocean Modeling System (ROMS). ROMS was configured in a one-dimensional setting using periodic lateral boundary conditions and meteorological conditions typical of northern Gulf of Mexico during hurricane events.

The four predictors we are primarily interested in are surface wind speed of the hurricane, bathymetry or water depth, the initial stratification of the ocean, and initial concentration of sediments. We performed a set of idealized numerical sensitivity experiments by selectively varying each of those predictors and understanding their impact on DO when the ocean is subject to hurricane-force winds.

For instance, in one set of simulations, we fixed the storm state, ocean stratification, and initial sediment concentration but varied the water depth (25 m, 50 m, and 100 m). For water depth of 25 m, the hurricane-force winds are able to completely mix the water column and produce near-homogeneous DO throughout the water column. On the other hand, for the case with 100 m water depth, the wind forcing is able to reach close to the bottom and resuspend sediments that consume benthic DO. However, in the case of 50 m water depth, in addition to resuspension of sediments, additional DO is consumed by the remineralization of dead phytoplankton and other detritus that get mixed down to the bottom. Thus, the maximum bottom hypoxia in this set of simulations was found in the case of intermediate water depth.

We performed similar experiments by varying one predictor at a time and estimating the corresponding impact on DO. Finally, using values of these predictors from all the simulations, we were able to fit a multivariate regression model to predict DO. The linear model performs reasonably well and explained nearly 46% of the variance in DO changes. This relationship can be used to estimate DO consumption due to hurricane-induced vertical mixing and sediment resuspension in climate models that do not explicitly resolve these processes.

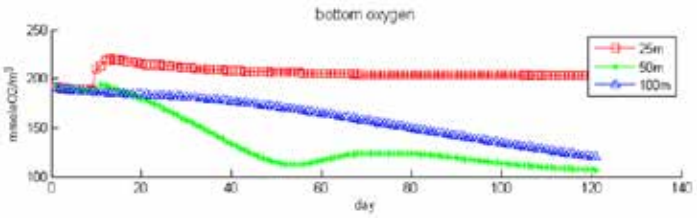


Illustration of the impact of hurricane-induced mixing on bottom DO for various water depths. The hurricane-force winds are applied for a day at day 10.

PN16066/2843

Understanding Polar Climate Sensitivity

Hansi A.K. Singh

The polar regions play a key role in moderating global climate, suggesting that climate change in the Arctic and Antarctic will have major consequences for the Earth climate system. This project aims to pinpoint the relative roles of atmosphere and ocean dynamic processes in determining the sensitivity of polar climates to anthropogenic forcings.

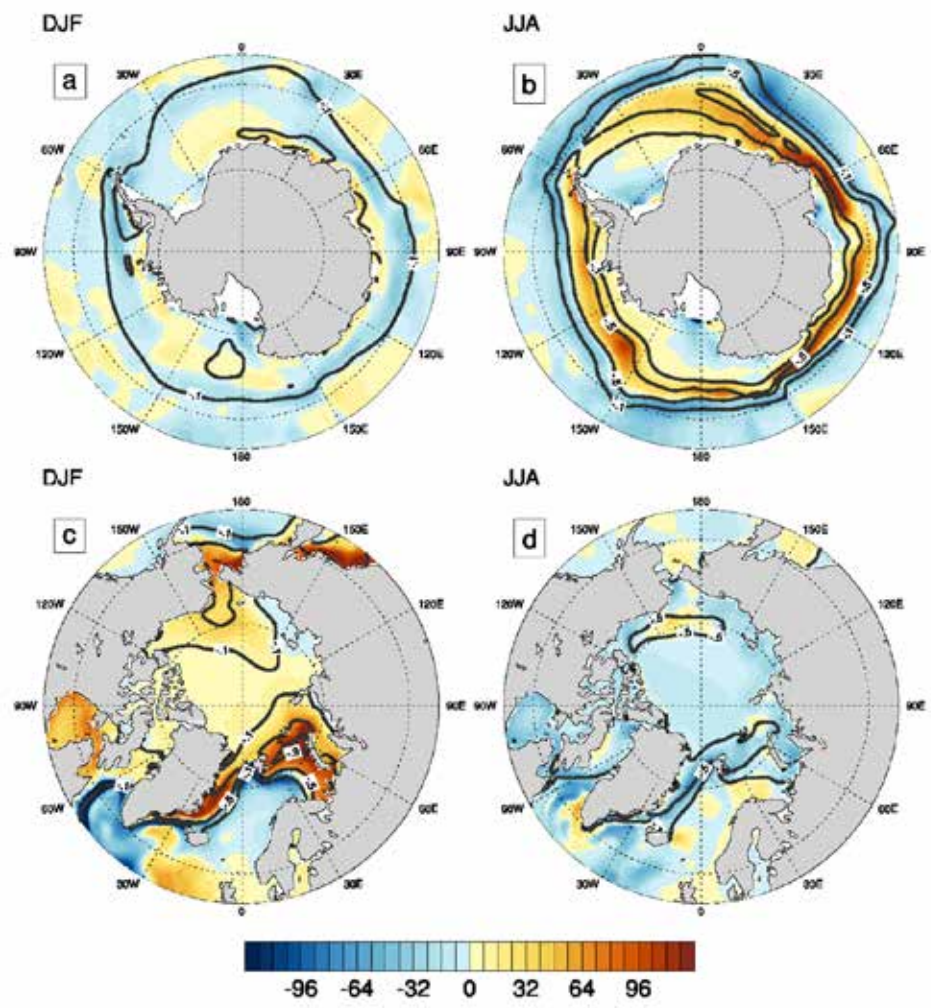
Previous studies have identified many local factors that enhance polar climate sensitivity, such as ice-albedo feedback processes, breakdown of winter temperature inversions, increased atmospheric precipitable water, and positive cloud feedbacks. However, the impact of large-scale dynamics on polar climate sensitivity, particularly changes in ocean and atmosphere energy flux convergence, remains an open question. The aim of this project is to isolate the role of ocean and atmospheric dynamics on the perturbation response of polar climates to forcings. Such an approach will highlight how differences in the general circulation in the Arctic and Antarctic, and the forced response of these circulations, affect the local climate sensitivity of each region.

In this project, we aim to use a range of varying-complexity models to study the many dynamic mechanisms that mediate polar climate sensitivity: single-column radiative transfer models, dry dynamical cores, moist dynamical cores, atmosphere-only models with gray radiation, atmosphere-only models with full radiation, atmosphere models coupled to a slab ocean without continents (aquaplanet) and with continents, and atmosphere-ocean models (fully coupled) with different configurations of polar orography. While polar climate sensitivity has been studied with the use of box models and fully coupled global climate models, there is very little study in the literature using models that fall between these in the complexity spectrum. Study with a range of models can be used to clarify a range of outstanding questions pertaining to polar climate sensitivity, including the impact of the mean state ocean and atmosphere energy transports (likely set by ocean, land, and ice sheet configurations) on the perturbation response; the role of changes in atmospheric moisture and dry static energy transport from the lower latitudes; the effect of robust poleward shifts in the general circulation; and the role of ocean dynamics (including heat uptake and heat flux convergence).

In FY 2017, we published two manuscripts on polar climate sensitivity. In the first, we showed inter-hemispheric parity between the high-latitude hydrologic cycle response to CO_2-doubling in the Arctic and the Antarctic: in both hemispheres, precipitation becomes more locally sourced (i.e., from *in situ* evaporation) in winter but more remotely sourced in summer. In the second, we used the Community Earth System Model (CESM) to link enhanced polar evaporation in winter to an enhanced, poleward-shifted pattern of ocean heat flux convergence, which provides energy for turbulent energy fluxes. Increased ocean heat flux convergence also warms the high latitudes and erodes the sea ice edge; concurrent midlatitude cooling enhances the polar lapse rate feedback, highlighting the link between large-scale dynamics and local feedbacks.

In FY 2018, we will look more closely at the roles of oceanic and atmospheric dynamics on transient (rather than quasi-equilibrium) climate sensitivity in the high latitudes. We are currently in the process of analyzing a set of CESM slab ocean experiments in which we evaluate the high-latitude climatic impact of the evolving pattern of ocean heat flux convergence with CO_2-doubling. While ocean heat uptake is similar in both hemispheres, we find that the Arctic is more sensitive to ocean heat uptake than the Antarctic. We will also consider whether high-latitude climate sensitivity due to natural variability can be used to deduce forced climate sensitivity and will quantify the state dependence of local climate feedbacks in the high latitudes. Finally, we will evaluate whether sea ice can be used as a predictor of the winter season atmospheric state and the extent to which polar climate change can be explained by sea ice loss. The dependence of the atmospheric state on sea ice will be evaluated using the single-column version of CESM.

PN16113/2890



The change in ocean heat flux convergence in the oceanic mixed layer due to CO_2-doubling in the CESM (colors; in W/m^2) and the change in sea ice fraction (contours at -0.1, -0.5, and -0.9). The poleward shift in the winter season ocean heat flux convergence is evident in the blue-to-red transition in December-January-February (DJF) in the Arctic and June-July-August (JJA) in the Antarctic.

PN16113/2890

Understanding the Role of Coastal Wetlands in Carbon Cycling – An Integrated Modeling-Observation Approach to Improve Regional Earth System Modeling

Zhaoqing Yang

Carbon fluxes among land-coastal and wetland-ocean interfaces are substantial components of the global carbon budget, yet remain poorly understood and quantified, largely due to the highly heterogeneous and dynamic nature of these interfaces. This project aims to quantify carbon fluxes and transport across the interfaces of wetland-estuary, air-sea and water-sediment, and to better understand the role of wetlands in coastal carbon cycling using an integrated modeling-observation approach.

Carbon movement from terrestrial systems to the oceans, through coastal wetlands, is a critical pathway of the global carbon cycle. There is growing research showing that these ecosystems may be relatively large carbon sinks globally. Coastal wetlands, in particular, have been persistent atmospheric CO_2 sinks over recent millennia but are significantly understudied and pose the greatest uncertainty and variability with regard to carbon cycling. Coastal wetlands are absent in Earth system models (ESMs) at global and regional scales. Therefore, there is an urgent need to develop a coupled wetland-ocean model that simulates the biogeochemical processes in the coastal wetland carbon pool, as well as fate and transport processes at wetland-estuary, wetland-atmosphere, and water-sediment interfaces.

The proposed study establishes a new capability of integrated tools with *in situ* measurements and coupled ocean-wetland models for simulating the fluxes and transport processes of carbon cycling in the coastal zone. The wetland model simulates carbon transport and transformation at the interface of terrestrial and ocean ecosystems and, thus, fills the current gap between land and ocean components of ESMs. The model addresses key physical transport and biogeochemical processes, as well as the interactions and feedbacks between them.

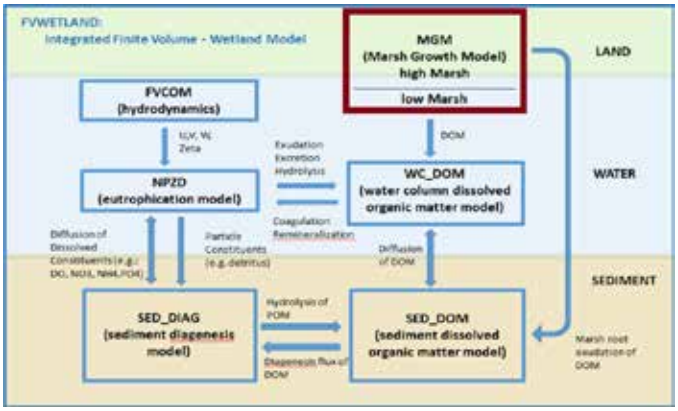
In situ measurements of key parameters were conducted in the water column at the dock of the Marine Sciences Laboratory in Sequim Bay and in a channel connecting the bay to a tidal salt marsh lagoon. High-frequency data for temperature (T), salinity (S), dissolved oxygen (DO), colored dissolved organic matter, and dissolved organic carbon (DOC), total suspended solids, and nitrate (NO_3) were periodically collected from spring to summer. Greenhouse gas emissions, including CO_2, CH_4, N_2O, and NH_3 in the marshland were simultaneously collected as the water column sampling. Collected data showed strong tidal signals, indicated the carbon flux and cycling in coastal wetlands is influenced by the physical processes. Data collected are used to parameterize the coastal wetland model and support model validation in Sequim Bay.

A finite-volume coastal wetland model (FVWETLAND), has been developed within the framework of Finite Volume Community Ocean Model, which includes several sub-models. A water quality model is included to simulate nutrient, phytoplankton, zooplankton, and detritus in the water column. A marsh and submersed aquatic vegetation growth model (MGM) is created to simulate the growth and mortality of marsh or submerged aquatic vegetation (SAV) plants dynamically based on available nutrient and light, as well as mortality characteristics. The state variables for the MGM include the plant leaf, stem, and root biomass per unit area of seabed. Light attenuation is affected by the plant height, in addition to background attenuation. A water column dissolved organic model (WC_DOM) and a sediment dissolved organic model (SED_DOM) are created to simulate DOM due to mortality of phytoplankton, zooplankton, and marsh/SAV. A sediment diagenesis model (SED_DIAG) is created to simulate the geochemical reactions within sediments and the exchange with water column based on receiving organic matter settled from overlying water column and decay of root materials.

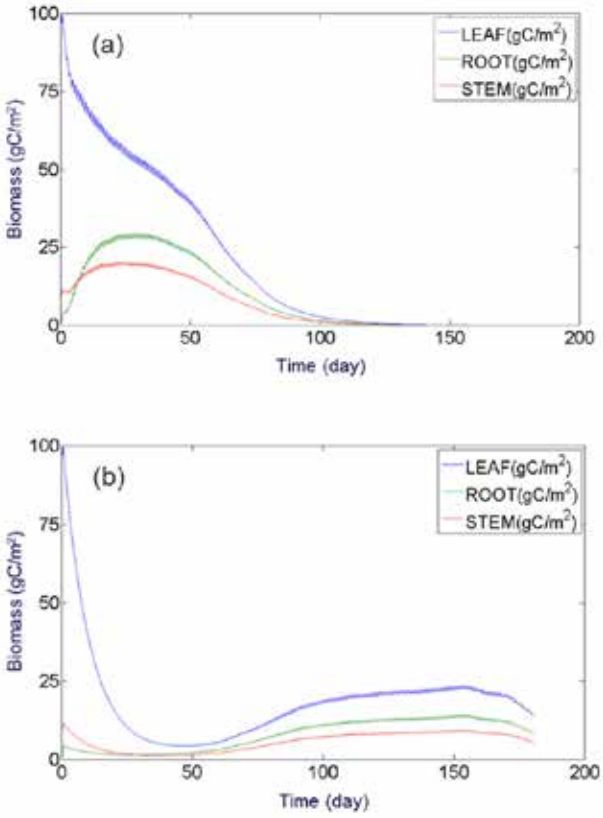
FVWETLAND was tested to evaluate the effect of marsh exudation on DOM flux from sediment bed into the water column in an idealized model domain, which mimics the spatial scale and physical-biogeochemical processes with mash

PN16027/2804

plants placed along the coastal line of the estuary. Preliminary model results demonstrated that the interaction between marsh growth and light attenuation, as well as marsh root mortality generated organic matter, are important processes that regulate the coastal carbon and nutrient cycles. Model results also indicated that, for freshwater marsh, salinity regime, together with light attenuation, can determine the differential survival pattern of marsh/SAV in coastal environment.



Schematic diagram of the coupled hydrodynamic-wetland model – FVWETLAND.



Simulated marsh biomass (leaf, stem, and root) response to different salinity regions with FVWETLAND: (a) high salinity region (> 25 ppt); (b) low salinity region (< 5 ppt).

Understanding the Stability of Organic Matter-Clay Systems in Presence of Aqueous Fe(II)

Ravi K. Kukkadapu

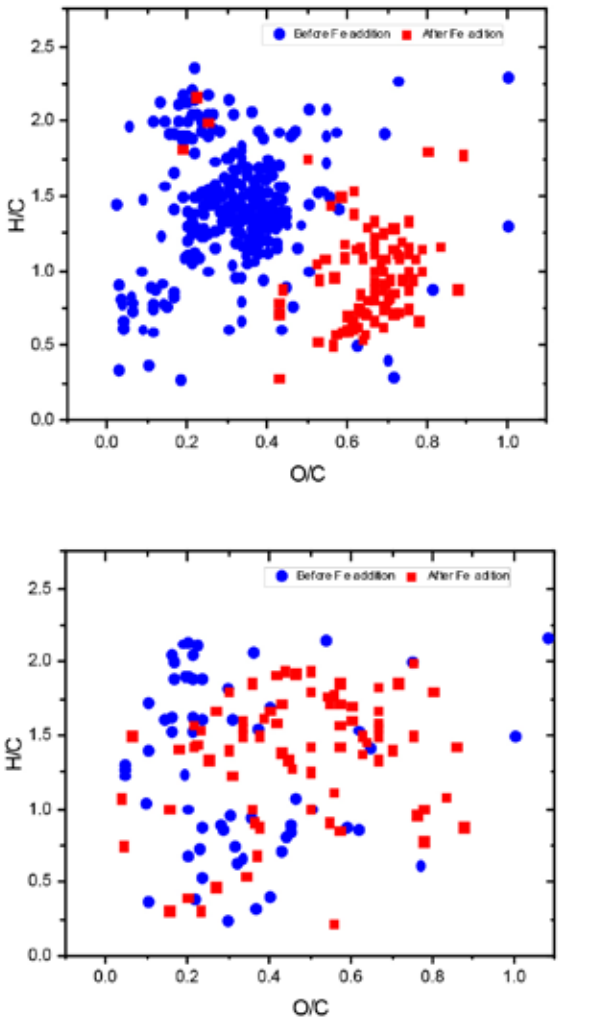
The fate and transport of C, as well as nutrients and metals in subsurface, is largely controlled by Fe metal and soil organic matter (SOM) transformations. An understanding of Fe and C coupled (bio)geochemical changes will strengthen the existing transport models to predict the movement of C in terrestrial subsurface systems.

A large fraction of the SOM is sequestered by Fe-containing clay minerals. The stability of such moieties under reducing conditions is most likely a function of both the type of clay, particularly its Fe content, and the nature of SOM, based on studies with different clay types (in absence of organic matter [OM]) and varying types of SOMs sorbed on a clay.

Clays with different characteristics exist in subsurface. In this study, two well-characterized reference clays from the Clay Mineral's Society (Fe-poor SWy-2 and Fe-rich NAu-2) that are coated with SOM (fulvic acid from the International Humic Substance Society) are reduced with aqueous Fe(II) in bicarbonate medium at pH 7. Fe-rich clay that is coated with SOM is also reduced with a Fe-reducing bacterium. These studies mimic interplay of Fe and C, as well as nutrient dynamics in subsurface.

Samples are analyzed for bulk Fe speciation by ^{57}Fe-Mössbauer spectroscopy, detailed secondary mineral identification by high-resolution transmission electron microscopy, and organic matter characteristics by FT-IC-MS mass spectrometry.

The clay-OM aqueous system revealed several discoveries. The first is that clay type dictates the soluble OM composition. Lignin-like compounds dominate Fe-rich/FA clay supernatant. We also found that lignin-like compounds are absent in Fe(II)-spiked Fe-rich/FA suspensions. This could be due to precipitation of Fe(II)-lignin-like compounds. In addition, high O/C compounds (tannin-like) are evident in the Fe(II)-Fe-rich/FA suspensions—likely due to release of these compounds from the reduced clay surfaces. Lastly, we found that Fe(II)-spike has little or no effect on OM composition for the Fe-poor clay system. This could be due to complexation of residual Fe by OM.



Van Krevelen diagram showing unique peaks in the supernatant before and after the addition of aqueous Fe(II). (Top: Fe-rich/NAu-2/FA system—blue circles are lignin-like compounds, and red circles are O/C rich, tannin-like. Bottom: Fe-poor/SWy2/FA system.)

Overall, the suite of experiments that comprised Fe-poor and Fe-rich clays coated with fulvic and humic acids, reduced either biotically or abiotically under varying pH and buffer conditions, has revealed that a clear understanding of the stability of clay-OM assemblages is critical to understanding the fate and transport of SOM.

PN16067/2844

Virtual Plant-Atmosphere-Soil-System (vPASS) 2.0: Mechanistic Process Components of a Virtual Plant Simulator

Yilin Fang

There is a critical need to improve the accuracy of plant process contributions to the carbon cycle in climate model projections. A systematic, mechanistic, and quantitative understanding of how plants invest and allocate carbon in the context of their immediate environment will be key to addressing this need. This research project will develop an integrated plant ecosystem modeling capability for plant performance as a function of genotype and environmental conditions.

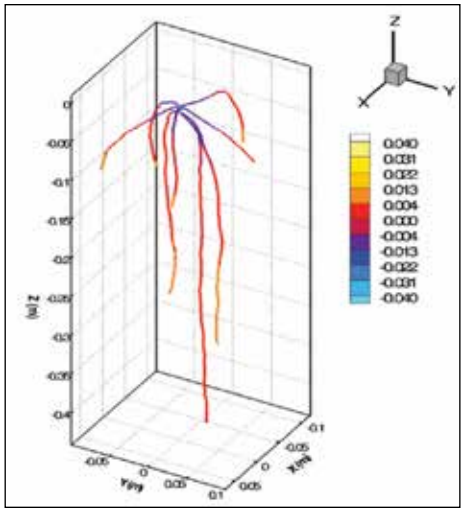
Brachypodium has been well studied for investigations of genomic potential for plant performance; however, no computational process models of *Brachypodium* structure and function have been published. This project will develop a novel prognostic Virtual Plant-Atmosphere-Soil System (vPASS) model that integrates a comprehensively detailed mechanistic single-plant model with microbial, atmospheric, and soil system processes in its immediate environment. This development will focus on *Brachypodium* sp. as a model for bioenergy grasses. There are three broad areas of process module development under this project: 1) incorporating models for root growth and function, rhizosphere interactions with bacteria and other organisms, litter decomposition, and soil respiration into established porous media flow and reactive transport models; 2) incorporating root/shoot transport, growth, photosynthesis, and carbon allocation process models into an integrated plant physiology model; and 3) incorporating transpiration, volatile organic compound emission, particulate deposition, and local atmospheric processes into a coupled plant/atmosphere model. Partnering with other research organizations, the developed model will be released as a community open source model.

Principle efforts have focused on planning and coordination, and the identification, acquisition, and testing of plant modeling software. The vPASS team has had ongoing discussions on data needs with the other focus areas involved in the *Brachypodium* experiments. A sound experimental design is key to the successful parameterization and testing of the vPASS simulation capability. To date, we have identified a single-plant growth and basic functionality model, BioCro. BioCro is adapted from a plant production model developed at University of Illinois at Urbana-Champaign. It is a detailed model that scales from leaf-level photosynthesis to ecosystem level carbon and water balance. From the model prediction, initial model test of the C3 canopy module of BioCro showed that 1) elevated carbon dioxide concentrations increase canopy assimilation, but have no effect on transpiration, and 2) assimilation decreases as angles of light incidence increase.

We generated root system architecture for *Brachypodium* using RootBox. RooTBox can easily create time dependent branched geometries of growing plant root systems. The resulting root systems can be coupled to arbitrary soil models such as nutrient and water uptake models.

We also developed a fully coupled root water and nutrient uptake model with eSTOMP, an open source, high performance software simulating variably saturated flow and reactive transport. Water flow within the root xylem and between the soil-root interface and root xylem is solved by discretizing the root system as a network of connected root nodes, which is the output from RootBox. Boundary conditions for roots are transpiration rates, which can be output from BioCro, and they are defined at the root collar. The coupled model can naturally simulate hydraulic redistribution by roots.

Ion uptake of root system is very selective. Some plants have been observed to accumulate larger amounts of K+ ions and to reject other ions such as Na+. We were able to simulate the accumulation of rejected ions in the rhizosphere due to the transpiration-driven uptake of water during daylight with our model and parameters from the literature.



Flow (mL/d) within the roots showing hydraulic redistribution at the end of dark period.

PN16103/2880

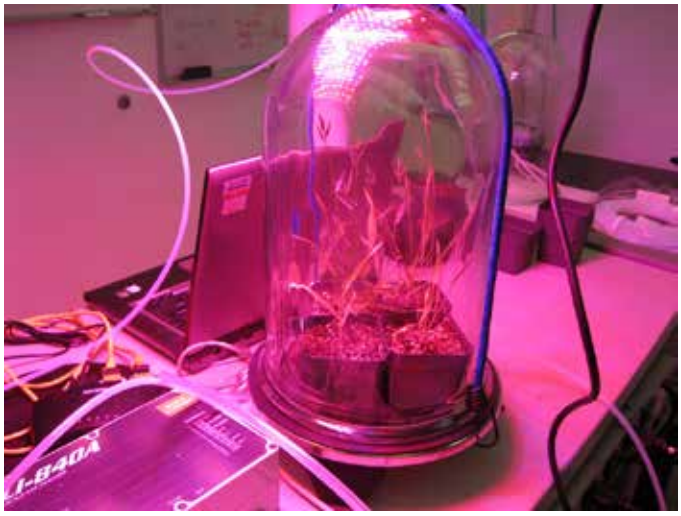
Volatile Organic Compound Emissions from the Plant, Soil, Microbial Ecosystem and Their Climate Implications

John E. Shilling

The goal of this study is to measure volatile organic compound (VOC) emissions of plant/soil/microbial ecosystem and how these emissions respond to stress events, including drought and mechanical injury. Ultimately, these studies will improve the understanding of plant emissions in a changing climate.

World-wide, ecosystems, including grasslands, face a range of challenges from global change, including direct effects of elevated atmospheric CO_2 levels, increasing temperatures, shifts in precipitation regime, and higher concentrations of ground level ozone. Little is known about how grassland emissions will respond to these climate-induced stresses. A small number of studies have shown that stress may alter emissions from trees, but there are almost no studies on grasses. This project aims to fill in knowledge gaps by focusing on *Brachypodium* grass, investigating: the relationship between plant genotype/phenotype, VOC emissions, the role of climate-induced stress on emissions, and the importance of microbial communities on these emissions. To accomplish these tasks, we plan to construct a chamber to capture, isolate, and measure VOC emissions from *Brachypodium* grass and then apply stress to the plants, measure the emissions, and compare them to emissions from the unstressed plants.

We have constructed a plant emission chamber and a system to regulate, control, and measure humidity and CO_2 levels in the chamber. Briefly, VOC-free air is mixed with pure CO_2 and moist air that is humidified by passage through a Permapure Nafion membrane saturated with water vapor. CO_2 and water vapor are measured continuously with a Li-COR Li-840 sensor. The temperature and humidity of the air exiting the jar are measured with a high-precision miniature humidity/temperature transmitter. A grow light hangs above the chamber to promote photosynthesis and, thus, VOC emissions. An Ionicon high sensitivity Proton Transfer Reaction Mass Spectrometer (PTR-MS), capable of detecting sub-ppbv VOC concentrations, sub-samples the chamber outflow.



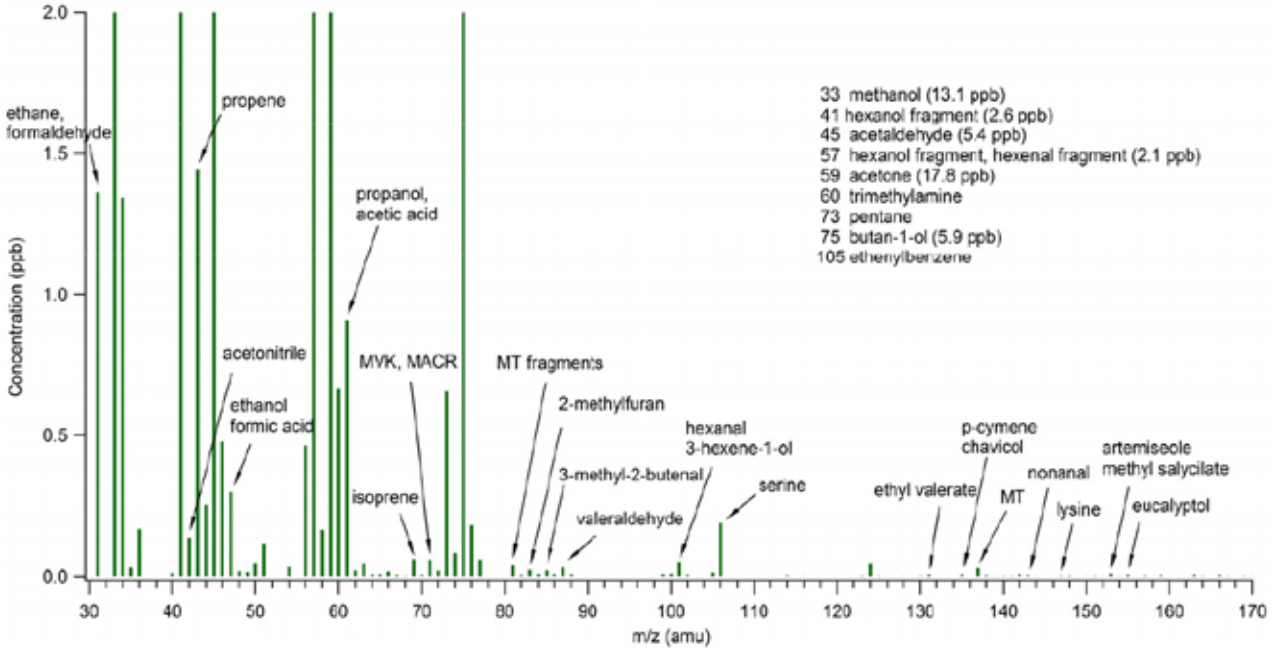
Chamber to measure VOCs emissions from *Brachypodium* plants.

We have tested the plant emission chamber to ensure that it is completely sealed from the background laboratory air and have begun making reproducible *Brachypodium*/soil/ecosystem emission measurements. Currently, we measure, quantify, and assign reproducible VOC emissions in real-time using the PTR-MS. Typically, acetone, methanol, acetaldehyde, and butanol dominate the emissions of unstressed plants. We also observe smaller amounts of ethane/formaldehyde, propene, propanol/acetic acid, ethanol/formic acid, serine, isoprene, monoterpenes (MT), and several other minor components. We have also begun to measure emissions as plants are stressed, initially focusing on mechanical injury due to cutting. The major components of the emissions from cut *Brachypodium* were methanol, acetaldehyde, acetone, propanol, and butanol.

We observed higher emissions of these compounds for more than 90 minutes compared to the values before clipping. For approximately 10 minutes, we observed that hexanal emission increase by a factor of 34. This is consistent with literature studies that have found substantially enhanced levels of these compounds in clipped pasture. We also assessed the *Brachypodium* plant's behavior at the sharp transition from light to dark and vice versa. The transition to dark

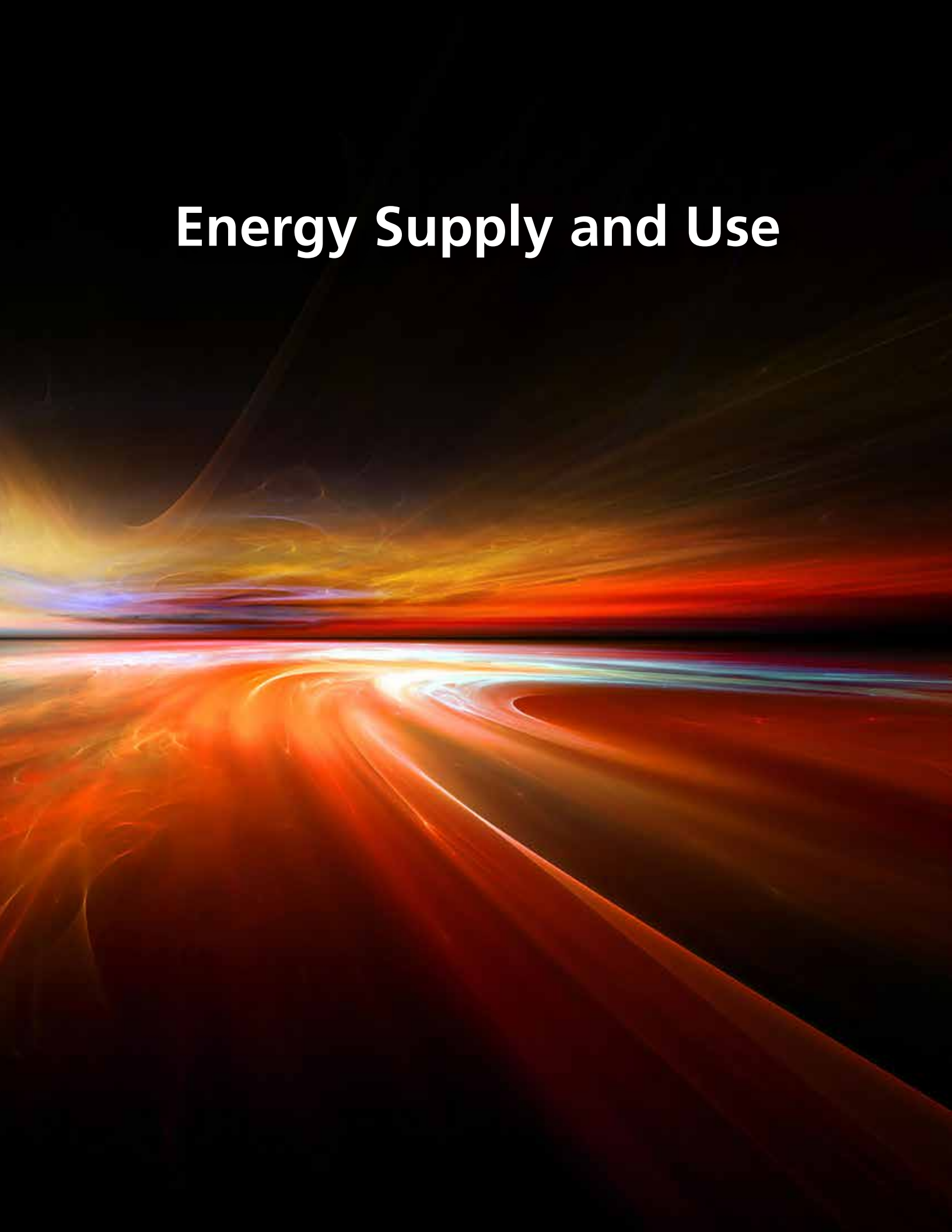
PN17005/2895

triggered bursts of isoprene, methanol, and acetone, while the transition to light resulted in similar bursts for methanol only. At the end of FY 2017, we plan to measure the emissions from *Brachypodium* that have been subjected to drought stress and compare them to a control group that has not.



PTR-MS spectrum of *Brachypodium* emissions.

Energy Supply and Use



Campus as a Laboratory

Paul Ehrlich

The ability to connect to and test on physical hardware enables control method developers to better validate the behavior of systems under control of the methods they develop and the models they use in simulations at scale.

This project continues the development of new physical test bed capabilities and enhancement of the existing Control of Complex Systems Initiative (CCSI) test bed, while ensuring that the resulting test bed is flexible, scalable, and expandable so that new, advanced control technologies can be experimentally tested. The project focuses on expanding the existing test bed infrastructure to include additional hardware, equipment, and buildings and includes testing and validation of the new test bed capabilities implemented. The project also includes documenting the components, structure, and operation of the test bed and providing support to experimenters in defining, configuring, and executing experiments.

During FY 2017, project work focused primarily in two areas: 1) integrating external data source connections into the test bed, and 2) re-integrating the relocated Building Diagnostics Laboratory into the CCSI Test Bed and supporting use of it for an integrated demonstration.

The external data sources integrated included data streams for wind and balancing authority load from Bonneville Power Administration, several data sources from the Midcontinent Independent System Operator, data from the New York Independent System Operator, data on system status of the California Independent System Operator, and price data from the Commonwealth Edison Company's (ComEd's) Residential Real Time Pricing Program. The data from these sources can be used as input to control experiments.

Re-integrating the laboratory focused primarily on establishing network connections and commissioning and enhancing the laboratory equipment and controls. This lab includes a complete commercial HVAC system (chiller, air handler, variable-air-volume terminals, and controls) that would normally serve a small office building or school. The lab was relocated from its previous location in FY 2017 as part of a PNNL move from some old facilities and enhanced with new controls, metering, and sensors, and the ability to false load a set of small test rooms conditioned by the HVAC system. This facility can readily be used to simulate a series of building sizes and system types. Testing in the facility is achieved with the use of a commercial control system that uses BACnet® communications. These controls are connected to the CCSI Test Bed via a VOLTTRON™ server and a bridge between VOLTTRON and the Framework for Network Co-Simulation (FNCS, pronounced Phoenix) co-simulation environment, enabling researchers to perform hardware-in-the-loop simulation experiments.

This flexible facility as part of the CCSI Test Bed began to support experimentation on new methods of control implemented on the HVAC system in the laboratory, the laboratory system in conjunction with operating buildings on the PNNL campus, and hardware-in-the-loop together with simulations that enable experiments on large-scale building-grid systems.

In FY 2018, the project will continue to expand the range of physical systems supported by the CCSI Test Bed, including light commercial HVAC, residential heating and cooling systems, Electric Vehicle Service Equipment, and battery storage.

Equipment in the Advanced Building Controls Test Facility.

PN17040/2930

Control Framework for Large-Scale Complex Systems

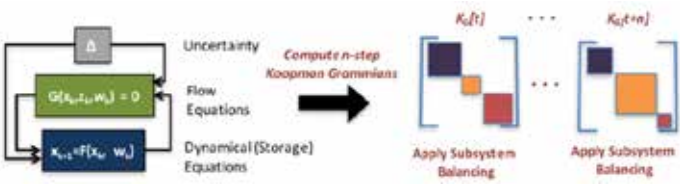
Enoch H. Yeung

This project is developing a framework for system analysis, decomposition, and engineering distributed control for infrastructure systems.

With rapid technological advancement in communication, computing, and control, dynamical systems become increasingly complex (e.g., large-scale, dynamic coupling, uncertainty), posing new challenges to the current control engineering paradigm.

This project develops analysis and synthesis methods to systematically architect and design control hierarchies for large-scale complex systems. The outcome of this investigation is a set of requirements that will inform control architecture definition and algorithm selection for a given system size and physical shape, as well as communication and computation architectures.

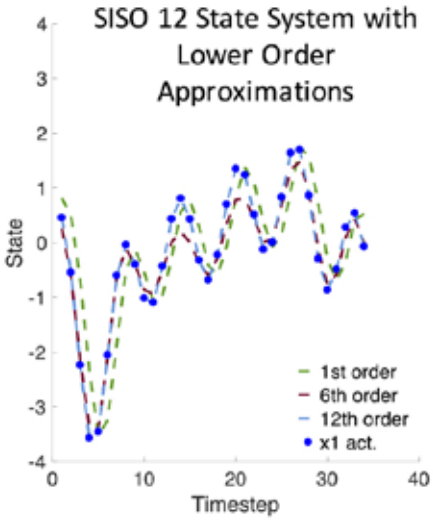
In FY 2017, we developed a new approach for nonlinear system decomposition, based on Koopman operator-theoretic methods for quantifying controllability and observability. We first extended the theory of Koopman operators to nonlinear systems with control inputs. We introduced Koopman grammians as a way to quantify controllability and observability for a nonlinear system.



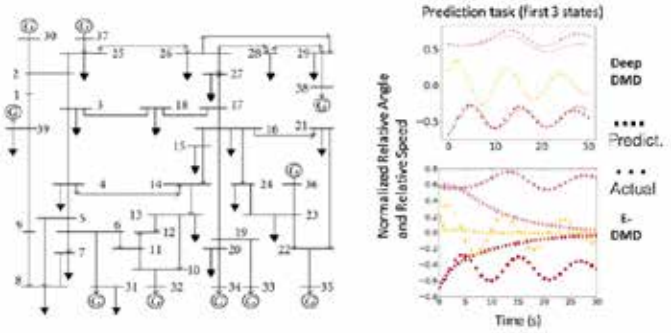
Schematic of model decomposition approach for infrastructure systems (e.g., multi-machine power transmission systems and water networks). The data (or models) of these systems are used to learn and compute Koopman grammians for the system, which are then subjected to a subsystem discovery process involving multi-way partitioning algorithms. The resulting subsystems are bounded-input and bounded output (or state) stable.

We then introduced the concept of a Koopman balanced realization, extending classical balanced truncation techniques from linear systems theory. Koopman controllability and observability can be used as criteria for model decomposition, to guarantee that the corresponding subsystems are internally stable, while robust to external disturbances. Moreover, Koopman grammians enable discovery of reduced-order models for nonlinear complex systems, as well as provable error bounds on their approximation fidelity.

We developed both model-based and data-driven approaches for Koopman input-output analysis. We can use a model of a nonlinear system and directly compute Koopman controllability and observability grammians. When models are incomplete, we have developed a data-driven approach for learning Koopman operators directly from data. Koopman operators are learned by constructing a dictionary of basis functions; we have developed a scalable and automated algorithm for learning this dictionary using deep learning and the TensorFlow framework. We showed that deep-learning-based methods for learning Koopman operators result in higher multi-step prediction accuracy (i.e., time-series forecasting), as well as increased scalability over state-of-the-art methods using extended dynamic mode decomposition.



A comparison of actual values and different approximations of a two-state nonlinear model with unstable oscillatory dynamics using Koopman



(Left) The 10 generator IEEE 39-bus benchmark system.

PN16063/2840

In FY 2018, we will focus on algorithm development and systems benchmarking. We will extend Koopman learning algorithms to model and analyze nonlinear systems with stochastic dynamics, as well as systems with hybrid and switching dynamics. We will then apply Koopman learning algorithms to capture dynamics of large-scale infrastructure models, use Koopman grammian decompositions to decompose stochastic hybrid systems into subsystems based on input-output properties, and translate this decomposition into a coordination and control architecture.

Data-Driven Decision Science

Yan Chen

Through the project, a machine-learning-based platform has been developed for analyzing energy incentive-occupant behavior using streaming sensor data. We envision this platform can potentially benefit energy incentive program evaluation and behavior analysis.

Incentive programs aim to affect occupant behavior for energy conservation or load shifting. Recent studies have been investigating methods for occupants' behavior analysis using machine-learning algorithms (such as regression model, decision tree). Those studies typically predict the energy usage purely based on historical data but fail to group different energy usage patterns. In other words, energy data has yet to be analyzed to identify the impact of incentive on behaviors. Therefore, this study aims to develop an incentive-occupants' behavior analytics, using machine-learning techniques (i.e., combination of k-means and kernel ridge regression) to learn and predict occupants' behaviors from streaming energy monitoring data.

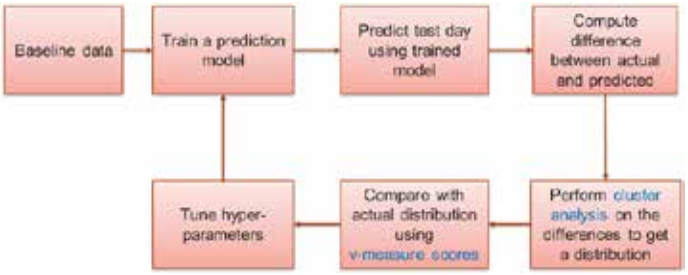
The framework developed in this study is composed with three major parts. The first part is to use training data to develop a model for predicting baseline energy consumption. Since the energy consumption changes also depend on other factors that are not related to the incentive, such as weather, we need the model to forecast the baseline energy consumption so that these factors are already included in the prediction. This prediction model was based on kernel ridge regression (KRR).

The second part is to compute the energy consumption difference between baseline model-based prediction and actual consumption for each premise of the population (e.g., a neighborhood of houses, an entire apartment building). The third part is to group and cluster the premises using *k-means*. Similar deviations (magnitude and shape) are grouped into one category for behavior change using clustering methods.

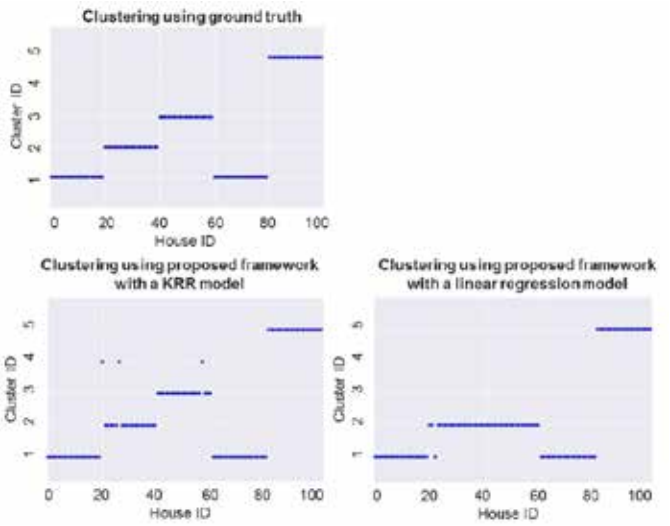
During the model training stage, the ground truth of the occupant behavior change is given with the training data. Therefore, the clustering results can be compared with the ground truth. The performance is evaluated by using the V-measure score. The prediction model may have hyper-parameters that can be set to be certain pre-selected values at the beginning of the process. These values are not necessarily optimal in terms of giving the best clustering results. Based on the V-measure values as feedback, the hyper-parameters can be updated until the best performance is achieved.

Simulation data was used to represent time series building power demand at 30-minute intervals for 100 residential houses during several weeks. With the pre-defined incentivized behavior impacts (ground truth) included in the modeling, we composed a pre-incentive dataset and a post-incentive dataset. Then, the data sets were used to obtain the optimal model hyper-parameters and evaluate the performance.

Evaluation of the framework shows the overall performance of this framework with the KRR model is better than the two other benchmarking methods (i.e., linear regression model and clustering on the raw input data). It can appropriately cluster houses into the relevant behavior groups. These results indicate that that this framework has a potential to be used to analyze the impact of energy incentives on occupant behaviors in real buildings.



The framework developed for energy behavior analysis.



Comparison of clustering results using ground truth, proposed framework with a KRR model, and proposed framework with a linear regression model.

Electrocatalytic Hydrogenation Process Development

Jamie D. Holladay

The overarching aim of this project is to efficiently stabilize and add energy density to bio-crude for the creation of renewable hydrocarbon fuels. To this end, we are developing new condensed phase electrochemistry capabilities that could increase the energy content while improving, and in some cases stabilizing, bio-oil by electrochemical processes.

PNNL has developed various technologies for the production of bio-crude; however, the resulting bio-crude requires high-pressure hydrogen and elevated temperatures for upgrading to a usable bio-oil. The operating conditions result in high costs and, for pyrolysis oils, challenges preventing coke formation. Electrocatalysis offers a new approach for upgrading the bio-crude at mild conditions. In addition, by its nature, it is amenable for moderate to small scale, systems while retaining high efficiency. This enables the development of modular systems, which can be deployed at point of use, greatly reducing the logistics problems, which are another challenge for bio-oils.

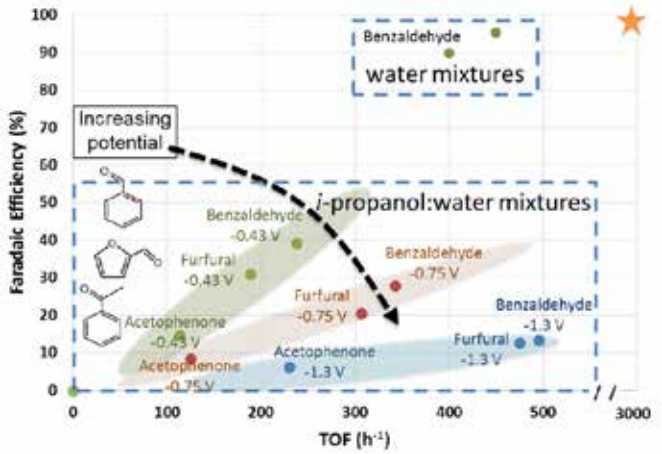
The approach here is primarily synthetic in nature, with the focus being on the determination of conversions, selectivities, and current efficiencies as a function of electrode composition and structure. Others will build on fundamental findings on simple systems to apply electrocatalytic hydrogenation (ECH) to increasingly complex mixtures. An outcome will be the identification of catalyst systems active for reduction of multiple functionalities found in biomass. Cell designs developed by other projects will be used with the best electrodes to determine conversion and selectivity, as well as space time yield rates under a set of standard conditions to be determined.

While PNNL has extensive capabilities in fuel cells for power production, similar capabilities for electrochemical synthesis in the condensed phase are lacking. We will create electrochemical systems and reactor capabilities, along with the expertise in electrocatalyst and reactor design, macrokinetics, deactivation rates, and performance parameters for operating electrochemical conversion systems.

During FY 2017, we developed the flow cell reactor capabilities, in particular, integrating the reference electrode into the system. Being able to isolate the cathode potential is essential to understanding the reaction kinetics and to evaluate different metals as catalysts. We found that integrating the reference electrode on the anode provided the most stable and consistent results. We were able to show that when using the same electrolyte and substrate, the activity, in terms of turn over frequency and faradaic efficiency, is similar when the number of active sites is accounted for.

Using the reactor, we tested a wide number of substrates and several catalysts. For acetyl-furan, hexanal, hexanone, glucose, and vanillin, both the Cu and Pd had minimal activity. The Pd was more active for the conversion of the other substrates than the Cu and had a higher faradaic efficiency. Interestingly, the Cu did not require higher overpotentials to achieve similar currents to the Pd. The activity of the other substrates was as follows: Benzaldehyde > furfural > acetophenone for both catalysts.

From this, we learned that the aldehyde proximity to the double bonds found in the benzaldehyde and fufural seems to improve the conversion. We hypothesized that the alkyl compounds were inactive, since they align with the active sites differently than the compounds with rings. We also observed that the Pd activity for the benzaldehyde, in terms of turnover frequency (TOF) and faradaic efficiency, was significantly lower for the flow cell compared to the reactions in the H-cell but



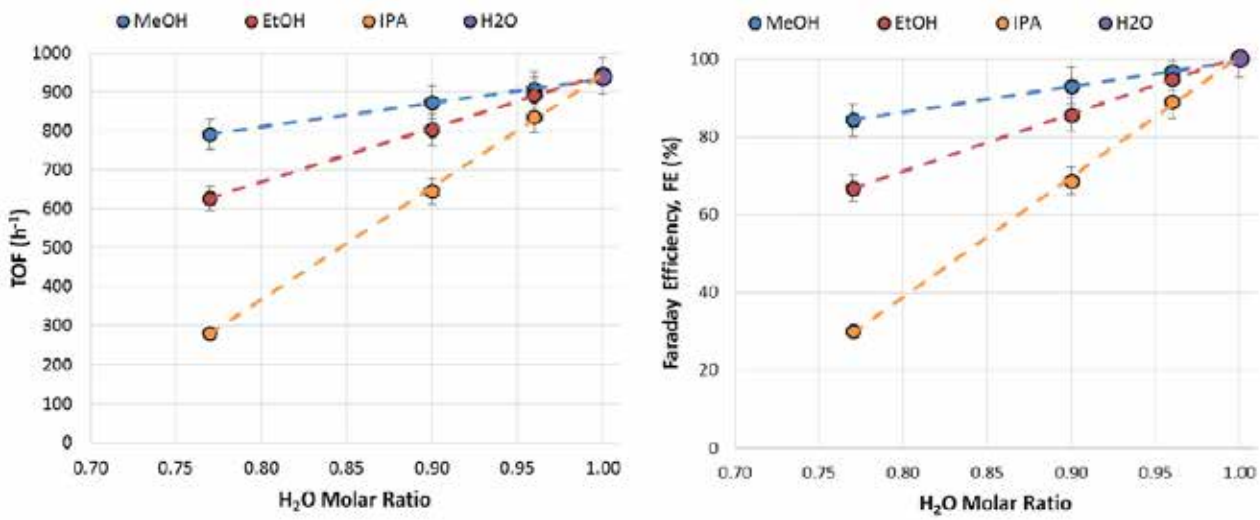
TOF and faradaic efficiency for a variety of substrates all tested in 50:50 IPA:H_2O (mass), 5% acetic acid, and 20mM concentrations. The star represents target TOF and efficiency.

PN17035/2925

using a different electrolyte. The major change was the addition of alcohol. We hypothesized that the iso-propanol (IPA) was covering active sites, displacing the organic substrates.

In FY 2017, we also began looking at mixtures. To begin with, we examined the impact of alcohols. This is important to understand, since to go to industry, relevant substrate concentrations alcohols must be added to increase the substrate solubility. We learned that 1) the more aliphatic alcohols had less impact on the TOF and faradaic efficiency than the non-aliphatic ones; 2) increasing the alcohol concentration resulted in lower TOF and faradaic efficiency; 3) addition of alcohol can shift the reaction order compared to just water; 4) at low benzaldehyde concentrations and high water content, the apparent reaction order is ≈0. As the concentration of benzaldehyde increases, the apparent reaction order increases up to ≈0.7.

During FY 2018, we will deepen our understanding of electrocatalytic chemistry and study increasingly complex catalysts, reaction environments, and mixtures. We will focus on progressively more complex mixtures of aldehydes, ketones, and phenolics; working on the development of more complex catalysts; and develop strategies to achieve oxygen removal from compounds that currently do not reduce.



TOF and faradaic efficiency variation with different alcohols and alcohol concentration. All tests were done with 20mM benzaldehyde and at -50mA.

Experimental Management for Controls of Complex Systems Test Bed

Mark J. Rice

Through the creation of this project's Experiment Management System (XMS), we are reducing researcher experimental overhead and improving the overall experiment process so that additional time and energy can be focused on the research, while simultaneously delivering a more concrete product.

Development and deployment of large-scale physical and virtual experiments on control for complex systems may encompass the description and configuration of up to millions of devices. Researchers often devote significant time and project resources to setting up experiments. A core element of our project is a powerful, flexible test bed that enables researchers who have developed new control schemes to examine performance and behavior in an experimental environment to capture the macroscopic behavior of increasingly complex interacting systems with an adequate degree of fidelity.

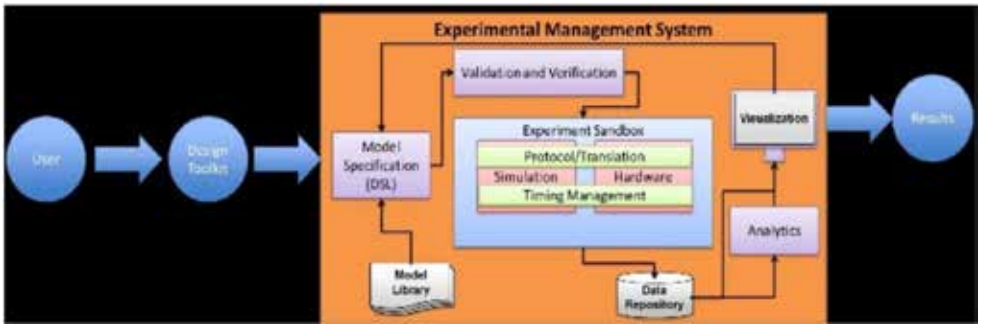
XMS is designed to provide functions such as experimental design assistance and automation, instrumentation guidance, data collection and curation, model cataloging and sharing, and experiment execution. XMS manages experiments performed in the test bed and provides visual analytics capabilities and graphical interfaces to support the management and monitoring of the test bed setup, experiment orchestration, instrumentation, and data analysis. We have defined the requirements and specifications needed to enable tools to be integrated into the test bed and experiments to be run to test and validate control theory and methods. The XMS has been built in a way that allows changes in modules (e.g., simulators) to be implemented without the need for updates to other modules.

In FY 2015, we built the framework to deploy control system experiments using multiple simulators from a Web portal. The project used OpenStack, a cloud computing management platform, to enable the deployment of multiple computers into an isolated network, a tenant, on which to perform experiments. The platform allows the XMS to replicate an experimental setup for other research to start a new study with the same infrastructure.

Within XMS, we built virtual machines (VMs) containing the common simulation software, GridLAB-D™ (an electrical distribution system simulator), ns-3 (a communication simulator), MATPOWER (a wholesale market simulator), and the framework for network co-simulators (FNCS), a PNNL-developed software that enables multiple independent simulators to work in a coordinated manner. Using images of these simulators and Arion, software developed by a related project that helps experimenters by automatically generating code specific to their experiments, we deployed a demonstration experiment to showcase XMS and other tools for an annual review. The experimental setup included multiple GridLAB-D VMs, an ns-3 virtual machine, and the FNCS virtual machine. This demonstration showed that we can create the environment (tenant) in less than 10 minutes to perform the experiment.

In FY 2016, we established communications between XMS instantiated in PowerNET and the test bed assets developed in the Complex Systems Control Test Bed project and the Agent-Based Test Bed for Complex Building Control Systems project. The project also worked on integrating the PNNL agent platform VOLTTRON™, which is used to connect physical devices, with the simulation side of the test bed. XMS is used to deploy new agents that implement different control schemes to VOLTTRON installations, including those for water heaters, electric vehicle charging stations, and the Building Diagnostics Laboratory. In addition, we created a repository for the data describing experiments and collected during experiments, which is hosted in the XMS and provided as a service to other projects that would like to mine the collected data.

In the final year of the project, FY 2017, we continued to build out capabilities for physical platform control, as well as the data collection capabilities. With the VOLTTRON platform, we fully tested an end-to-end connection where a VOLTTRON instance deployed on a VM in the test bed communicated with and collected data from a VOLTTRON agent deployed to interact with a physical building controller. Finally, we worked with the Integration project to create hooks to change our data collection and storage to archive to institutional data archival systems.



The workflow of performing an experiment in the test bed.

PN15001/2676

Hardware Integration Platform for the Control of Complex Systems Initiative Test Bed

Jereme N. Haack

This project is developing an integrated framework for simulation, hardware, and experimental control by adding functionality to the PNNL-developed VOLTTRON™ framework that is specific to supporting other related PNNL projects.

This project works closely with other related projects to integrate all of their results/products into a seamless Control of Complex Systems Initiative test bed platform. Products integrated include those focused on simulation, experiment management, data storage, and a hardware test bed for realistic test deployments. This project develops enhancements to the VOLTTRON platform needed to facilitate easy implementation and execution of experiments that include physical devices and systems, integration of a greater variety of physical devices/systems than possible originally, providing an easy-to-use environment for building new control modules (agents), and integrating physical devices/systems with the simulation environment and the Experiment Management System (XMS).

VOLTTRON serves as the integration point for the physical (and emulated) devices in the test bed. Together with the co-simulation environment, it will establish communication between the physical test bed environment and the simulation environment. It will also support communication for experiment configuration and data storage with XMS. Test bed developers will use VOLTTRON to communicate with the hardware devices, build and deploy modules that implement their control methods for the hardware, and set up topologies required to implement their control methods for experiments.

In FY 2016, the project integrated custom drivers developed for water heaters to bring them into VOLTTRON-platform compliance. This allows them to be used more generally by the VOLTTRON community and increases their maintainability.

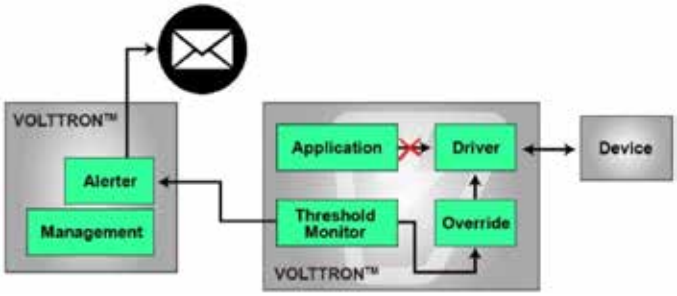
The MATLAB® bridge, developed in FY 2016, is a major addition to the VOLTTRON platform that allows for agents to interact with simulated buildings and the testing of MATLAB-based control algorithms. Additionally, this capability generated interested from the greater VOLTTRON community.

Integration with simulations was further enabled by building an integration point with a PNNL simulation project. This powerful combination allows large-scale simulations to interact with hardware being controlled by the VOLTTRON platform.

Cross-network data and control streams were also demonstrated in FY 2016. Data collected on the isolated facilities network were directed to a protected network via the VOLTTRON platform. Control actions inside the protected network were then sent out to the PNNL network to update agents being deployed in the test bed. This capability was further developed in FY 2017 to provide flexible control and collection capabilities across multiple networks.

Also in FY 2017, existing demonstration, utility, and example code was updated to work with the latest version of VOLTTRON. Agents taking in external forecasting and utility signals were rewritten as standard VOLTTRON agents, simplifying support and making them usable by the wider VOLTTRON community.

An integrated demonstration of XMS and VOLTTRON utilizing new features of the platform was developed. In this demonstration, XMS sends an experiment definition to VOLTTRON, which then deploys the application and configures a Threshold Monitor agent in VOLTTRON. This issues an alert when sensor readings exceed limits, causing an override agent to lock out application control and reset devices to a known state. Another agent sends out email to experimenters and test bed administrators.



XMS (3.4)/VOLTTRON™ experiment deployment, monitoring, and override.

PN16012/2789

Integrated Control Testing Under Complexity

Jacob Hansen

This project investigates prototype-stage advanced control methods to understand their behaviors within large-scale interconnected energy systems via simulation and hardware-in-the-loop (HIL) experiments of unprecedented scale and complexity.

Power and energy systems are undergoing substantial changes, from highly predictable systems with central generation, to systems with high penetration of non-dispatchable, variable, renewable energy resources, increasing the overall system uncertainty and decreasing the controllability. These changes have necessitated the deployment of increasingly advanced control schemes to balance supply and demand, both short- and long-term, while maintaining operational costs at a competitive level. Such advanced control schemes are often designed using simplified models of the power system without considering the full complexity present in interconnected power systems, how the solution might work within the context of legacy control systems, or the structural position within a larger architectural framework. Neglecting this complexity makes them highly susceptible to unforeseen and potentially undesirable effects on the energy system as a whole. At low penetration rates of renewable generation, these assumptions are quite practical; however, envisioning a more active and distributed grid of the future requires examining the interactive nature of high renewable penetration cases in greater detail.

Previous efforts have taken on the task of developing advanced control methods targeted towards the power grid. The resulting algorithms have been tested on small test systems with assumed boundary conditions, showing tremendous performance improvement according to the objective of the control system. However, to truly assess the control system, this project evaluates the same control system using the Control of Complex Systems Initiative (CCSI) test bed. The CCSI test bed is capable of simulating large-scale power systems using co-simulation capabilities to capture interactions in the power grid from the distribution system to bulk power operations.

In FY 2017, this project evaluated two control systems using the CCSI test bed. The first control system, which was developed by a project focused on distributed control, takes a transactive control approach consisting of price responsive residential heating and air conditioning systems that react to changes in the wholesale locational marginal price (LMP) of electricity. The LMP was assumed to be a boundary condition and the wholesale markets were not modeled. In previous demonstration efforts, this boundary condition was removed, and a simple transmission system was modeled to create a co-simulation of transmission and distribution. In doing so, price oscillations were observed in cases with high penetration of flexible loads.

To address these price oscillations, this project added a new control layer that collects distribution system flexibility information and communicates it to the wholesale market. This allows the wholesale market to dispatch the flexible load and eliminate the aforementioned price oscillations. Furthermore, this project also expanded on the complexity and scale of the systems for which the co-simulation capability could evaluate control systems. The current evaluation is able to simulate an interconnect-scale power system (e.g., the Western Interconnection, which serves the western United States) with highly detailed distribution system load modeling along with appropriate existing control systems to capture more realistic power system operations.

The second control method evaluated was developed in a project focused on the impacts of communication networks on the performance of distributed control. The method uses a distributed consensus algorithm that is robust to communication imperfections. The algorithm was tested using simple models in a single simulator to represent the distributed control system.

As part of this project, a distributed agent co-simulation framework was developed with every agent modeled as a separate simulator. This enables evaluating the control algorithm in a truly distributed manner. Because these are actual distributed simulators, they can even be installed in physically separate systems. Using the control theory from the earlier project, this distributed co-simulation framework was used to solve the day-ahead economic dispatch problem. Individual agents could communicate with each other, and all messages between agents were routed through an advanced discrete network simulator, ns-3. Short and long network delays, and both power line and fiber optic communication infrastructures were considered. These technologies were used to evaluate the impacts of communication

PN17057/2947

imperfections on the control system performance. Simulation studies showed that communication imperfections do impact the time to solution but, overall, the control system can converge to correct solutions when ample communication links are present between the distributed agents.

In FY 2018, this project will focus on two major efforts. First, it will continue to evaluate the transactive control system, adding additional complexity in the wholesale markets by modeling not only real-time but also day-ahead interactions. Furthermore, larger scale simulations will be performed to increase the realism of the power system cases examined. Secondly, the project will evaluate a building-to-grid control method by performing HIL simulations with both unoccupied and occupied buildings on the PNNL campus.

Measurement and Verification in Controlled Complex Systems

Jim Follum

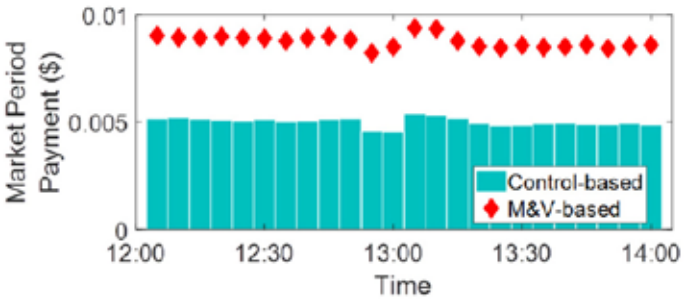
As revolutionary control architectures are implemented in complex systems, their physical and financial impact must be fairly, quantitatively, and efficiently evaluated. To meet this need, measurement and verification (M&V) principles and approaches are being extended to fit emerging control approaches and address the challenges of large, complex systems. The resulting M&V methods will restrain the costs of verification, provide ongoing monitoring capabilities, and better link incentives with measured performance.

Effective control of large-scale infrastructures, such as energy, building, and transportation systems, can improve efficiency and reliability, but associated methods of ensuring that the desired improvement is realized are largely unexplored. Traditional M&V methods require detailed, manual, and at times subjective analysis of datasets collected over long periods of time, making them ill-suited to large, complex systems. The objective of this project is to extend M&V principles and approaches for application to complex systems. The resulting methods can be automated to allow ongoing performance monitoring and impact quantification at each level of the control architecture. Adoption of these M&V methods will accelerate market deployment of emerging control strategies by mitigating risks for investors, increasing trust among participants, and providing an avenue to evaluate contractual obligations.

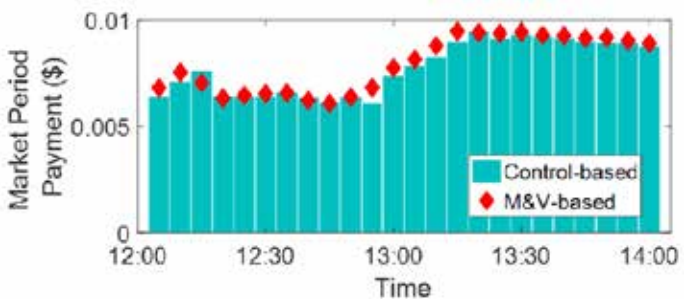
The first M&V approach that has been developed is based on data mining methods. These methods, such as neural networks and support vector machines, are used to establish the baseline behavior of a complex system from measurements collected before an advanced control method is implemented. After the control is deployed, additional measurements are compared against the baseline to estimate the physical and financial impacts of the advanced control method. The data mining approach was tested with field measurements from a building on the PNNL campus.

As a complementary approach, an M&V method that compares the expected behavior of distributed control devices with measured performance in near real-time was developed. The method is particularly useful when deployed alongside a transactive control, because expected device behavior is made explicit in bids. Experiments in a building on the PNNL campus demonstrated the method's ability to identify poorly modeled air conditioning zones and validate transactive-control payments.

The experience of deploying M&V in a commercial building made the need for an M&V tool clear. The project team will be identifying resources to develop a software tool implementing the methods developed in this project. It will leverage the VOLTTRON™ platform to collect measurements for analysis and return results to the user. The tool will allow the user to monitor the performance of an advanced control method at various levels of the architecture and provide verification for incentives paid out to control participants.



Shown here are the initial results obtained for an air conditioning zone when an integrated M&V method was deployed alongside a transactive control scheme in a commercial building. The significant disagreement between the control- and M&V-based payments helped identify a poorly tuned model used by the controller.



Shown here are the results obtained for the air conditioning zone after the model was retuned. The improved agreement between the control- and M&V-based payments indicates that the control is operating properly and provides validation for the transactive control payments.

PN16002/2779

PACiFiC: Proactive Adaptive Cybersecurity Framework for Control

Paul M. Skare

The power grid relies on Operational Technology (OT), as do numerous other high consequence process control systems. The goal of this investment is to demonstrate, for the first time, that OT can, in principle, be made secure by developing adaptive systems—ones that are proactive, flexible, and dynamic in response to cyber threats/events.

OT systems consist of process control systems and their associated network infrastructure. This effort will build upon the study of combined cyber-physical-social dynamics in high consequence systems; in doing so, we will explore OT systems like buildings, transportation, grid, and other critical control systems. Our research and development leverages the known physics of physical systems to bound cyber space. Process control limitations, physical boundaries, chemical, and material laws will limit detection and response margins to provide OT with security, resilience, and control. Additionally, leveraging cognitive, social, and psychological understanding will add an additional dimension by which cyber space can be bound. Merging physical and social models with a cyber model can turn an, as of yet, untraceable problem of cyber security detection and automated response into an achievable approach.

We are building a reference system for generic (sector agnostic) OT that does not exist today and serves as a model for government and industry sponsors to extend cyber-physical testing to include simulated, virtualized, emulated, and physical computing devices.

During the first year, we worked on establishing Secure Design and Development principles. *Secure Design and Development Principles for Control Systems* presents a holistic approach to design and develop computerized systems for use in control systems. It documents best practices from various disciplines, including traditional Information Technology (IT) and OT, to securely design, develop, and implement control systems. The goal is to present clear, concise, and actionable recommendations for designing, building, deploying, operating, maintaining, and withdrawing control systems, all in a secure manner.

Control systems are specialized computer systems that interface directly with the real, physical world. Unlike IT systems that operate on data from keyboards and databases, control systems use sensors to measure physical properties, such as temperatures, flows, and pressures in the physical world, and use actuators to move valves and drive motors. Control systems often use feedback mechanisms (closed loop control) that use sensors to monitor the results of the actuators and to further refine actuators until the desired physical process characteristics are achieved. The control systems then continue to use the sensors to monitor the physical process for external actions (i.e., those not from the control actuators) and constantly make adjustments with the actuators to keep the physical process within desired parameters.

While some sections in the developed document may be more applicable to particular users than others (i.e., programmers versus hardware designers), readers are encouraged to read the entire document to establish a comprehensive understanding of control systems and their applications.

Additionally, PACiFiC invested in various tasks named PACiFiC-starters. These efforts were explorations in high risk, high reward topics that we identified as having applicability to immediate needs, future development, or alignment with our own objectives for next year.

AlphaClock: Machine Learning for Enhancing the Timing Resilience of Cyber-Physical System. The AlphaClock team has completed a detailed literature/technical review to identify existing techniques for enhancing the timing resilience in cyber-physical systems. By using the smart grid as a use case, we compared existing techniques' capabilities of detection and mitigation for GPS jamming and spoofing. It also provided insights on how the proposed measurement-based methodology can overcome those shortcomings. Moreover, time sequence synchrophasor data has been generated through power system dynamic simulations to formulate the training input and output for machine learning process. Currently, we are working on the model selection and training for the detection of GPS jamming and spoofing.

Peripheral Authentication and Authorization for Legacy OT Networks. The Peripheral Authorization goal is to investigate the feasibility of creating an in-line device that can

PN17095/2985

independently verify the identity of a user and authorize access to communication resources based on work schedule and permissions. PNNL's Secure Software Center assisted in the high-level software design through the Class, Responsibilities, and Collaborators process. Further development of the software design and identification of software interactions revealed a threat model with 91 potential vulnerabilities. The threat modeling methodology is based on Microsoft's STRIDE tool.

The primary method of investigation is team thought experiments bounded by the initial assumptions of the system and operating environment.

Risk-Informed Autonomous Adaptive Controller. A framework for risk-informing a cyber controller in an operational setting was described, considering a business following enterprise-control integrated system architecture, as laid out in IEC 62264-1 Standard. A combination of event trees for tracing attacks through various business functions, fault trees for causative analysis associated with each business function, and cyberattack graphs to provide linkages to fault trees was used to trace and map cyber elements with business operations. The framework was complemented with a methodology to formulate reconfiguration alternatives given some contextual attack information in real-time. The usefulness of the framework was highlighted through a methodology for estimation of risk associated with impending attack paths along with estimates for adaptive actions to confuse an adversary.

Secure Operating System Design for Control Centers. Through this task, we developed an initial design and found supporting technologies suitable for making our proposed capability work. From our design, some of the capabilities procured for the test bed networking equipment were driven from the needs of this capability to allow leveraging some Cisco capabilities to accelerate the prototyping. We have high confidence the proposed capability can be developed by creating new functions in QubesOS that provide a secure proxy virtual machine (VM) that bridges communication between the application VM and the physical network interface to enable the additions of metadata to inform a Cisco Integrated Security Engine switch how to properly segment traffic.

TMBR – Threat Model Based Response. The goal is to decompose threat information into base components of functionality that can then drive mitigation actions. Currently, we are pulling data on malware to utilize in our processes. It appears that the data set there has a large number of duplicates that are different in non-critical elements, further distorting the analysis. We are supplementing this with additional data sets.

Automated Detection and Response. In this task, we have been testing the various areas of indicators and actuators. This includes developing a python library that can interface with the devices configuration management console to extract relevant information and to change settings for the software based indicators and actuators. For the side channel indicators, we have developed an Arduino-based sensor platform for putting inside the device to monitor the temperature, vibration/movement, light, pressure, etc. For the hardware indicators and controls, we have been applying JTAG and Bus analyzers to study what capabilities are on the hardware. We have found the JTAG port and interface with the FPGA on the board but have been unable to read or write anything to it. Finally, we are using existing tools to capture and analyze communication to the device for the network-based indicators.

Going forward, we will be exploring the next step for following on the design principles and the tasks identified in the PACiFiCstarters.

Reducing Emissions Associated with Vehicle Cold Start Via the Coupling of Exothermic Carbonization Reactions with Conventional Aftertreatment Catalysis

Kenneth G. Rappé

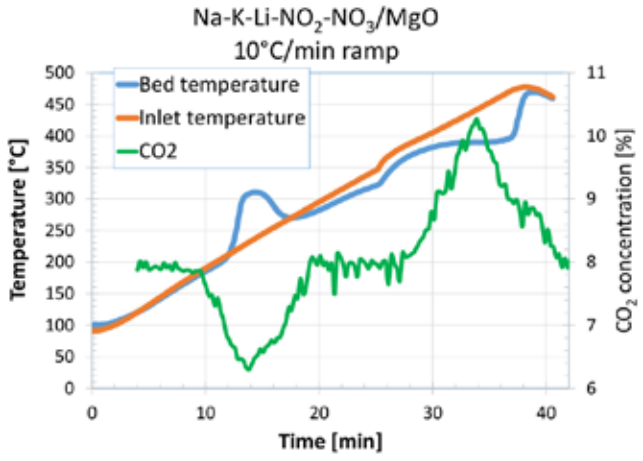
Pollutant emissions released from vehicle engine exhaust during the cold-start period represent up to 95% of the total hydrocarbon and NO_x emissions released over the entire driving cycle. Therefore, this work is focused on developing technology that can passively add heat to aftertreatment catalysts, heating them up quicker and shortening the cold-start period, thereby decreasing cold-start emissions and total emissions.

Current aftertreatment catalysts function inadequately at low temperature and fall short of the emission conversion targets identified by DOE and industry. Storage and release materials are being actively pursued to bridge the gap between vehicle start-up and catalyst light-off, including hydrocarbon traps and PNAs (passive NO_x adsorbers). However, significant challenges remain with these materials, including sufficient trapping capacity of emissions and the matching of release behavior to conversion. This project is focused on developing a class of adsorbents that can passively heat aftertreatment catalysts. This is accomplished by using exothermic CO_2 capture to produce heat at low temperature and release that CO_2 at higher temperatures in the exhaust, thereby functioning reversibly and repetitively.

The adsorbent material technology developed in this work is based on more than 3 years of previous PNNL development work. The adsorbents are composed of bulk metal oxides promoted by molten salts that can act as phase transfer catalysts to significantly facilitate the CO_2 reaction in a highly exothermic fashion. The capability of a molten eutectic component to dissolve bulk MgO facilitates a dynamic MgO dissolution and precipitation equilibrium providing activated MgO accessible to CO_2 reaction that otherwise is not. Although the eutectic component exists as a molten phase during CO_2 capture, the bulk material can still be handled as a solid sorbent. The highly exothermic nature of the sorbents provides heat strategically during the cold-start period of vehicle operation, thus shortening the "effective" cold-start duration and decreasing the emissions released. These materials, when operating properly, can represent a paradigm shift in the ability to decrease emissions during cold start.

The initial efforts in this work focused on determining the optimal CO_2 adsorption and release temperature windows, optimizing the total CO_2 adsorption capacity and associated exotherm, and demonstrating the feasibility of the approach with aftertreatment catalysts. Eutectic systems were pursued with melting points < 150°C to target capture CO_2 at the same temperature. These included the ternary $NaNO_3/KNO_3/NaNO_2$ system, the ternary $LiNO_3/NaNO_3/KNO_3$ system, and the quaternary $LiNO_3/NaNO_3/KNO_3/NaNO_2$ system. Thermogravimetric (TGA) analysis in model environments helped guide sorbent synthesis and screening to identify windows of proper material operation. Results of this effort showed feasibility to lower the temperature of the onset of CO_2 capture to < 150°C.

Testing of the most promising sorbent system from the initial TGA work showed excellent CO_2 capture under realistic cold-start conditions on a bench testing apparatus using synthetic exhaust. The results of the capture was an effective and usable exotherm that began at 200°C and resulted in an

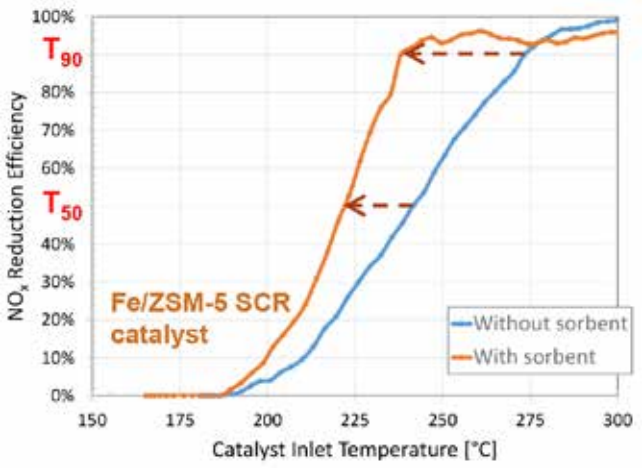


Sustained temperature rise of the adsorbent system during realistic simulated cold-start conditions resulting from ~30 wt% CO_2 capture.

PN17068/2958

extremely fast temperature rise of more than 100°C. This temperature rise is only usable if it is sustained, and results showed that, with increased temperature ramping during the simulated cold-start test that better represents reality during deployment, it actually led to a more sustained net temperature benefit.

By coupling the adsorbent with an aftertreatment catalyst, the impact the light-off curve could be measured. The experimental system chosen for demonstration was based on a NO_x-selective catalyst reduction catalyst. In this demonstration, the adsorbent (or sorbent) decreased the T50 value of the catalyst (the temperature at which 50% of the NO_x is reduced) by approximately 20°C and the T90 value by approximately 35°C. This represents up to a 15% reduction of NO_x emissions released during cold start and up to a 12% reduction of NO_x emissions released over the entire drive cycle. Similar results are expected with the catalytic system used to reduce hydrocarbon emissions.



Impact of the exothermic CO_2 capture by the sorbent on aftertreatment catalysis light-off.

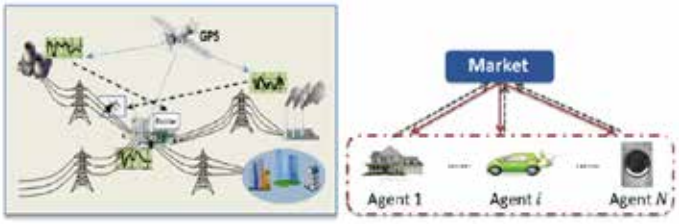
Resilience in Large-Scale Distributed Control Systems

Wei Zhang

This project aims to develop approaches to analyze, quantify, and improve resilience of distributed control systems. The research led to new analysis and design tools for distributed control of the electric power grid.

In distributed and physically interconnected systems, the effects of local physical faults or cyberattacks propagate through the physical and communication networks and impact the rest of the system, presenting a challenge for the resilient operation of the overall system.

This project aims to develop theory and algorithms to analyze and quantify resilience of distributed control systems in the context of power system applications. We proposed a universal resilience index that quantifies the resilience of a distributed control system as the performance degradation under abrupt change in the system structures. Using this definition, we have studied the resilience analysis and quantification problem for two different types of distributed control systems that are important for power system applications: wide area control (WAC) systems for the bulk power grid and transactive systems for coordination of distributed energy resources.



Two classes of distributed control systems studied in this project (Left: WAC; Right: transactive).

WAC is an effective means to improve the stability of the future power grid with increased penetration of renewables and distributed energy resources. All types of WAC rely on real-time Phasor Measurement Unit (PMU) measurements sent over communication networks. We studied WAC resilience under a class of cyberattacks that may cause loss of communications from the PMU sensors.

We showed that for a given stable wide area controller, attacking certain communication channels can cause instability of the closed-loop system. We also proposed to use H2 norm as the performance index for a WAC system and designed a resilience index based on the H2-norm degradation under the worst attack. Finally, we developed an efficient algorithm to compute the resilience index for any given WAC. In addition, we showed that the algorithm is guaranteed to converge to a solution that quantifies the resilience of a WAC system.

Our results have been tested through realistic simulations on the IEEE 39 bus system. The developed framework, along with the algorithm, provides a formal way to compare different WACs, from both a performance and resilience perspective, and helps identify the most critical communication links to be protected from cyberattacks.

At the beginning of FY 2017, we started working on the resilience analysis problem for transactive systems. Such systems employ market mechanisms to engage and coordinate distributed energy resources. Existing research has mainly focused on improving performance of such systems. Resilience analysis of a transactive system is a wide open problem that has not been adequately studied in the literature. In this project, we studied a multi-period double auction market that is similar to the one used in the American Electric Power gridSMART® demonstration project.

We proposed to model the bidding decision of a participating asset as the result of a certain utility optimization problem and model the market clearing process as a social welfare optimization problem. Given the model, we developed a mathematical formulation that allows us to study price stability over multiple market periods. Under this formulation, we derived analytical conditions on the assets' bidding function and their dynamic models under which the resulting market price signal is stable.

Our theoretical study provides insights on how to design the bidding functions to avoid large price oscillations. It also provides a foundation for us to study and understand the kinds of cyberattacks or failures that may cause price instability.

PN15007/2682

Standardized Contracts for Energy and Reserves from Responsive Loads

Jacob Hansen

Electric power markets have become increasingly complex over the years. This project designed and tested a simplified market clearing concept that ensures safe and reliable operations of the power grid.

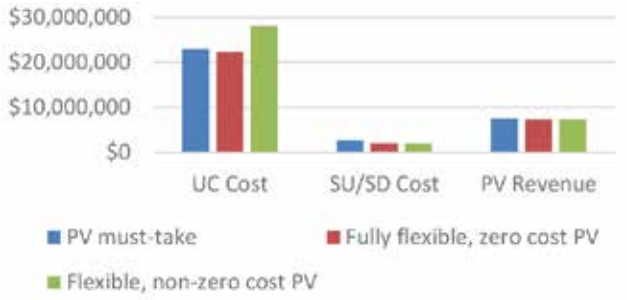
The wholesale power markets have undergone substantial change over the past few years, in both structure (i.e., ownership and technology aspects) and architecture (i.e., operational and oversight aspects). The market operators have often responded to these changing needs by defining new products and services to ensure reliable and stable system operations. However, too many products and services, competing for the same resources, can lead to thinly traded markets, causing distortion of the price signal and potential for market manipulation by market participants. In addition, the current market designs do not adequately value and compensate for the flexibility that various resources provide on a day-to-day basis.

A new market design based on the two-part pricing concept was introduced in literature, which involves the use of *standardized contracts* to procure capacity for provision of both energy and reserves. In this formulation the Independent Service Operator contracts with and compensates the market participants (both generators and demand-side assets) for making their capacities available in the day-ahead market, while paying them a performance payment based on actual production in the real-time market (RTM). Numerical results are included validating the theoretical results.

With a standardized contracting mechanism, resources do not submit multiple bids to provide different grid services. Instead, the resources make a single capacity bid, which means that the independent system operators do not have to operate multiple commodity markets. The proposed mechanism allows for clear representation and evaluation of the flexibility offered by various resources in very comparable terms.

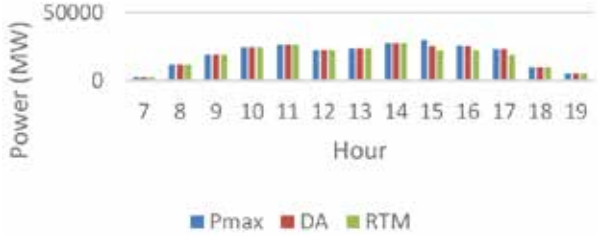
In this project, we continued to refine and test the proposed market design based on a standardized contracting mechanism for all technologies. The simulation results show that the proposed formulation is able to solve the combined unit commitment (UC) and economic dispatch problem, with and without demand-side flexibility, along with different reserve margins. In addition, increasing flexibility of the demand side leads to a more economical solution with lower total cost. The performance of the standardized-contract-based UC process was also compared with the conventional UC, based on conventional resource bidding. It was observed that the total cost does not increase if the SC-bids can be constructed in a realistic manner.

One of the use cases analyzed involved representation of photovoltaic (PV) capacity in markets using standardized contracts. Essentially, instead of treating PV as a must-run capacity and netting it out of load, PV resources were allowed to bid in an availability price and a performance price. The idea was to see under what conditions the market would commit PV but not dispatch it fully, so that the capacity may be available later to either ramp-up or down, as needed. An interesting outcome of this test case was that the flexibility offered by PV made very little impact in day-ahead market, where the UC decision was dominated by high costs to start-up (SU)/shut-down (SD) baseload units. However, more impact of flexibility on commitment schedules was seen as the PV penetration increased in the market.



Power Purchase Agreement vs. Standardized Contract: 45% PV

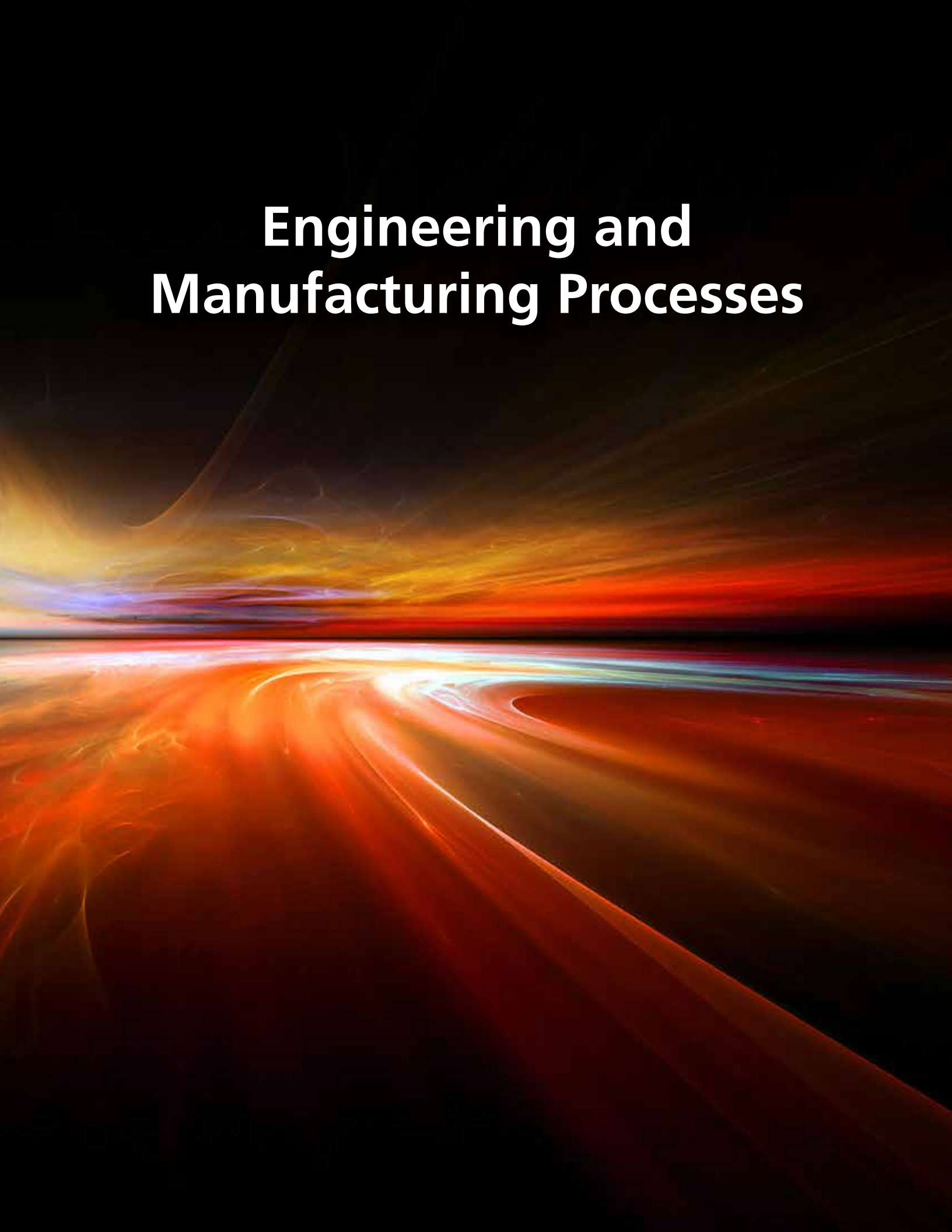
In general, under all kinds of scenarios, the real benefit from PV's flexibility was realized in RTMs, when PV could be ramped up/down to deal with forecasting errors associated with load or even solar ramps. The latter was not surprising, but the lack of substantial effect on UC decisions was a revelation. This discovery allowed the project to infer the general value of flexibility in market and system operations.



PV Pmax vs. Day Ahead Dispatch vs. RTM Dispatch

PN17037/2927

Engineering and Manufacturing Processes



Bulk Thermally Stable Nanocomposite Processing

Scott A. Whalen

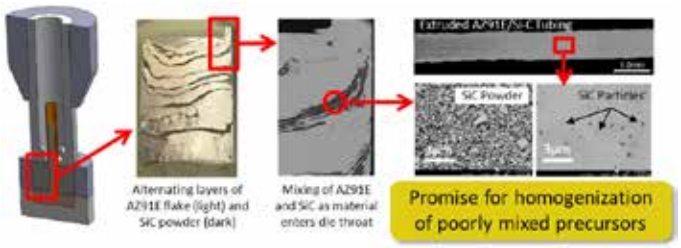
In this project, we are developing the underlying science and engineering processes for fabricating bulk materials with thermally stable nanocomposite microstructures. Our research in far-from-equilibrium shear processing is envisioned to enable fabrication of industrial-scale products, with game-changing material properties that have heretofore only been achieved in powder, flake, thin-film, or very small specimens.

The goal of this project is to use shear-assisted processing and extrusion (ShAPE) to consolidate and extrude, in a single step, nanocomposite Mg-X and Al-X precursors (in billet, powder, or flake form) without any loss of microstructure in the resulting bulk extrusion, thereby maintaining property improvements derived from nanoscale features in the precursor. Current Assisted Pressure Activated Densification (CAPAD) is also being used for the same purpose in FeSi alloys to consolidate bulk nanostructured soft magnetic materials. The processes developed and material properties achieved have already led to quality journal publications in the fields of magnets (FeSi), thermoelectrics (Bi_2Te_3), and lightweight metals (Mg Alloys).

A key accomplishment in FY 2017 was completing the installation and run-off testing—at PNNL—of the world's first ShAPE machine and new turn-key CAPAD instrument at the University of California-Riverside. The ShAPE process was used to consolidate a variety of different precursors into extruded tube and rod. A few highlights follow.

AZ91 flake, fabricated by melt spinning at PNNL, was consolidated into tubing with an outer diameter of 7.5 mm and wall thickness of 0.75 mm. Billets of as-cast AZ91 were also extruded to the same tube dimensions and the microstructure was compared between tubes formed with AZ91 flake and as-cast billets. Microstructural analysis showed < 3 µm grain size in the tubing and that ShAPE was effective at breaking up millimeter-scale second phases and dispersing them as micron-size particulates in the bulk tubing. In other experiments, Al-1,5,10 at. % Mg billets (a composition to have improved thermal stability) were extruded into 2.5 mm diameter rods where subsequent heat treatments and grain size analysis will be compared with thermal stability simulations.

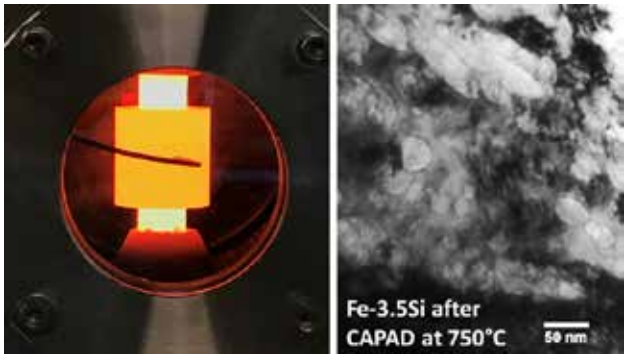
Other experiments were designed to assess the degree to which the ShAPE process can homogenize "dirty" starting material. To examine this, precursors of the AZ91 flake ball mixed with 1 and 5 wt. % alumina, and yttria and silicon carbide were formed into the typical tube size to create a metal matrix composite.



An example of ShAPE's ability to homogenize precursors. Here, 20 alternating layers of AZ91 flake and SiC powder are extruded into a tube with dispersed SiC particles.

Other alloys successfully processed via ShAPE include Mg_7Si, which is too brittle to extrude by conventional means, and Mg-Dy-Nd-Zn-Zr rare earth magnesium.

CAPAD was used to consolidate FeSi soft magnetic milled powder and flake into densified pucks with a 16 mm diameter and 1 mm thickness. Melt pun flake of Al-Mg was also consolidated.



Shown here is one exciting result of our experiments: Fe-3.5Si powders with 9 nm starting grain size grew to just 16 nm after consolidation to 95% density at 750°C.

In FY 2018, the project will polish off some experimental gaps and then focus on publications and leveraging the successes of FY 2016 and FY 2017 into follow-on projects.

PN16022/2799

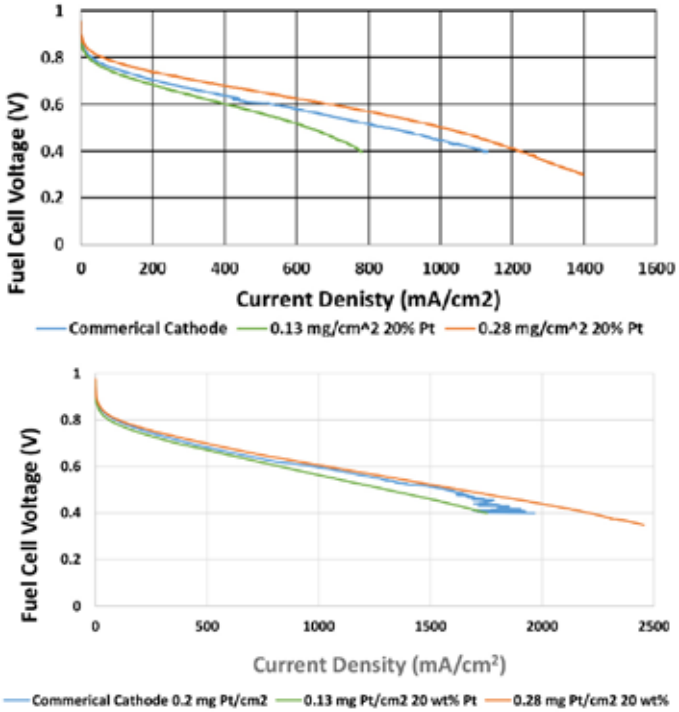
Membrane Electrode Assembly (MEA) Innovation for Fuel Cells

Yuyan Shao

This project is focused on improving a proton exchange membrane (PEM) fuel cell key capability: MEA and platinum group metal (PGM)-free catalysts.

MEA performance was improved by 200% in O_2 and 25% in air over previous results. This was achieved through component innovation and MEA assembly protocol improvement.

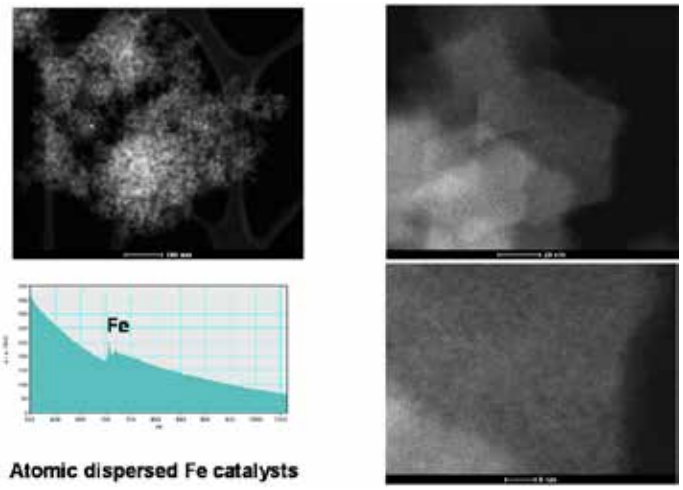
We first used PGM catalysts to optimize cathode electrode preparation, MEA assembly protocol, and MEA operation protocol. This includes ionomer/carbon ratio, catalyst ink (solvents), hot pressing, MEA assembly pressure, H_2/O_2 (H_2/Air) stoichiometry, relative humidity (RH), etc. Our new MEA demonstrates better performance than the one using a commercial cathode in both the H_2/Air and H_2/O_2 operation. The MEA capability has been successfully developed for PGM catalysts, which can be leveraged for future PGM-free catalyst development.



Fuel cell MEA performance of H_2/Air (Up) and H_2/O_2 (Down). H_2/Air (H_2/O_2) stoich 1.5/3, fuel cell temperature and dew point = 80°C, gas inlet at 90°C.

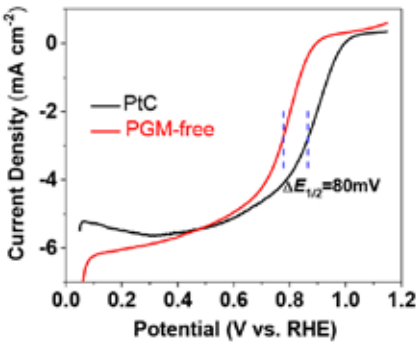
According to DOE tech-eco modeling, PGM catalysts and their application account for almost 50% of total fuel cell cost at large-scale production. To replace PGM catalysts with low-cost, PGM-free catalysts is critically important for fuel cell commercialization. However, today's PGM-free catalysts are still much poorer than PGM catalysts in terms of activity and durability.

In this project, we have made progress on PGM-free catalysts for PEM fuel cell application. Through catalyst precursor innovation (i.e., metal-organic frameworks structure and morphology innovation), we are able to develop atomically dispersed PGM-free catalysts. In these catalysts, metal nanoparticles were not observed. However, metal (e.g., Fe) was observed under electron energy loss spectroscopy (EELS), indicating single atom Fe catalysts.



Atomically dispersed Fe-based PGM-free catalysts. Transmission electron microscopy images show no metal nanoparticles and EELS show Fe element, indicating single Fe atoms are dispersed in carbon matrix, most likely in the form of Fe-N_4, which is considered to be the most active sites.

Our PGM-free catalyst demonstrates high oxygen reduction activity, with half wave potential of only approximately 80 mV lower than that of benchmarking Pt catalysts.



Oxygen reduction curve of PGM-free catalyst in comparison with Pt/C catalyst.

PN17003/2893

Novel Heat-Treatment for Automotive Castings

Aashish Rohatgi

The goal of this project is to develop a novel, energy-efficient, heat-treatment technique for metallic alloys. We anticipate that this technique will enable development of high-strength, light-weight, and low-cost aluminum and magnesium automotive castings that will lead to better fuel economy and reduced green-house gas emissions.

Heat-treatment of metallic alloys is essential to obtain desired functional performance such as strength, toughness, fatigue life, and corrosion resistance. Heating a material activates atomic-scale phenomena that produce the microstructures responsible for the material's functional performance. These microstructural changes often occur very slowly, which, therefore, require many hours of heating. Such extended heating is energy intensive and increases the product cost. For example, the required time in a traditional oven/furnace for heat-treating aluminum (Al) castings can be quite long (up to 24 hours). Therefore, some metal castings, despite their potential to reduce a car's weight, may be uneconomical (due to added costs of heat-treatment) for the mass-market automotive industry. In this project, we propose to develop a novel, energy-efficient, heat-treatment approach to accelerate the heat-treatment process and, thus, lower the associated energy requirements, time, and cost.

In FY 2016, an aluminum casting alloy, referred to as a 319 alloy in the aluminum industry, was used as the test material in this work. The as-received material was machined into test samples and subjected to different heat-treatments using both conventional furnace heat-treatment and the new method. The heat-treated samples were analyzed using metallography, scanning electron microscopy, hardness testing, and differential scanning calorimetry.

The microstructure of the as-received 319 alloy was a typical as-cast microstructure with aluminum dendrites, Al-Si eutectic, and various intermetallics comprising Al in different proportions with copper, iron, manganese, and silicon. The hardness of the as-received, as-cast material was measured to be approximately 45 RB (Rockwell B scale), while the hardness of the solutionized (495°C for 8 hours) plus aged material showed higher hardness of approximately 73 RB. The key result from the heat-treatment experiments was that the sample that was solution-heated by the new method, followed by aging, could achieve a similar high hardness with only 1 hour of solution heating, supporting our hypothesis of the advantage of the new approach.

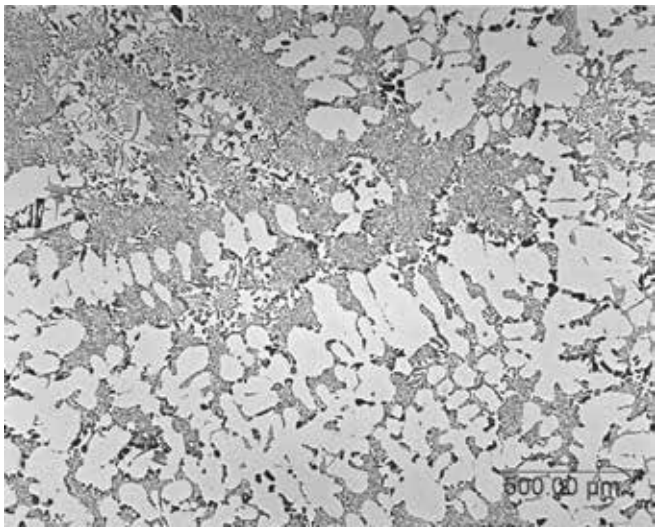
In FY 2017, we extended our work to evaluate if aluminum alloys other than alloy 319, such as alloy A357, will respond favorably to our new heat-treatment approach. For heat-treatment of alloy A357, the standard solutionization temperature is 540°C for 10 hours, followed by multi-step aging. The hardness of the as-received A357 alloy was approximately 12 RB, which, after standard heat-treatment, increased to approximately 52 RB. When the alloy A357 was solution-heated by the new method, followed by multi-step aging, it could achieve a hardness of approximately 50 RB with only 1 hour of solution heating. This result is qualitatively similar to that achieved with alloy 319 in FY 2016. However, the hardness of the samples heat-treated by the new method showed quite a bit of sample-to-sample variation, which was unexpected and requires further investigation.

In addition to performing heat-treatment experiments, hardness measurements, and microstructure characterization, commercial software from Ansys was used to develop a multi-physics model to simulate the specimen heating. Thermal and other physical properties of the Al alloy used in the model were obtained from the technical literature. Finite element solvers in the Ansys model were used to predict the sample temperature distribution as a function of the process parameters. While, in FY 2016, uniform specimen geometry was modeled, the work in FY 2017 extended the modeling approach to non-uniform specimen geometries.

The heat-treatment models, for both uniform and non-uniform specimen geometries, demonstrated that iterative feedback between the solvers allowed the simulation results to converge within several iterations. Good correlation between the measured sample temperature and the model predictions were obtained with a heat-loss coefficient of 5–10 $W/m^{2}\cdot°C$ to account for convective heat losses from the sample surface.

PN16042/2819

In summary, a new method for heat-treatment of alloys has been developed that, in Al alloys, shows promise to reduce the heat-treatment time, while still achieving similar mechanical performance as obtained through conventional heat-treatment. Additional research is necessary to optimize the process parameters to increase the robustness of the new heat-treatment process. A multi-physics model has also been developed that is capable of predicting the specimen temperature for different specimen geometries.



Optical micrograph of the as-received aluminum alloy 319 sample showing the as-cast microstructure that typically needs to be heat-treated to improve its mechanical properties.

Reactor and Process Design

Robert S. Weber

We are inventing and characterizing new chemical reactors that can convert regional- and community-scale quantities of waste into fuel or chemicals. The reactors must be able to deal with a wide range of feedstocks, they must be inherently safe, and they should not rely on external facilities for co-reactants or heat rejection and supply.

PNNL has shown that hydrothermal liquefaction can treat a wide range of feedstocks and can produce a bio-oil that required much less upgrading than other ways of deconstructing biomass. We determined that there was value in exploring the use of electrochemical processes for the further conversions need to transform the bio-oil from hydrothermal liquefaction into fuel or higher value chemicals.

Examples from our own work and from the literature showed that electrochemical reduction offered adequate reaction rates and at least some of the necessary generality. In addition, an electrochemical reactor necessarily both oxidizes (removes electrons) on one side of the reactor and reduces (adds electrons) on the other side. Therefore, the two compartments could, in principle, be coupled to upgrade the bio-oil and polish the water that was employed as a reactant and a carrier in the upstream hydrothermal liquefaction.

This project was commissioned to devise reactors that could perform the required electrocatalysis by providing effective contact between the oil-water emulsions, the solid catalysts, and the electrode surfaces.

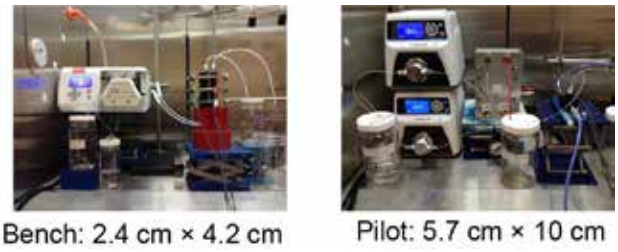
In the first year of the project, we demonstrated that electrolysis cells, derived from work at PNNL on redox flow batteries, could be adapted to this new application. We have constructed two sizes of cells and have shown that their performance accorded with the targeted flow regimes and reaction kinetics.

The carbon felts that we recommended as the catalyst supports proved to be very effective static mixers, distributing the inlet stream across the width of the reactor to give nearly plug-flow residence time distributions, for both scale of reactors.

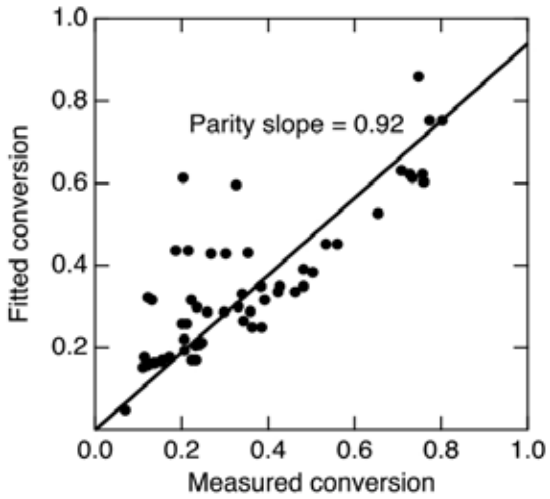
We also showed how to equip the electrolysis reactors with a reference electrode, so the results in the flow systems could be correlated with those obtained in the batch reactors used more conventionally by the electrochemistry community.

Finally, we have constructed a very preliminary techno-economic model, which indicates that an electrolysis reactor should cost only about 15–25% of the total target cost of a waste-to-energy system that could handle the envisaged 1–100 barrels of oil per day.

We are now fabricating a third, larger reactor to permit measuring transport rates and kinetics rates at a scale that could be readily numbered up to contend with pilot-scale deployment. We will then have a gamut of reactors extending from bench to production scales, along with the quantitative models that will permit us to accurately estimate performance and economics of chemical transformations under varying scenarios.



Reactors constructed and tested during the first year of this project.



Parity plot showing that reaction rates determined in the flow reactors could be accommodated by Butler-Volmer kinetics.

PN17023/2913

Scalable Processing of Nanostructured Materials

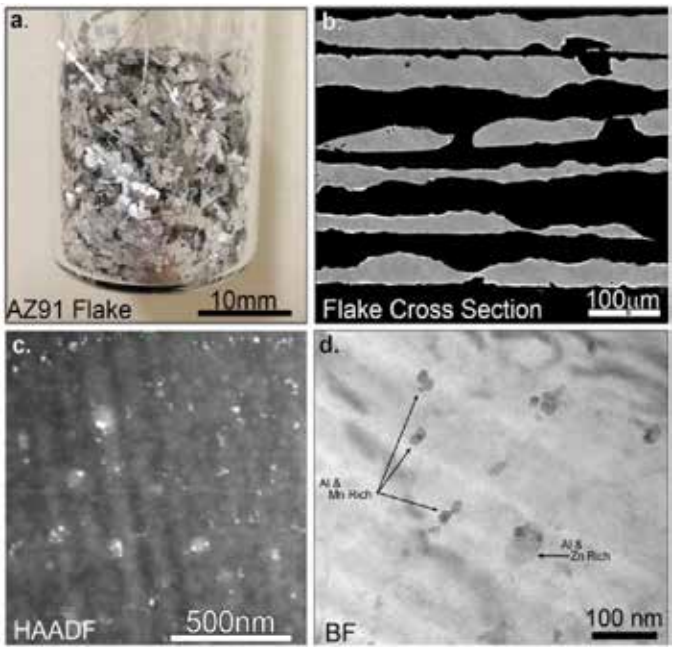
Nicole R. Overman

We are developing novel and scalable processing methods to fabricate unique structural and magnetic nanocomposites. This work is being performed in an effort to advance high-throughput manufacturing technologies and design new lightweight structural alloys and soft-magnetic materials.

Prior research on far-from-equilibrium alloy development has revealed several key factors that can have profound effects on improving material properties such as strength, toughness, corrosion resistance, and in some cases, magnetic properties. The basis for these advances has relied heavily on microstructural control and refinement. The underlying science behind them boils down to an ability to impart maximum disorder into the alloy at a microscopic level. The difficulty with this approach is in stabilizing the disordered or far-from-equilibrium material enough so that it can be retained, despite extreme environments encountered during processing and in service.

The current research is part of a strategic thrust area focused on production and stabilization of unique nanocomposite materials that can only be produced using specialized processing methodologies. This work bridges the research efforts of two sister projects under the nanocomposites thrust area, focused on 1) modeling thermally stable far-from-equilibrium alloy compositions and 2) severe shear deformation consolidation approaches through the production of far-from-equilibrium precursor materials.

The key goals of this project are first to provide rapidly solidified (RS) material for consolidation testing and, secondly, to understand and identify key processing parameters that can be exploited to produce and homogenize these novel precursor materials on the nanoscale level. Processing regimes under investigation target two primary fabrication approaches: melt spinning and ball milling. Using this multifaceted approach, we are able to take advantage of both thermodynamic and kinetic effects to produce precursor materials for consolidation that exhibit novel alloy compositions with enhanced homogeneity, processability, and thermal stability.



As-produced AZ91 RS flake material is shown in (a), along with a cross-sectional view of the flake via scanning electron microscope backscatter imaging (b); transmission electron microscopy images (c, d) show the presence of nanoscale second phases.

Our approach in years one and two has focused on optimizing structural refinement through systematic evaluation of processing parameters for both of these methods.

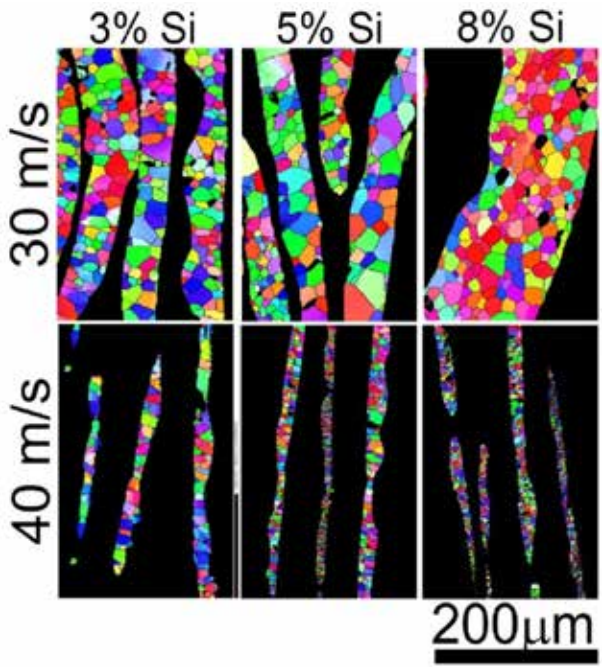
Significant progress has been made in the key material systems targeted. Demonstrator alloys were identified and fabricated based on their ability to serve as game-changing material systems. AZ91, a widely used magnesium-based alloy, was chosen for preliminary process development, as well as Fe-Si (3–9wt %). Both alloys were fabricated using melt spinning. Evaluation of the melt spun precursor material has revealed key fabrication insights, showing fine distributions of nanoscale second phases.

Ball milling research has resulted in the development of generalized curves defining optimal milling time for grain refinement and magnetic properties optimization.

A unique outcome of the work performed has been an increased understanding of the effect wheel speed can have on the as-produced microstructure of melt spun ribbon. We have shown that, in the case of the Fe-Si ribbons, the grain size and structure can be significantly refined when the through thickness width of the ribbon is decreased.

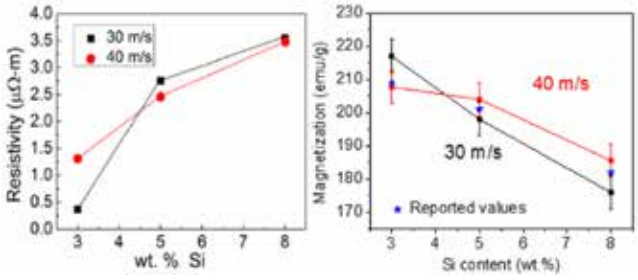
PN16014/2791

This result was found to be highly dependent on the material system under evaluation; the Mg-based ribbons did not show the same behavior.



Electron backscatter diffraction results of ribbon cross sections show grain orientation of AZ91 RS Fe-Si ribbons. Increasing the wheel speed was shown to dramatically refine both the flake thickness and grain structure over a wide variety of compositions.

In Fe-Si systems, we have demonstrated an ability to tailor the as-processed microstructure based on the wheel speed employed during processing. Evaluation of resistivity and magnetization of the Fe-Si ribbons has clarified the impact wheel speed can have on these properties, at varied silicon concentrations.



Resistivity and magnetization trends as a function of increasing silicon content. A cross-over point is seen at 4–4.5 wt. % Si, where wheel speed is observed to have a more dramatic effect on the measured properties at lower Si concentration.

Additionally, Mössbauer spectroscopic studies indicated Si content had a significant influence on the number of Fe sites, relative abundance of the Fe species, and internal magnetic fields/structural environments. This result is significant, because it illustrates the ability to affect changes in the electronic structure of the material.

Preliminary studies on the thermal stability of the melt spun AZ91 ribbons following ball milling have shown precipitation of $Mg_{17}Al_{12}$ equilibrium phases occurs at milling times greater than 20 hours, illustrating a need for thermally stable alloy chemistries.

Specific to FY 2017, production of three thermally stable binary alloys (identified via thermodynamic modeling) has been accomplished for Al-1,5,10 At. % Mg alloys. To date, successful flake processing at scale has resulted in the fabrication of approximately 3 kg of RS Al/Mg-based precursor material, which has supported over 30 consolidation experiments.

FY 2018 efforts will focus primarily on 1) further streamlining of thermal stability assessment in the material systems under investigation, 2) obtaining improvements in the magnetic properties of Fe-Si-based materials through ball milling and heat treatment experiments, and 3) continued precursor fabrication of alloy chemistries identified in sister projects.

PN16014/2791

Stochastic Distribution Control for Complex Networked Traffic Flow Systems

Hong Wang

This project has developed a generic framework for the modelling and control of networked traffic flows using the stochastic distribution control concept. This will provide a foundation for effective tool suites that enable real-time control of traffic flows to achieve smooth traffic flows with minimized energy consumption.

Networked traffic flows are the common transportation structures in cities and heavy traffic areas, where traffic signals at road junctions and sensors distributed in road networks constitute a feedback control mechanism which, together with actual traffic flows that are randomly distributed in nature both in the space and time, form a large-scale, multivariable, stochastic distribution control systems.

Networked traffic flows.

The effective operation and control would not only ensure a smooth traffic flow across the whole concerned area, but also have a significant impact on energy savings and pollution minimization. As a result, modelling and control of traffic flows has been a subject of study for the past decades. In terms of modelling, traffic flow density modelling has been carried out using both the first principle modelling tools and data-driven-based approaches. On the other hand, distributed sensor (e.g., traffic cameras, etc.) placement for the networked traffic flows has also been advanced, and it has been shown that optimized sensor placement can improve traffic flows when they are installed in the area of most uncertainties in urban traffic networks.

In terms of control, intensive research has been done on adaptively or intelligently tuning traffic light switching patterns to cope with traffic volume flows. However, little has been reported on the real-time feedback control of stochastic distributions of whole field networked traffic flows.

A preliminary study has shown that stochastic distribution control is a useful tool that can help to develop distributed traffic flow control, where the main purpose is to ensure that the probability density functions of geographically distributed traffic flows can be made to approach a uniform distribution. This forms the main objective of this study, where the ultimate purpose is to form a preliminary framework to effectively control the traffic signals in a distributed, real-time feedback, and collaborative way across the field road network, so that the traffic flows distribution are made as uniform as possible, while achieving minimized energy consumption and emissions of populated traffic flows.

Traffic queue length model: This task looks into the modelling of traffic flows for a signaled corridor, where traffic queue length models have been established and tested via simulations. This model takes the input as the timing of traffic lights and the output as the queue length.

Signaled corridor traffic flow.

A first principle model has been developed and tested via simulations.

Simulation results for corridor traffic flows, where red, green, and yellow colors stands for the switching on of the red, green, and yellow signal lights, respectively.

Stochastic distribution control – a preliminary framework: Using this model together with the following traffic flow density model,

PN17059/2949

$$\frac{\partial k(x,t)}{\partial t} + \frac{\partial k(x,t)w(x,t)}{\partial x} = Bu(t) + noise(x,t)$$

$$u(t) = \begin{bmatrix} Time & in & Green \\ Time & in & Yellow \\ Time & in & Re\ d \end{bmatrix} \quad (1)$$

a preliminary framework has been established that can realize the real-time feedback control using stochastic distribution theory. Using such a model, a B-spline stochastic distribution model can be obtained

$$\gamma(y,u_k) = a_1\gamma(y,u_{k-1}) + \Lambda + a_n\gamma(y,u_{k-n})$$

$$+ C(y)D_0u_k + C(y)D_1u_{k-1} + \Lambda + C(y)D_mu_{k-m} \quad (2)$$

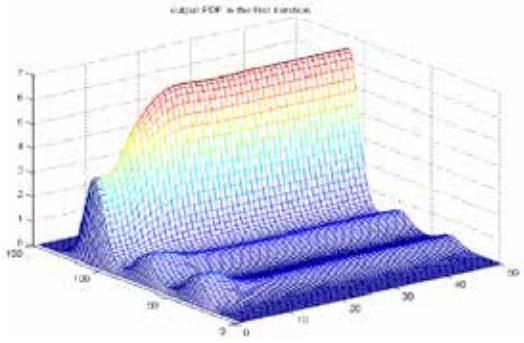
where $\gamma(y,u)$ is the traffic queueing length probability density function at time k per controlled by traffic light timing represented by u_k. By minimizing the following functional distance:

$$J = \int_a^b (\gamma(y,u_k) - g(y))^2 dy + Ru_k^2 \quad (3)$$

the following closed loop control strategy that controls the traffic lights timing can be obtained.

$$u_k = -\frac{\int_a^b \pi(y,k)C(y)D_0 dy}{R + \int_a^b (C(y)D_0)^2 dy} \quad (4)$$

It has been shown that equation (4) constitutes a preliminary closed loop control framework, where for corridor traffic flow, an illustrative response can be shown.



An illustrative response of traffic length distribution.

Virtual Fish for Biological Design of Hydropower Turbines

Marshall C. Richmond

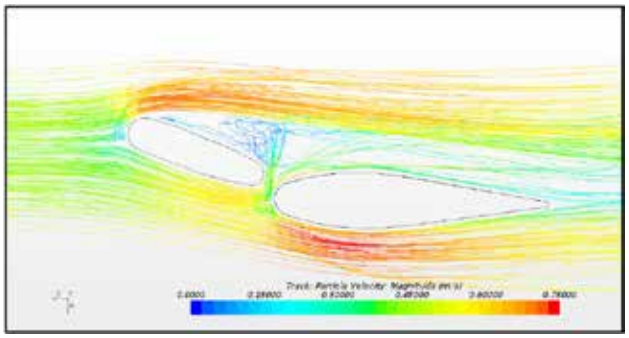
In order to increase hydropower capacity and maintain existing capacity, predictive simulation methods are needed that can reliably estimate the risk to fish passing through turbines, spillways, and other flow passage routes at hydropower facilities. This project addresses this challenge by providing an experimental validation data set for software tools that will support biologically based design, operation, and evaluation of hydroturbines and other passage routes at hydropower facilities before the commencement of expensive new construction, replacement, or repair actions.

A critical challenge currently facing the development of increased hydropower capacity is the need to alleviate, mitigate, or otherwise minimize the adverse environmental impacts of hydropower to levels acceptable to regulatory agencies. This project addressed this challenge by supporting the development of software tools that will aid biologically based design, operation, and evaluation of hydroturbines and other passage routes at hydropower facilities.

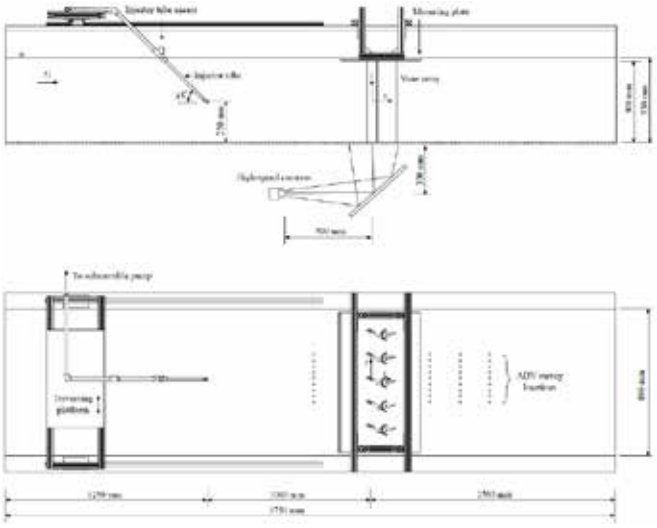
Currently, the biological performance assessment (BioPA) toolset developed at PNNL approximates fish trajectories through a turbine by simulating the trajectories of massless particles using computational fluid dynamics (CFD). This work focused on increasing the physical realism of the BioPA software toolset by including the effects of particle mass (referred to as inertial particles), which are more representative of actual fish, and performing laboratory observations, initially in the presence of simplified geometries and eventually in physical hydraulic models of turbines, to validate the computational methods. This addition of particle mass effects will greatly increase the commercialization potential of the software.

In order to achieve close comparisons between the CFD/BioPA solutions and experimental observations, a simplified flow condition was identified and recreated in the Albrook Hydraulic Laboratory at Washington State University. After the laboratory flow conditions were fully characterized using an acoustic Doppler velocimeter, the CFD solution was solved using identical conditions so that the behavior of the particles around representative turbine geometry can be validated against the laboratory data. Because inertial particle motion in turbulent flow is a stochastic process, the CFD solution was validated using experimental observations of statistical distributions to describe the likelihood of collision occurrences.

The trajectories of two particle shapes (cylindrical and spherical beads that were near neutral buoyancy) were observed experimentally and using CFD for a range of collision geometries, flow speeds, and release locations. The collision geometries used were a single cylinder and an idealized array of stay vanes and wicket gates. The particle passage through the vane arrays were tested at two different wicket gate angles. The trajectories were recorded experimentally using dual high-speed cameras (500 frames/second). A total of 24 tests were performed, with at least 400 particles released for each.



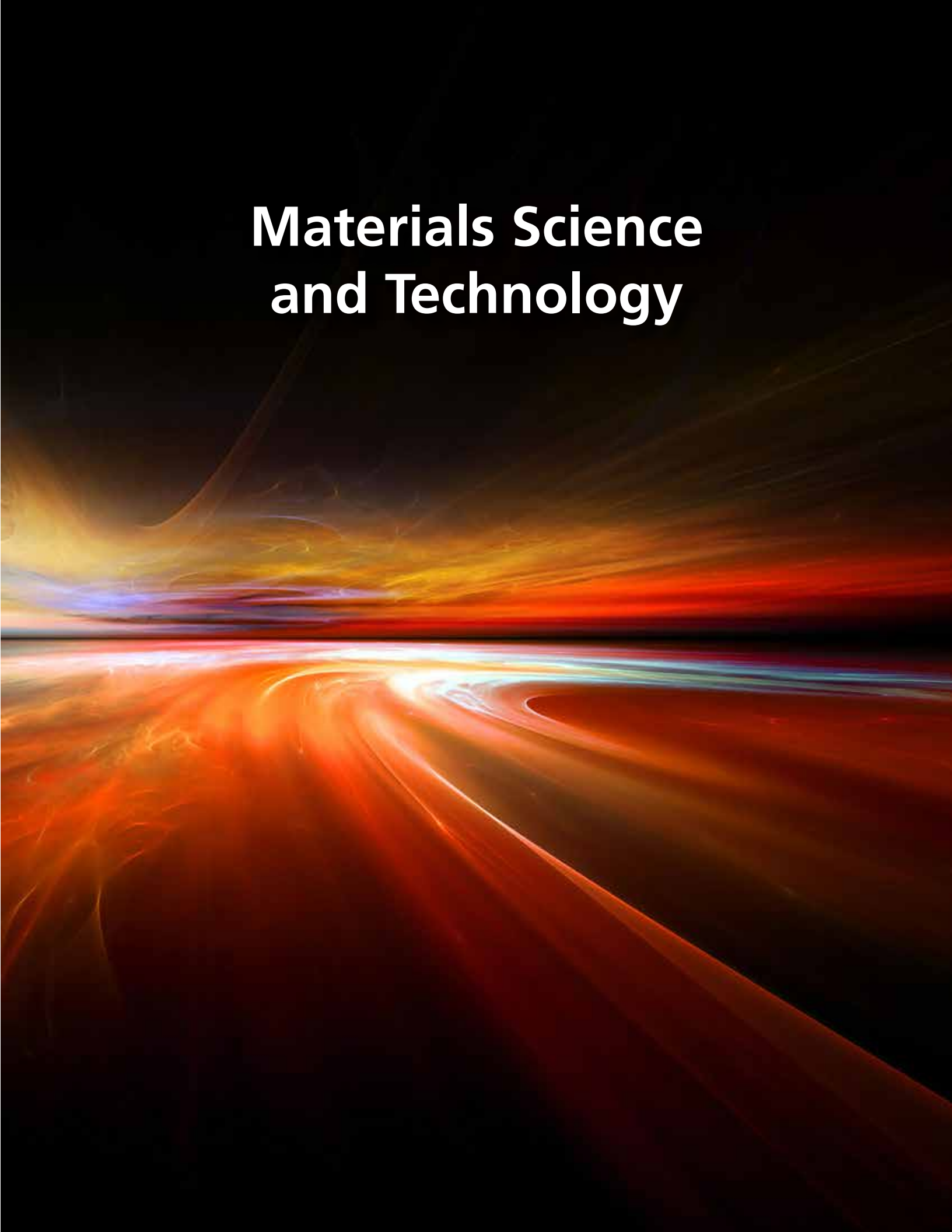
CFD calculations of Virtual Fish bead trajectories around stay vane and wicket gate.



Schematic of test section of recirculating water flume showing bead injector tube and vane array installation.

PN17081/2971

Materials Science and Technology



Bridging Length Scales in Complex Oxides: From Point Defects to Defect Superstructures

Peter V. Sushko

This project provides fundamental understanding of defect formation and interconversion mechanisms in multicomponent materials to advance their functionality by means of controlled synthesis and processing. Our aim is to reveal how cooperative effects that underpin materials functionalities emerge from the properties of individual building blocks.

Development of fundamental understanding of the roles of heterogeneity, interfaces, and disorder in materials systems represents a transformative opportunity to move from ideal materials to the complexity of real systems under realistic conditions. The overall aim of this project is to advance fundamental understanding of how local inhomogeneity (i.e., deviation from an ideal composition and structure) gives rise to cooperative behaviors that underpin materials functions. In particular, we seek to 1) reveal coupling between zero-dimensional (0D) defects and 1D, 2D, and 3D structures; 2) investigate the effects of defect concentration and spatial distribution using theoretical and computational methods; and 3) identify control parameters that enable utilizing this cross-scale coupling in energy-storage and conversion processes.

For the purpose of this project, we selected materials systems and associated phenomena of ever-increasing complexity. Our results provide mechanistic insights into the origin of their functionalities, optimal synthesis, and processing needed to harness their properties and functions in practical applications.

The high electron mobility of $SrTiO_3$ makes it an intriguing candidate for solar applications. However, its usefulness is limited by the large optical band gap (3.2 eV) and propensity to form oxygen vacancies that undermine transport properties. To break through these limitations, we designed a new material: $Sr_{1-x}La_xTi_{1-x}Cr_xO_3$. This work reveals association mechanisms of intentional defects and suggests a path for imprinting 1D wires in thin films.

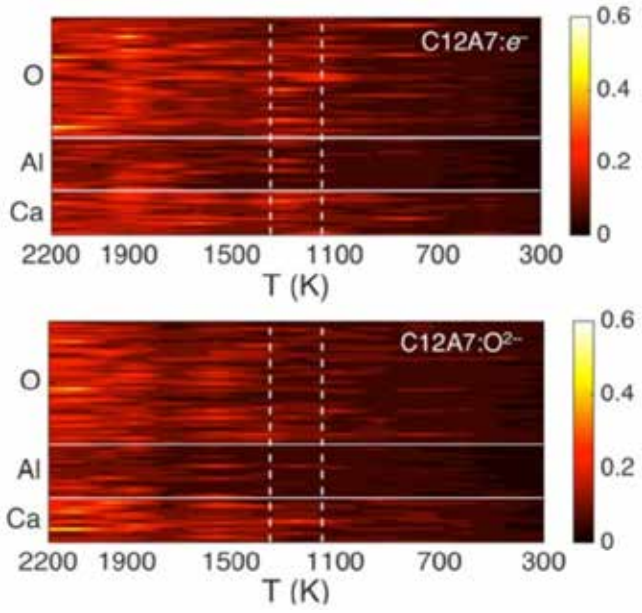
Replacing Li^+ with a bivalent or trivalent ion is a promising route to improving storage capacity and energy density. Mg is an especially attractive alternative to Li due to its large specific capacity, small ionic radius, and high abundance. Transitioning from Li^+ to Mg^{2+} requires new electrode and electrolyte materials. We identified b-SnSb as a promising electrode material and investigated the mechanisms of Mg incorporation and extraction using *ab initio* methods. Together with the experimental data obtained separately, these results allowed us to create a model for b-SnSb electrode evolution and predict optimal electrode structure. These results point to nanostructured Sn electrodes in the form of about 30-nm particles as a potentially optimal Sn electrode for Mg-ion batteries.

Novel approaches to efficient NH_3 synthesis are actively sought in order to enable ammonia production at compact facilities. The key is the development of a high-performance catalyst that enhances dissociation of the N≡N bond. Using *ab initio* simulations, we established an atomic-scale mechanism for how C12A7:e^- promotes electron transfer to N_2 molecules and significantly reduces the dissociation energy of N_2. Further experimental studies demonstrated that our results not only reveal the factors responsible for enhancement of catalytic activity, but also redefine the bottleneck step in NH_3 synthesis. This study also suggested that electrons not associated with any of the lattice ions, often referred to as *electron anions*, can be chemically active depending on their local environment.

We extended the concept of electron anions for catalytic applications to another class of materials: metal hydrides. Using data from our experiments, we proposed a class of materials in which electron anions can be formed without undermining the overall stability of the lattice.

One of the key challenges in materials design is identification of degrees of freedom that ensure robust control of the properties of interest. Small changes in composition of crystalline materials can lead to large changes in their electronic properties such as optical absorption and electrical conductivity.

PN14067/2651



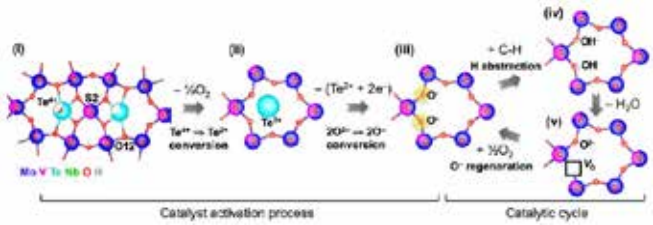
Atomic speed (Å/fs) distributions of individual atoms for C12A7:e^- (top) and C12A7:O^{2-} (bottom) during quenching indicates lower T_g in the electron-rich systems. The mobility of Al atoms ceases by T ~1900 K; Ca and O atoms remain mobile until a substantially lower temperature.

Similarly, a small concentration of (co)dopants in amorphous materials can be used to control their optical properties. However, delicate control of viscosity and glass transition temperature in multicomponent materials (T_g) has remained elusive due to the collective origin of these properties. Substitution of atomic anions with electron anions in materials to form electrides introduces an additional degree of freedom.

We demonstrated that electrons in $[Ca_{12}Al_{14}O_{32}]^{2+} \cdot (2e^-)$ dramatically change dynamics of atoms when this material is transformed to a molten state. Similarly, during cooling, when the melt undergoes a transition to a glass state, the concentration of the electron anions strongly affects the value of T_g. Extending the concept of electron anions to other amorphous materials would provide a powerful instrument for the design of glasses with finely tuned characteristics.

Catalytic function emerges from a synergistic interplay between the catalytic material and its reactive environment. Our modeling of MoVTeNbOx catalyst revealed that its function emerges from a cooperative process, in which 0D defects—oxygen vacancies—form under catalytic conditions and convert Te^{4+} species located in 1D lattice channels into Te^{2+} ions.

This realization of a 0D-1D-2D cooperative lattice effect demonstrates the inter-related nature of processes in complex oxides and opens new strategies for designing highly active oxidation catalysts. The mechanisms of the catalyst activation and of the catalytic cycle provide a path forward for testing our theoretical model and design of new catalysts.



Schematic of the cooperative activation step and catalytic function in $MoVTeNbO_x$. (i) Thermal treatment in a reducing environment induces Te^{4+} to Te^{2+} conversion. (ii) Emission of Te^0 results in the formation of chemically active O^- species. (iii) In the presence of alkane molecules, these O^- species abstract H atoms and form surface OH^-. (iv) The OH^- groups recombine and leave the surface. (v) Interaction of the resulting vacancies with O_2 regenerates O^-, providing continuous activation.

We concluded this project with an attempt to create a model for aggregation and ordering of oxygen vacancies into quasi-two-dimensional planes observed in $SrCrO_{2.8}$, which forms from $SrCrO_3$ upon its anneal in vacuum. In spite of recent advances in *in situ* characterization of the materials transformations, this transition occurs too fast to capture with conventional tools. Yet, understanding of the atomic-scale pathways of this transformation is a key to potentially controlling and utilizing properties of multiple metastable states in $SrCrO_{3-x}$. Results from our experiments suggest that it is possible to stabilize quasi-1D vacancy aggregates by disrupting pathways for their transport through a judicious use of dopants.

Having the ability to exercise such a detailed control of complex arrangements of point defects is paramount to the science of synthesis of atomically precise materials that exhibit new electronic states and structural motifs and, in turn, can enable new types of devices, sensors, and catalysts.

PN14067/2651

Bulk Nanostructured Alloy Optimization: Designing for Processing and Thermal Stability

Aashish Rohatgi

This project will help develop the next generation of advanced materials for fuel-efficient transportation and high-efficiency power generation systems. These advanced materials possess a specially designed nanostructure to counter undesirable microstructural changes that otherwise limit the use of traditional metals and alloys at high temperatures.

Nano-sized microstructural features can produce significant improvement in structural and functional properties of materials. For example, nano-sized precipitates, present within a matrix of micron-scale grains, act as obstacles to dislocation motion and form a key basis of strengthening in a wide range of engineering materials such as steels, super-alloys, aluminum (Al) and magnesium (Mg) alloys, etc. Even greater strengthening is predicted by the Hall-Petch relation if the grain size itself is also reduced to within a certain nanometer (nm) size range.

In soft magnet materials, both theory and experiments show that a composite of soft iron (Fe) grains (approximately 10 nm) embedded in a ferromagnetic amorphous matrix with approximately 2 to 3 nm separation will give the best magnetic performance. Unfortunately, nanostructures are inherently unstable due to the high density of interfaces and, driven by the need to lower the interfacial energy (the driving force), coarsen rapidly even for temperatures moderately higher than room temperature.

Traditionally, coarsening is inhibited by a "kinetic" approach, where grain boundaries are pinned by nano-sized precipitates. However, this approach does not lower the interfacial free energy, but only lowers the grain growth rate and can be overcome at higher temperatures and/or at longer times. In order to reduce/eliminate the driving force, several research groups have proposed a "thermodynamic" approach, wherein segregation of solute atoms to the grain boundaries is used to lower the interfacial free energy, thereby preventing coarsening of the nanostructure and leading to a better thermal stability behavior.

In this project, the thermodynamic approach will be adopted and existing analytical models will be used to identify thermally stable alloys. We will develop high-throughput experimental methodology to evaluate the thermal stability of these alloys and use atomistic modeling techniques for a fundamental understanding of how solute-atom and grain boundary interactions lead to enhanced thermal stability.

This research will follow an integrated experimental and modeling approach in order to achieve its objectives. We will use existing thermodynamic stabilization models to down-select a list of potential nanostructured alloy candidates that are predicted to be thermally stable based on criteria such as reduction in grain boundary energy, enthalpy of mixing, enthalpy of segregation, etc. Experimentally, we will use magnetron sputtering to create nanostructured thin films of model-predicted Mg-, Al-, and Fe-based compositions. Thermal stability of these nanostructured compositions will be evaluated by subjecting them to elevated temperatures and characterizing the resulting grain growth, solute distribution and diffusion, and precipitate formation. The experiments will be accomplished using analytical techniques such as transmission electron microscopy (TEM), atom probe tomography (APT), X-ray diffraction, etc. Finally, we will use atomistic and molecular dynamics simulations to understand interactions between solute atoms and grain boundaries.

In FY 2016, a lattice Monte Carlo method was used to evaluate thermal stability of nanostructured Al alloys, and it predicted an Al-Mg system to be a potential candidate. Accordingly, thin films of Al-Mg alloys were fabricated by magnetron sputtering and procedures developed to handle and anneal them, and analyze them by TEM and APT. Additionally, density functional theory calculated formation energies were used to derive pairwise interaction energies of Fe-Fe, silicon (Si)-Si, Si-Fe, Fe-vacancy, and Si-vacancy pairs and evaluate diffusivity of a vacancy and Si atom in Fe lattice.

In FY 2017, nanostructured Al-Mg thin films with different Mg content were evaluated in as-sputtered and annealed conditions using TEM and APT. The results showed that, while both pure Al and Al-Mg films underwent grain growth during annealing, Mg-containing films exhibited reduced grain growth (69%) compared to pure Al films (100%). This reduced

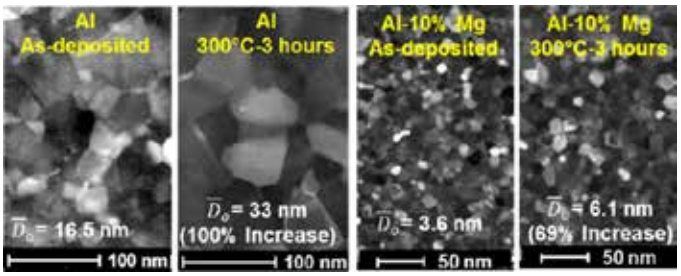
PN16013/2790

grain growth in Al-Mg supports the hypothesis of the effectiveness of Mg additions in increasing the thermal stability of nanostructured Al-Mg alloys.

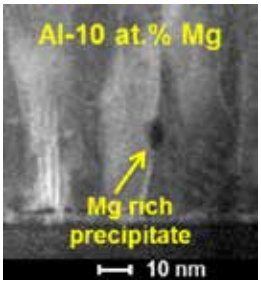
Additional compositional analysis using APT and TEM cross-sectional observations suggested that the reduction of grain coarsening in Al-Mg alloys resulted from the retarding effect of both Mg solutes at grain boundaries (thermodynamic stabilization) and Mg-rich precipitates (kinetic stabilization), with the former playing a major role. Thus, these experimental results successfully validated the computational predictions of the Monte Carlo method from FY 2016.

Additionally, in FY 2017, we used a self-learning kinetic Monte Carlo method to determine the diffusivity of Si and a vacancy in the bcc Fe matrix at various Si concentrations. Local atomic neighborhood activation barriers for vacancy hops (vacancy-atom exchange processes) were calculated on-the-fly using the pairwise interaction energies. Results showed that Si diffusivity is very minimally affected by the magnetic order/disorder state of the bcc Fe matrix. A small increase in the activation barrier from ordered to disordered state is compensated by an increase in the diffusion pre-factor. The results provide further understanding of Si diffusion behavior in α-Fe-Si and offer insight for Si-steel alloy design and processing.

In FY 2018, we will evaluate the thermal stability of nanostructured Mg-based alloy films (e.g., Mg-Zinc [Zn]) in a manner analogous to the nanostructured Al-based films studied in FY 2016–2017. Nanostructured Mg-Zn thin films will be fabricated and analyzed for solute segregation and grain growth kinetics to evaluate the thermal stabilization potential of Zn solute.



Plan view TEM images of nanostructured Al and Al-10 atomic (at.) % Mg in as-deposited and annealed (300°C for 3 hours) conditions, showing lower level of grain growth in Al-Mg compared to pure Al. $\bar{D}_0$ indicates average grain size.



TEM cross section image of the as-sputtered Al-10 at. % Mg film showing an example of ~10-nm size Mg-rich precipitate along the grain boundary.

Carbon Rods with Unexpected Humidity-Driven Water Expulsion

David J. Heldebrant

The goal of this project is to provide the scientific principles critical for understanding controlled nucleation without growth. Such knowledge will push the boundaries of scalable solution synthesis of nanocomposite particles with precise control over particle size, shape, and morphology.

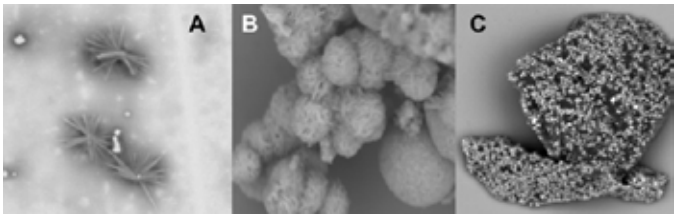
This project focuses on the Predictive Science of Synthesis, with the aim of understanding the structure and interaction of water at interfaces of nanomaterials to control the adsorption/desorption and transport of water. Studies in FY 2017 were to characterize the growth mechanism and surface chemistry of the rods. Synthetic efforts included studying how purity and concentration of reagents impacted the rod's geometry and expulsion behavior to narrow down the controlling factors of the geometry on material behavior. The goal was also to identify optimal synthesis conditions to enable a larger scale synthesis of rods to be used in larger scale testing.

We also studied the surface chemistry and depth profiling of the rods to decipher how the surface chemistry and geometry of the rods interact with water and impact the onset of evaporation (desorption). We continued the use of advanced microscopy and spectroscopic techniques to characterize the interfacial and surface chemistry before and after exposure to water.

Various reaction conditions were investigated to determine the optimum conditions for the formation of rods that have solvent cavitation behavior. We had previously investigated the effect of the $FeCl_3$, its concentration and its purity. In general, all $FeCl_3$ sources afforded rods that were approximately 20 micron in length and 100 nm in diameter. Small differences were observed from batch to batch. During FY 2017, we focused on the effect of switchable ionic liquid (SWIL) purity.

We investigated the effect of SWIL purity to determine its effect on rod formation. Vials containing the different CS_2-SWILS (same composition, but their quality is different because of their bench life) were heated in a sealed autoclave. Interestingly, the age of CS_2-SWIL has a pronounced effect on rod formation. Freshly prepared SWIL always resulted in the formation of rods, irrespective of the $FeCl_3$ source, although some minor difference was observed in batch to batch samples. On the other hand, when we used aged CS_2-SWIL, carbon rods formation was not observed, even after extensive heating. When the aged CS_2-SWIL was used, a sticky, dark, polymeric-type material, which was not rods, was observed. These experiments corroborate our previous result that CS_2-SWIL plays a major role in the formation of carbon rods.

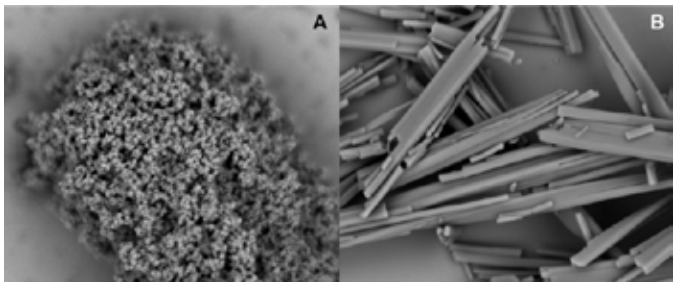
Although carbon-based rods synthesized from fresh CS_2-SWIL can adsorb water at low humidity and spontaneously expel about half of the adsorbed water when the relative humidity exceeds a 50–80% threshold, the overall water adsorption capacity was low. In an aim to increase the overall water adsorption capacity in these rods, we investigated changing the CS_2-SWIL. Scanning electron microscopy (SEM) images confirm the formation of hierarchical star-like aggregates with rod-type structures being its components with one variant. When the reaction mixtures were heated at a higher temperature, formation of porous spheres was observed, and no rods formation was observed. Another variant leads to the formation of uniform MnS nanocrystals in a glassy carbon shell with no rods formation observed.



SEM image (A) hierarchical star-like aggregates formed when trimethylglycine (TMG) base was used instead of 1,8-Diazabicyclo[5.4.0]undec-7-ene (DBU) at 1,050°C; (B) porous spheres formed at 160°C; (C) formation of 50 nm MnS nanocrystals in a glassy carbon shell.

We also studied the effect of synthesis temperature when freshly synthesized CS_2-SWIL was used. The mixtures of CS_2-SWIL and $FeCl_3$ were heated in an autoclave at different temperatures. The temperature has a pronounced effect on rod formation.

PN16020/2797



Iron rich carbon-based materials synthesized at different temperatures: (A) at 105°C and (B) 160°C.

In an aim to understand the mechanism of water expulsion ambient pressure X-ray photoelectron spectroscopy (XPS) measurements were obtained on hierarchical star-like aggregates to determine the surface composition and chemical state of the constituents in presence and absence of humidity and characterized in ultra-high vacuum. Very small changes were observed, and more experiments are needed to confirm the changes in surface composition and chemical state of the constituents.

Studies are underway to understand the effect of rod morphology and surface chemistry on the unique water expulsion behavior. Although we have hypothesized that this behavior may be attributed to confinement mediated solvent cavitation, we do not fully understand the mechanisms for water expulsion in these systems. In recent work, we observed that freshly synthesized rods using fresh CS_2-SWIL exhibit water expulsion behavior. We now believe that the age and temperature of CS_2-SWIL are responsible for such behavior.

We hypothesized that the surface composition of the carbon nanorods drives the water expulsion and adsorption. In order to confirm this hypothesis, secondary ion mass spectrometry (SIMS) depth profiling was used to analyze the surface and bulk compositions in the carbon nanorods with and without good water expulsion properties.

SIMS spectral analysis based on depth profiling shows that compositions are significantly different at the surface and deep into the carbon nanorods. The data suggest the carbon nanorods surface color change may be related to the water repelling and uptake properties; that DBU potentially is a main component of the nanorods structure; and that the structure and composition of carbon related bonds are different at the surface and deep into the nanorods affecting the surface properties of these rods.

In FY 2018, we plan to finish the carbon nanorods data analysis and report findings in a paper.

PN16020/2797

Computational Methods to Rapidly Design Glasses with Targeted Properties for Application to Nuclear Waste Vitrification

David K. Peeler

This project developed and tested a mathematical framework for integrated computational materials design of nuclear waste glasses with uncertainty propagation. This expandable and modular code will allow for state-of-the-art scientific waste form development.

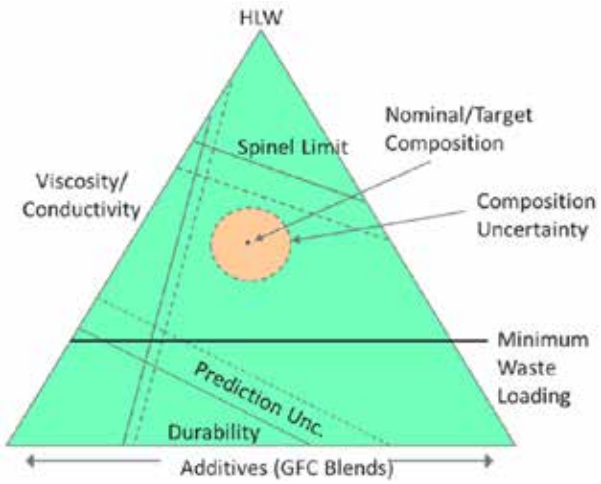
Currently, waste glasses are designed using empirical and semi-empirical models in order to predict glass properties and property uncertainties. Moving away from these empirical models and towards integrated computational materials design (ICMD) will allow for more rapid improvements in the composition design and make it easier to satisfy tailored glass properties. ICMDs provide opportunities for breakthroughs in glass formulation, including treatment efficiency, integration of additional waste streams, and enhanced glass properties, which are too complicated for the current method.

Our objective was to develop and test a concept tool for ICMD of nuclear waste glasses in order to enable breakthroughs in glass formation that are not possible using current techniques. Implementing ICMDs will help to expand possible waste glass compositions and result in more enhanced glasses with improved properties and efficiency.

The ICMD code was written in MATLAB™. MATLAB provides powerful software tools that allow for quick optimization of waste glass and calculation of glass properties in an expandable modular environment. MATLAB integrates well with Microsoft Excel™, through Simulink, and allows the user to run programs directly from Excel. Additionally, reading information from Excel is a simple process, as is writing information to Excel.

Existing, semi-empirical models and uncertainty descriptions were used as a test case in the ICMD code and used for validation of the calculations. Optimal glasses were designed for an example blended high-level waste stream and compared to the previously designed glass showing improved maximum allowable waste loading and process flexibility.

We plan to add advanced scientifically developed glass models to the ICMD to improve model prediction and glass design efficiency. The software will be verified and validated for use under appropriate nuclear quality assurance programs.



Nominal schematic for glass formulation optimization showing processing envelope (central green polyhedral) surrounded by property limits (solid grey lines) with prediction uncertainties (dashed grey lines) and nominal composition with composition uncertainties (tan circle).

PN17012/2902

Developing High Energy Cathode Material for Near-to-Market Advanced Li-Ion Batteries

Jie Xiao

This project has systematically studied the synthesis-structure-performance relationship for high energy density Ni-rich cathode material $LiNi_{0.8}Mn_{0.1}Co_{0.1}O_2$ (NMC811). A cost-effective approach has been identified to synthesize NMC811 cathode material that has great potential to be used in next-generation, high energy density battery technologies.

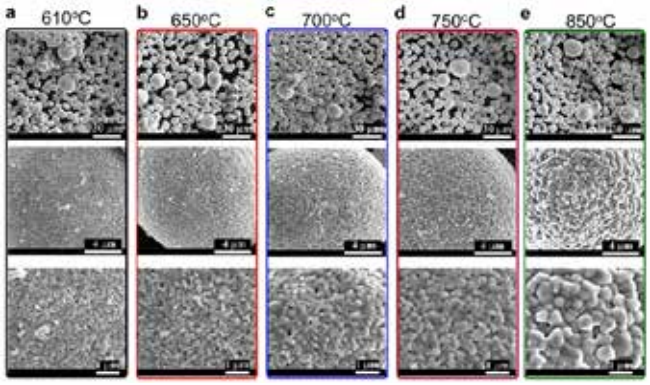
With the state of the art lithium ion batteries (LIBs) used in automobile industry, the driving range of electric vehicles is still far behind those of conventional vehicles using internal combustion engines. Therefore, there is an urgent need to further improve both energy density and power capability of LIBs, to improve the driving range and decrease the recharging time. A lot of effort has been made in the research of high energy density electrode materials, particularly the cathode materials, which are still a limiting factor on the total energy density of LIBs. This project aims to develop a Ni-rich layered cathode material (i.e., NMC811) that is able to deliver a high discharge capacity, representing a significant improvement in energy density compared to traditional $LiCoO_2$.

The synthetic conditions, such as calcination temperature, calcination atmosphere (e.g., O_2, air, etc.), and heating/cooling rates, play an important role in manipulating the material structural properties and morphologies and, thus, influencing the electrochemical performances of as-prepared Ni-rich cathode materials. This project carefully investigated the synthesis route of NMC811 and identified a simple approach to prepare high-performance NMC811.

During FY 2017, the materials synthesis has been focused on a co-precipitation method, followed by high temperature calcination. The heating temperatures are found to largely affect the material layered structural property, particle size, and, thus, electrochemical performances of NMC811 cathode. Scanning electron microscopy (SEM) images indicate that, while secondary particles remain almost unchanged, the higher calcination temperature, the larger the primary particles. The primary particles also change from rod-like to cubic structures as temperature increases. The electrochemical performances of the NMC811 prepared at different temperatures further reveal that the primary particle size and the layered structure characteristic of NMC811 play a critical role on the material long-term cycling stability. The optimal calcination temperature for synthesizing the NMC811 is identified to be approximately 650–700°C.

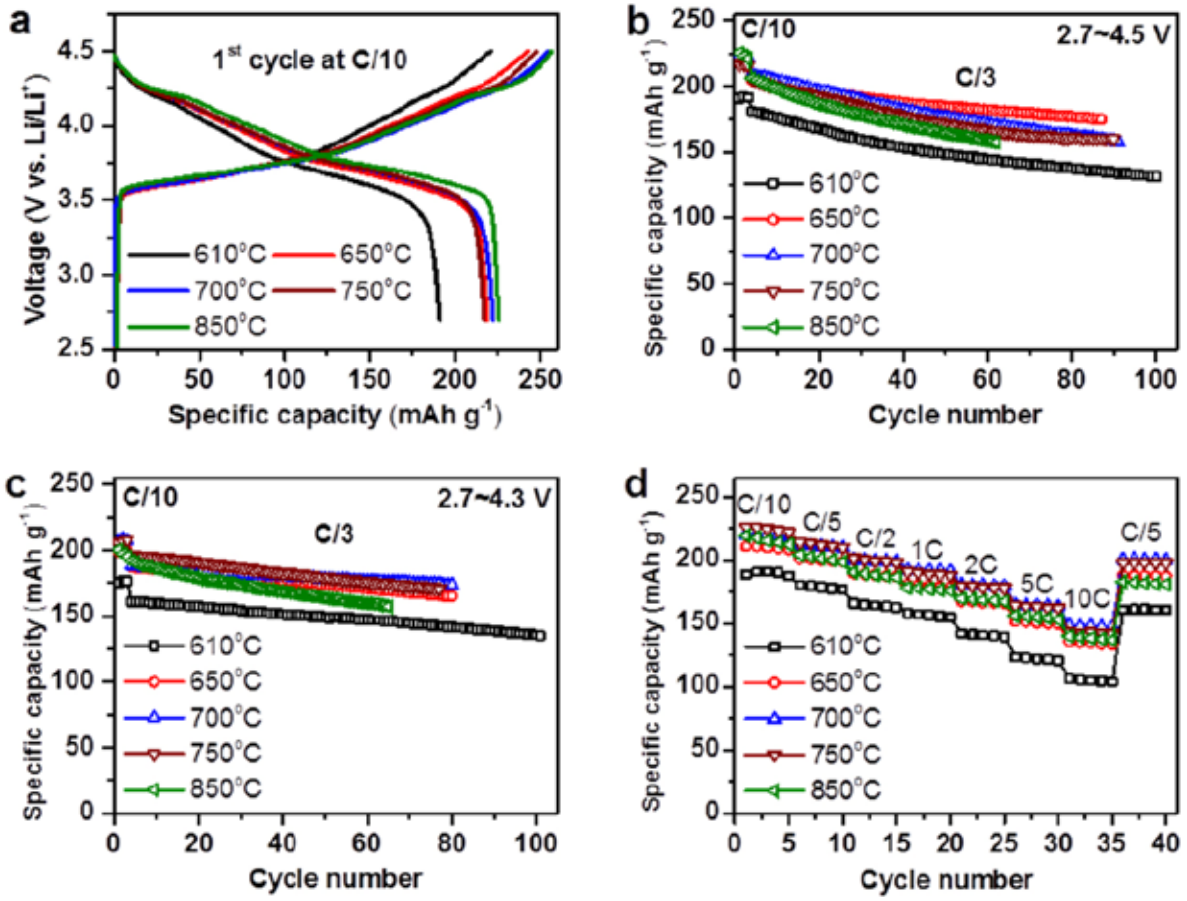
More importantly, the technology for large-scale synthesis of $NixMnyCoz(OH)_2$ precursor using a continuously stirred tank reactor has also been developed to facilitate the next stage of materials development and collaboration with industry.

In FY 2018, we aim to scale up the synthesis of NMC811 materials and begin to study the storage issues of NMC811 through surface/interface modifications. Despite its high energy density, Ni-rich NMC811 is moisture sensitive, which brings concern about material storage, transportation, and operation in ambient conditions. Meanwhile, high-Ni content in NMC811 catalyzes electrolyte decomposition, especially at high voltages, affecting the long-term cycle life. A surface coating and infusion technology will be adopted to enhance both the interfacial stability and structural integrity of NMC811. We will also evaluate the optimized NMC811 in pouch-type cells, with the aim of achieving a prototype pouch cell with high energy density.



SEM images of the $LiNi_{0.8}Mn_{0.1}Co_{0.1}O_2$ cathodes prepared at different temperatures.

PN17097/2987



Cycling performance and rate capability of the $LiNi_{0.8}Mn_{0.1}Co_{0.1}O_2$ cathodes prepared at different calcination temperatures.

Developing *In Situ* Capabilities for Interfacial Characterization Using Synchrotron Light Source

Theva Thevuthasan

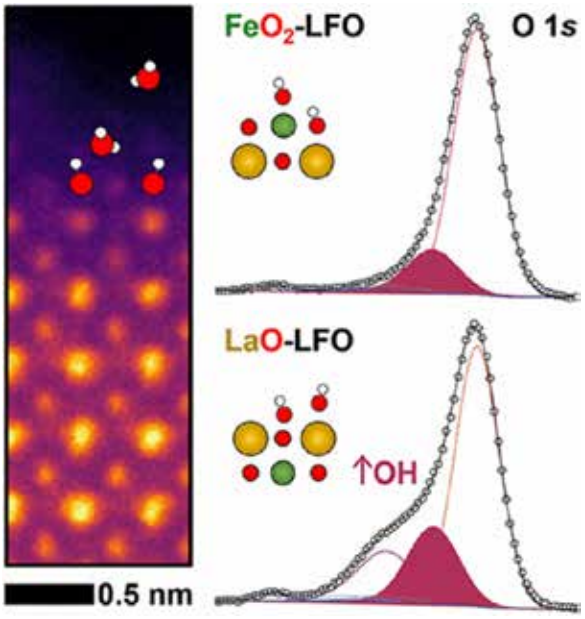
This project develops capabilities that can be utilized to observe material and chemical transformations under realistic conditions.

Developing *in situ* capabilities for synchrotron measurements can help address the science behind interfacial processes. In particular, this work will combine PNNL's core scientific expertise in nuclear magnetic resonance, transmission electron microscopy (TEM), and atom probe tomography with synchrotron light source capabilities available at the Advanced Light Source (ALS) at Lawrence Berkeley National Laboratory.

The main objective of the research is to engage synchrotron light source capabilities to address critical mission areas at PNNL. Our initial efforts focus on 1) understanding redox pathways and coordination geometries in iron oxide and aluminum (oxy) hydroxide materials using *in situ* X-ray absorption capabilities; 2) studying well-ordered oxide surface interactions with water vapor and understanding photochemistry on these surfaces through *in situ* ambient pressure (AP) X-ray photoelectron spectroscopy (XPS) experiments; 3) understanding the chemical environment and coordination geometries of elements in materials relevant to chemical conversions and energy storage under operating conditions using scanning transmission X-ray microscopy (STXM), X-ray absorption near edge structure (XANES), and extended X-ray absorption fine structure (EXAFS); and 4) determining the chemical and structural properties of new materials synthesized at PNNL using STXM, XANES, and EXAFS.

In our AP-XPS work, we synthesized (001)-oriented epitaxial films of LFO distinctly terminated on (LaO)+ or (FeO_2)− planes to probe the influence of perovskite surface termination on chemical reactivity. We have demonstrated that LaO terminated LFO (LaO-LFO) is more reactive toward water, forming hydroxyl species at lower relative humidities than its FeO_2-terminated (FeO_2-LFO) counterpart, consistent with density functional theory (DFT) calculations that indicate a greater stability of hydroxylated LaO-LFO.

In the *in situ* XANES and EXAFS work, we have utilized the gaseous *in situ* cell located on beam line 6.3.1 at ALS to carry out various experiments associated with correlating chemical and topographical arrangements of copper sites in oxide-based



TEM micrograph of the film is presented, along with ambient pressure XPS O 1s scans from water adsorbed FeO_2 and LaO terminated surfaces of LFO thin film.

catalysts to the reaction efficiency by 1) identifying the chemical and bonding nature of substituents with host material by *in situ* XANES and validating with time-dependent-DFT-derived spectroscopic signatures, and 2) studying local structural evolution of active metal sites through XANES and EXAFS under *in situ* reaction conditions and at varying copper concentrations.

In addition, we have carried out Al and O K-edge X-ray absorption spectroscopy, along with Fe and Cr L-edge studies using the solution *in situ* cell capabilities at beam line 6.3.1. Characteristic features in X-ray absorption spectra are expected to provide unique geometrical information on Al in tetrahedral (undistorted and distorted), square planar, and octahedral coordination environments.

During these measurements, we have identified the following challenges: 1) the use of 1-mm x 1-mm Si_3N_4 windows with 500-nm thickness, typically used with *in situ* cells, inhibits X-ray intensity, 2) the current beam size is too large for the windows, as only 20% of the light is utilized (thus decreasing sensitivity), and 3) we are unable to detect signals from small amount of elements, possibly limiting the utility in studies of catalytic systems.

PN17063/2953

Development of Hierarchical Porous Structured Materials for Energy Storage Applications

Luis Estevez

In this project, we are developing hierarchically porous carbon materials with tailorable properties. This ability to tune the morphology (down to the atomic level) will allow us to target the particular prerequisites of certain electrodes for energy storage applications—specifically, lithium ion battery (LIB) anodes and supercapacitor electrodes.

The United States transportation sector is overwhelmingly powered by the internal combustion engine (ICE). This results in roughly a quarter of our total energy use being in the form of petroleum. Not only does the use of this fossil fuel have detrimental environmental effects, but it is, coincidently, similar to the amount of petroleum we import as a country every year. Thus, there has been a great demand for a move towards an electric vehicle (EV) powered fleet, particularly as petroleum becomes scarcer. A major impediment to this goal is the fact that the current energy storage technology of choice for EVs, the LIB, is roughly an order of magnitude away in terms of energy density and at least that in terms of power density.

In order to bridge these gaps, this project will advance two distinct energy storage technologies that can be utilized in an EV (potentially even together in a hybrid EV): supercapacitors (SCs) and silicon-based LIBs.

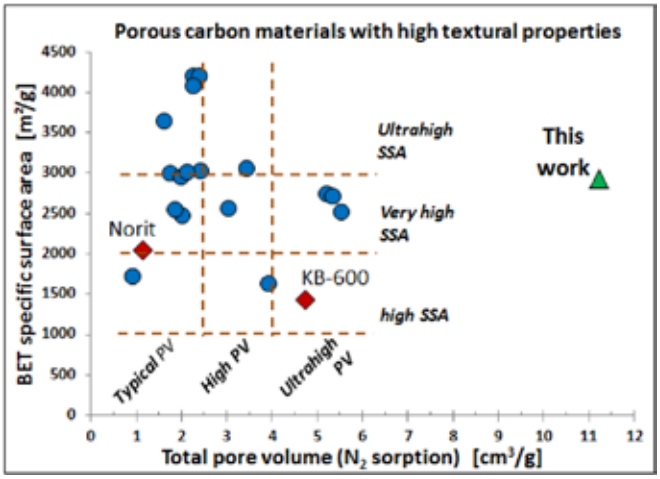
SCs are one of the few energy storage technologies that can compete with the very high power densities present in an ICE (typically about 1–10 kW/kg for both systems). Unfortunately, the most common SCs, electric double layer capacitors (EDLCs), can usually only muster energy densities at least two orders of magnitude below ICEs (typically 1 to 10 Wh/kg for SCs and more than 1,000 Wh/kg for an ICE).

Since the mechanism for energy storage in an EDLC is the adsorption of the ions present in the electrolyte onto the electrode surface, high surface area electrodes are typically employed with specific surface area (SSA) values that are typically greater than 1,500 m^2/g. The most common SC electrodes are usually activated carbons (ACs) that are electrically conductive and meet the large surface area prerequisite through an extensive network of small, nanometer-sized (typically less than 5 nm) pores. The problem is that these commercial ACs typically have a ramified porous structure that can give rise to kinetic barriers for faster moving ions, reducing the power density of the SC in order to increase the energy density. This is especially true for non-aqueous electrolytes, such as ionic liquids, where the widened voltage window increases the energy density, but the bulkier ions present can be kinetically stymied by the small micropores present in the high surface area ACs typically utilized for SC electrodes.

In our project, we have synthesized a hierarchically porous carbon (HPC) materials platform with controllable porosity in the three porosity length scales (as designated by International Union of Pure and Applied Chemistry)—micropores (less than 2 nm), mesopores (2–50 nm), and macropores (greater than 50 nm)—by utilizing three different templating mechanisms: CO_2 activation, hard colloidal templating, and ice templating (respectively), all three of which are quite tunable. This has enabled an open structure that will allow the HPCs to be maximized for both power density and energy density by using the interconnected mesoporous network to mitigate kinetic barriers.

This exquisite control of the micro- and mesoporosity has resulted in an HPC material with both high SSA (from micropores) and large amounts of void space, typically measured as pore volume (from mesopores). This morphological control has allowed us to produce HPCs with a combined SSA and pore volume of > 2,500 m^2/g and > 10 cm^3/g—values unheard of in the same material until now. We have extended our morphological control down to the atomic scale, resulting in two distinct types of HPC materials: an HPC with an atomically disordered carbon named HPC-S and an ordered/more graphitic HPC material denoted as HPC-G.

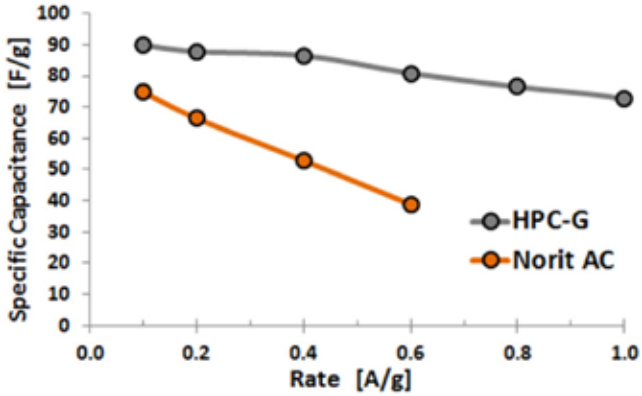
PN15004/2679



Various porous carbon materials from the literature (blue dots) and commercially available (red diamonds) compared with our HPC-G material (green triangle), plotted for surface area (y-axis) vs. pore volume (x-axis).

Our HPC-S materials with such combined high textural characteristics are rare, but similar properties are even more striking in a graphitic HPC material, such as our HPC-G system. The well-ordered graphitic structure around the pores was verified via transmission electron microscopy, revealing the graphitic sheets to be spaced approximately 0.34 nm apart. Bulk characterization methods were also employed, such as Raman spectroscopy, X-ray diffraction, and thermal gravimetric analysis, and further confirmed the graphitic structure.

When utilized as a SC electrode with a solid membrane based ionic liquid (IL) electrolyte, our HPC-G carbon continuously outperformed the high surface area Norit AC. At lower rates, where the kinetic impediments are vastly minimized, our HPC-G electrode was found to have a higher specific capacitance. As the rate increased, the separation between the performances of each widens dramatically, while the Norit AC's capacitance drops by nearly half. Past a certain point, the Norit AC electrodes cannot register a capacitance due to kinetic limitations in the electrolyte, while the HPC-G material continues to demonstrate a good capacitance, revealing the advantage of HPC-type morphology. This type of performance is good when compared to other solid based ILs and even, surprisingly, compares reasonably when contrasted to some liquid based IL systems in the literature.



The specific capacitance as a function of charge/discharge rate for both HPC-G and Norit AC based SC electrodes.

While SCs can help bridge the ICE gap in terms of specific power, employing a silicon-based anode in an LIB is one way to bridge the gap in terms of energy density. Silicon has over 10 times the theoretical capacity of graphite, but the major impediment to its use as a battery electrode is the fact that the Si anode pulverizes and breaks apart during charge/discharge cycling. This pulverization is due to the extreme volume changes during the lithiation process. Researchers have found that using smaller silicon particle sizes can mitigate this pulverization, but the extreme volume change can often physically remove the Si particles from the electrode, quickly diminishing the discharge capacity at every cycle.

To ensure the silicon particles remain bound to the electrode during cycling, we have employed our HPC material as a scaffold for the attachment of Si nanoparticles (NPs). We have utilized a facile method to functionalize the HPC material for covalent attachment to the Si NPs, resulting in an HPC-Si composite anode. We found that it is necessary to add only a small amount of the functionalized HPC material to the HPC-Si composite electrode to dramatically mitigate the loss in capacity that typically occurs in the silicon nanoparticle based anodes. Our results consistently demonstrate the HPC-Si composite anode maintains more of its initial discharge capacity compared to the Si nanoparticle based cells (without the functionalized HPC added) as the cells cycle. Typical results are approximately 75% capacity retention for the HPC-Si cells versus 35% capacity retention for the bare Si cells after 100 cycles.

For this project, we have achieved unprecedented control of the HPC morphology, resulting in an excellent material for supercapacitor electrodes. Our work led to many productive collaborations this year, including publications and a conference presentation.

PN15004/2679

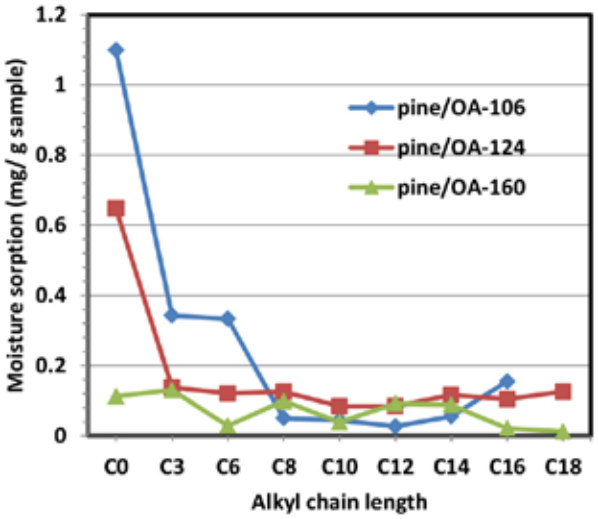
Development of Second Generation Aggregators for Cleanup of Crude Oil in Sea Water: Mixed Fatty Acid-Modified Sawdust Materials

Yongsoon Shin

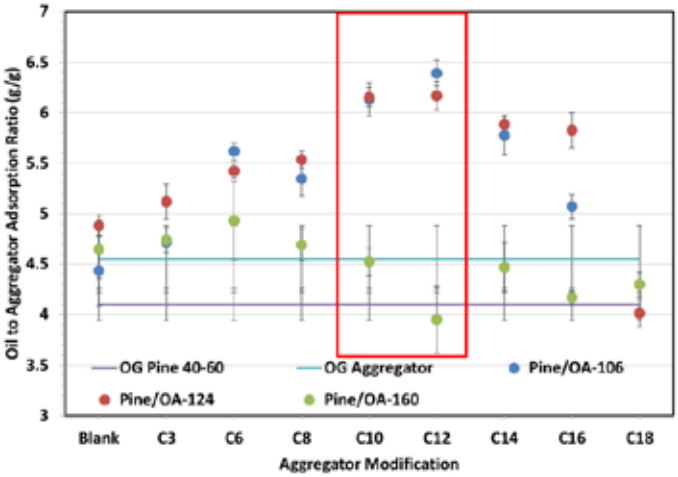
***This project includes the design, synthesis, and application of novel product materials to be used in the treatment of oil spills in aquatic environments. The product entails chemical modification of sawdust or wood flour to render it with hydrophobic, oleophilic, and buoyant properties. The resulting product is herein referred to as an aggregator. The primary purpose of the aggregator is to attract and absorb crude oil to create stable slick of sufficient thickness that it may be burned* in situ.**

The herding agent in products used to treat oil spills in aquatic environments is typically a liquid, and it is sprayed onto the surface of the water as a ring to surround the spilt oil. Current herding agents in use are liquid chemicals, and these suffer from several limitations; for example, they work best with thin (< 3 mm) slicks under calm sea conditions (no waves, < 5 mph wind), function only for an hour before diffusing into the surrounding water, do not work well when ice is present, and may have toxic properties. Wood-flour-modified materials are cheap and show excellent buoyance in waving sea water. Furthermore, their modified surfaces attract spilled crude oils effectively and easily form thick slicks.

First, we synthesized three-base materials, which are oleic acid (OA)-modified pine flour materials with different OA coverages. They are hydrophobic, but they still have uncovered surface hydroxyl groups. Second, fatty acids with different alkyl chain lengths have been introduced to cover surface hydroxyl groups. At this stage, max surface coverage has been reached. All the aggregators have been characterized by moisture sorption, Fourier-transform infrared spectroscopy, X-ray diffraction, thermogravimetric analysis, and tested using a real crude oil. Mixed fatty acid-modified aggregators showed higher sorption behavior than OA-modified samples. Aggregators are stable and recyclable.



Progress of mechanical alloying in Al-Ti-Cr system as a function of mia OA-modified pine aggregators become more hydrophobic when the second fatty acid is introduced.



C12 modification on 106% OA-modified pine aggregator showed best crude oil sorption.

PN17085/2975

Experimental and Theoretical Investigation of Core Level X-ray Photoemission Line Shapes for Unstable Transition Metal Cations

Scott A. Chambers

In this work, we have carried the measurement and theoretical understanding of core level X-ray photoelectron spectra of oxides to a new level. We can now discriminate between transition metal cations with various valences in different lattices based on the detailed shape of the spectrum.

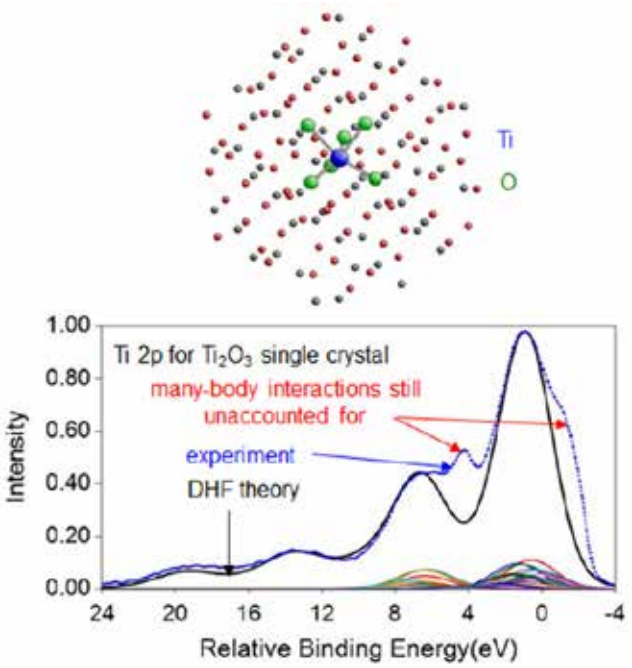
Prior to this project, the Ti 2p and 3p X-ray photoelectron spectra (XPS) of unstable Ti^{3+} had never been measured on a reliable sample with sufficiently high energy resolution to identify the many complex components that are present. Additionally, these core level line shapes had never been calculated with the only fully *ab initio* theoretical method that does not rely on empirical input (other than structural data for the lattice).

We initiated the project to show that this approach could be used successfully to understand these complex line shapes, using Ti^{3+} as a test case. We were successful in terms of demonstrating proof of principle.

This project consisted of experimental measurements and theoretical calculations of the Ti 2p and 3p spectra for structural Ti^{3+} in single-crystal Ti_2O_3. We used hard XPS by sending samples to the Diamond Light Source in the United Kingdom for measurements and utilized relativistic Dirac-Hartree-Fock (DHF) theory for theoretical predictions of the line shape.

The Ti 3p spectrum was fully accounted for using DHF theory, and the Ti 2p spectrum was mostly accounted for, with the exception of features on the low binding energy sides of the $2p_{3/2}$ and $2p_{1/2}$ peaks, which changed in intensity as the material was taken through its semiconductor-to-metal transition at T ~500K. DHF calculations with a single Ti cation in the embedded cluster could not reproduce these features.

We suspect that these shoulders result from screening of the core hole by conduction band electrons, and the screening changes with carrier concentration. This many-body effect cannot be accounted for unless the embedded cluster is expanded to include at least two Ti cations, and doing so is the focus of future work on this topic.



TiO_6 embedded cluster within a Ti_2O_3 array of point charges used to calculate the Ti 2p and 3p spectra for single-crystal Ti_2O_3 (upper). Comparison of measured and calculated Ti 2p spectrum show excellent agreement for all features except those at ~4.5 and -2 eV (lower). Approximately 18,000 distinct final states contribute to the total theoretical spectrum, the most intense of which are clearly visible.

PN17093/2983

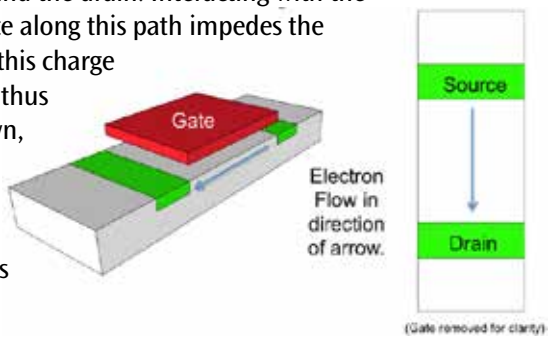
Free Space Transistors for Advancing the Art of Software Defined Radio

Ethan Farquhar

The objective of this proposed work is to collaboratively acquire expertise in the design, fabrication, and use of Free-Space Field-Effect Transistors (FS-FET) for the dual purposes of advancing the science of new nano-technological devices and applying these advances to various national security application spaces such as the further miniaturization of software defined radio systems and harsh-environment sensing systems.

The invention of the transistor, and the subsequent miniaturization and integration of these devices, has driven revolutionary changes to society over the course of the past generation. However, these devices have speed limits that cannot be exceeded. Because of these limitations, researchers have been looking to technologies other than traditional silicon semiconductor processes to develop ever faster and lower-power processing.

In standard complementary metal–oxide–semiconductor (CMOS) transistors, electrons are able to flow from the source terminal to the drain terminal, with the current being modulated by the field from the gate terminal. The electrons, though, are forced to conduct through the material between the source and the drain. Interacting with the crystal lattice along this path impedes the progress of this charge carrier, and thus slows it down, limiting the speed (frequency) at which this transistor can operate.



Normal CMOS transistor topology.

In contrast to standard CMOS, free-space transistors show the promise to run possibly as much as three orders of magnitude faster. This is due to the fact that electrons are ballistically transported through the channel instead of through a resistive channel. Also due to this configuration, there is a potential for greater power gain than with a standard MOSFET.

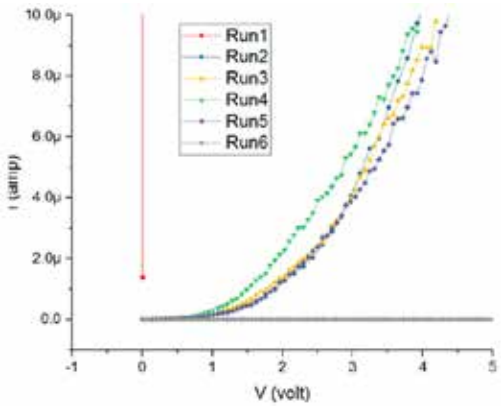
Early devices fabricated and described by others have demonstrated cutoff frequencies for the FS-FETs as high as 450 GHz. This frequency is high enough to give coverage over much of the radio frequency spectrum and is even of relevance in the millimeter wave band (30–300 GHz).

Due to difficulties that we had with the technical limitations of the clean room facility, we started looking for new ways of achieving the nano-gaps that we were attempting to produce. We developed the novel solution to fabricate devices whole (as before), but then, instead of trying to "burn" or ash away the thin part of the device, we are using a helium-ion microscope that had the capability of not only imaging the devices that we were making, but also cutting a small channel in our devices.

This switch proved to be extremely useful and also represents a novel use of the technology and approach to solving our problem. We were able to exceed our goal of producing gaps in the device of around 50 nm by reaching a 25 nm gap.

This led to the development of a new topology for a completely planar transistor. This development yields a device that could lead to incredible density of devices. If a transistor can live completely in a single metal layer of a modern CMOS process, then not only can transistors make use of the X and Y dimensions of the chip, but also in the Z dimension.

To make this planar transistor, we took the original structure that we started with and copied that structure but turned it 90 degrees. What we fabricated resembles a plus sign. Then the He ion microscope was used to make two cuts, giving us two symmetric structures at 90-degree angles. We used one set of the structure as the source and drain. We then use the other two terminals as gates to modulate the current flow.



Drain to source current for device fabricated with 30-nm gap.

These results are promising and demonstrate that it is possible to eject electrons across the source and drain.

PN16028/2805

Fundamental Insights into Gamma-Radiation Effects at Complex Oxide-Water Interfaces from First-Principles Simulations

Sebastien N. Kerisit

This project utilizes state-of-the-art computational techniques to provide molecular-scale insights into the effects of radiation fields on the reactivity of mineral phases relevant to the processing of radioactive waste.

Aluminum is a major component of high-level waste sludge stored in underground tanks at the Hanford Site and is principally present as aluminum oxyhydroxide phases, with gibbsite and boehmite as the most abundant phases.

An important unknown is the effect of the radiation field present inside the tanks on the properties of aluminum-bearing solid phases and, specifically, on their dissolution behavior, as these phases will be dissolved by caustic leaching prior to vitrification. Therefore, this project focuses on boehmite and gibbsite as model systems and aims to develop a fundamental understanding, from first-principles simulations, of the reactivity of these phases in the presence of a gamma-ray field.

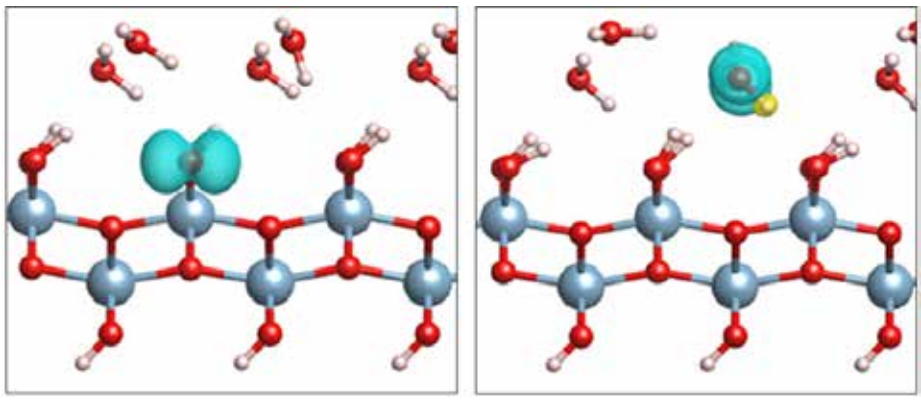
Enhancement of radiolytic production of dihydrogen gas (relative to pure water) has been observed experimentally at boehmite surfaces but not at chemically similar gibbsite. Although this observation offers a clue to their response to radiation, the elementary processes that give rise to this difference in surface reactivity between the two phases have yet to be elucidated.

As radiation particles pass through a solid phase, such as boehmite or gibbsite, they create electronic excitations that result in the formation of free electrons and holes, or carriers. One of the most important properties that control carrier formation is the electronic band gap of the material, the minimum energy cost to excite a valence electron and create a single electronic excitation. Yet, this quantity is unknown for boehmite and gibbsite.

This fiscal year, we performed quantum-mechanical calculations to predict the band gap energy of boehmite and gibbsite and found it to be similar for both materials and close to 8 eV. We confirmed this theoretical result experimentally by electron energy-loss spectroscopy and X-ray photoelectron spectroscopy. Interestingly, previous work showed that a band gap energy of 5 eV maximized the radiolytic activity of oxide particles, as determined from the production of dihydrogen gas. Not only is this value different from that we determined for boehmite and gibbsite, but both materials show very different radiolytic activities despite very similar band gap energies. Our results point to the fact that radiolytic activity is not a simple function of the band gap energy, as previously thought. We, therefore, need to look beyond carrier formation to their diffusion and eventual fate in the materials to understand the radiolytic activity of boehmite and gibbsite.

In aluminum oxyhydroxides, electrons can diffuse as delocalized carriers through the conduction band. Our calculations show that electron mobility through this process is high and direction independent. In contrast, the same calculations indicate that holes couple strongly to lattice vibrations in these materials and can, thus, localize on lattice sites, specifically on oxygen sites. Consequently, their mobility is low and strongly direction-dependent because of the layered nature of the boehmite and gibbsite lattices.

One conclusion of this work is that this stark difference in mobility will result in extensive charge separation and lead



Reactive electronic holes at the boehmite-water interface. Holes localized on surface oxygens (left) can react with water to form adsorbed hydroxide radicals (right).

PN16062/2839

to the accumulation of holes at edge surfaces that are perpendicular to the atomic layers in both lattices. The contrasting radiolytic activities of boehmite and gibbsite might, therefore, originate from differences in the reactivity of their edge surfaces.

Once they reach the surface, our calculations show that carriers can react with adsorbed water molecules to form very reactive species such as hydroxide radicals. As we continue this work into the next fiscal year, we will explore differences in interfacial reactivity between basal and edge surfaces, and between boehmite and gibbsite. Through this new molecular-scale understanding, this project is expected to yield insights into how radiation has affected aluminum oxyhydroxide particles in Hanford waste tanks over the years and also help guide the development of improved models of dihydrogen gas generation.

Fundamental Investigations of Photoelectrochemical Water Splitting of Model Oxide Electrode Surfaces

Kelsey A. Stoerzinger

This project provides fundamental understanding necessary to rationally develop oxide-based heterostructures that effectively absorb the solar spectrum and efficiently utilize solar-generated carriers to generate chemical fuels, such as hydrogen and oxygen gas, from water splitting.

Solar radiation represents an immense and renewable energy resource; however, storage of this intermittent source is critical in enabling grid integration and utilization in remote applications. One approach is utilizing the sun's energy to form chemical fuels through water splitting.

The efficiency in driving the hydrogen and oxygen evolution reactions is governed by that of electron-hole pair generation, propagation to the electrode surface (without recombination), and electrochemical conversion. Metal oxides constitute ideal aqueous interface layers because they are much more stable in solution than traditional semiconductors; however, fundamental understanding of the energetics and processes that occur at the surface is lacking.

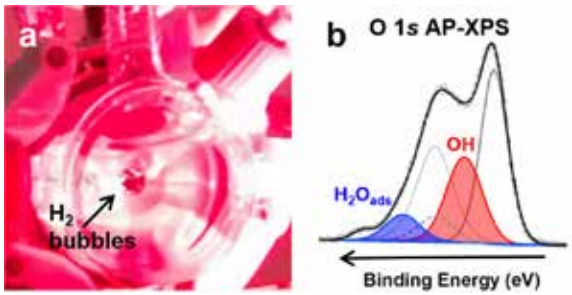
Our goal is to fabricate well-defined oxide-terminated heterostructures using molecular beam epitaxy (MBE) that absorb well in the visible portion of the solar spectrum, have low defect densities (to minimize recombination and carrier trapping), and exhibit band bending and band alignments that facilitate carrier transport to and through the electrode surface to the aqueous medium. The outcome will be an important new set of design principles, along with useful materials, for photoelectrochemical water splitting.

In FY 2017, we have established a testing methodology of oxide thin films and heterostructures thereof in an inert atmosphere linked to materials growth and characterization facilities in the Environmental Molecular Sciences Laboratory, providing unprecedented control over surface preparation and its characterization following photoelectrolysis. Photoelectrochemical (PEC) performance is evaluated in a custom PEC cell that electrically contacts and chemically isolates oxide MBE films for measurement. After PEC measurement using a newly procured electrochemical testing station in a nitrogen box, illuminated by high-power LEDs of select wavelength, the sample can then be directly transferred back into vacuum for characterization of the surface species present, eliminating the potential for contamination from ambient atmosphere. We have demonstrated this methodology for *n*-$SrTiO_3$/*p*-Ge (001) heterostructures to drive the hydrogen evolution reaction assisted by 630 nm (visible) light.

Mechanistic investigations of the water/oxide surface interaction have shed light on the functional groups that form on oxide surfaces *in situ*. Using ambient pressure X-ray photoelectron spectroscopy (AP-XPS) at the Advanced Light Source, we have probed the speciation resulting from interaction of oxides with water and electronic band bending at the surface. The absolute core-level binding energies measured with the surface saturated in water vapor yield valuable information of the effect that the aqueous environment and the associated adsorbates have on the presence and strength of built-in potentials in the near-surface region of the oxide. Likewise, the detailed line shapes, particular those for oxygen, reveal which species are present at the surface.

Studies of the photoanode $LaFeO_3$ terminated with the LaO or FeO_2 layer of the (001)-oriented perovskite crystal structure found notable differences in chemical reactivity with water. The incorporation of Ti into Fe_3O_4 was found to dramatically reduce the propensity to dissociate water and form hydroxyl groups.

In FY 2018, we will fabricate films of tunable band gap, band edge position, and carrier density in order to directly relate the PEC performance to the electronic structure. Parallel measurements with AP-XPS in a humid environment will shed light on the reaction intermediates and their influence on surface band bending, providing *in situ* understanding of the electronic states. *In operando* AP-XPS measurements will elucidate the electrode speciation during operation, further bridging the gap between typical surface science techniques the reaction environment.



a) Photograph *n*-$SrTiO_3$/*p*-Ge photocathode evolving hydrogen bubbles during water splitting, illuminated by a 630-nm LED during cyclic voltammetry. b) Oxygen core-level spectra of n-$SrTiO_3$/*p*-Ge using AP-XPS, probing the surface speciation in a humid environment. As the humidity increases, water dissociates to form hydroxyls (OH) groups and also adsorbs to the surface molecularly (H_2Oads), evident as species of distinct binding energies in XPS.

PN16105/2882

Hierarchical Framework Materials by Advanced Material Design

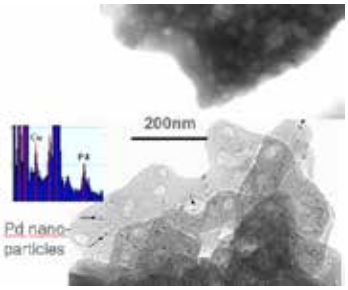
Miroslaw A. Derewinski

The goal of the study is the exact tailoring and crafting of micro- and mesoporous hierarchical materials to facilitate and improve the utilization of porous materials in transformation of larger substrate molecules in the condensed liquid phase, using PNNL's advanced capabilities in in situ spectroscopies and microscopies.

The application of microporous materials such as molecular sieves based on Si and Al in industrial processes is mainly facilitated by the high homogeneity of the active site and large surface area. The accessibility to the active site is a function of the pore size and poses significant challenges in respect to mass transport limitations. Additional challenges, such as good hydrothermal stability of these silica based sieves in hot liquid water, ubiquitous in the liquid phase transformation of biomass derived feedstocks, require smart synthesis design and material control to produce highly efficient catalysts.

To address these challenges, we attempt to first use knowledge-based, bottom-up approaches of synthesizing robust materials with interconnected hierarchical pore systems spanning multiple length scales. This can be done via hard and soft templating approaches or the recrystallization of ordered disordered meso-/macroporous materials. In a second step, we investigate in a top-down approach how changes to the polarity of porous surfaces in micro- and mesoporous structures can influence the selectivity and activity in the catalytic process and stability of the framework.

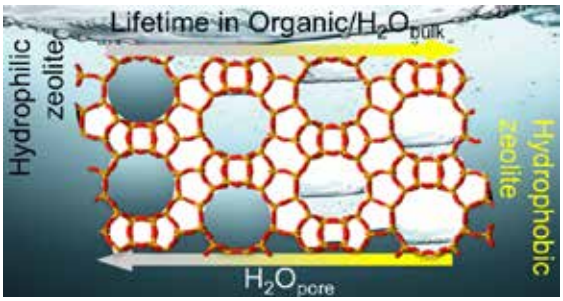
The bottom-up approach of designing hierarchical framework structures was initiated in FY 2017 by synthesizing zeolites of the Mordenite Framework Inverted structure in the presence of carbon porogens. This results in the embedding of the carbon in the porous structure of the zeolite and subsequent calcination generates a mesoscopically structured material with defined pore structure. The prepared hierarchical material could be used as an active support for the deposition of highly dispersed metal clusters trapped inside mesovoids/mesopores.



Pd nanoparticles embedded in nanovoids.

The controlled recrystallization of amorphous meso-/macroporous materials was employed to synthesize hierarchical framework structures containing nanoparticles of zeolite dispersed on the wall of meso-/macropores. Generated strong acid site functions enable converting larger substrate molecules due to the easier accessible pore structure.

In FY 2017, we also achieved insight into the deactivation and degradation of zeolite catalysts during aqueous phase reactions. This was augmented by previous work focusing on the crucial role of structural defects that showed that removal by healing of defects can have significant improvement on zeolite robustness. Similarly, the application of this technique of healing to readily available nanoporous materials comprehensively showed the effectiveness of this treatment for improving catalyst lifetime in aqueous phase reactions. It was shown that it is the hydrophilicity of the zeolite surfaces, dictated by the concentration of acid sites in the framework, which is most crucial for affecting framework collapse. By tuning the polarity of the surfaces, we could significantly stabilize the material. In addition, FY 2017 resulted in the successful synthesis of highly robust and hydrophobic zeolitic materials of low defect concentration by smart synthesis design.



A schematic overview of the effect of surface polarity and intraporous water concentration on zeolite lifetime in aqueous phase reactions.

Modifying the diffusion pathways of molecules in zeolites as a means to separate undesired water from the substrate was attempted by crafting a core-shell material consisting of a hydrophobic shell surrounding a hydrophilic, highly reactive core. It was shown that intrinsic activities were not altered, whereas the resultant lifetime of new material in hot water environment was significantly/largely prolonged.

PN17046/2936

Ion Implantation and Characterization of Epsilon Metal Phase Formation in Ceria

Ram Devanathan

The aim of this project is to develop new understanding of noble metal particle formation in irradiated nuclear fuel by studying a non-activated surrogate material. This work establishes the technical basis for improving the safety and economics of nuclear power to ensure U.S. energy security.

The mechanism of formation of metallic phases and their morphological evolution are key knowledge gaps in our understanding of nuclear fuels. The phase diagram of this metallic system is five-dimensional and the reactor conditions under which the phases form are far from equilibrium. However, our understanding is limited to a small region of phase space under equilibrium conditions.

Efforts are currently underway at PNNL and at external institutions to perform careful irradiation of nuclear fuel under controlled conditions in a test reactor, recover metallic phases that may be present, and characterize their structure and chemistry. Due to the long duration, cost, and radiological concerns associated with fuel irradiation, as well as the complex physical processes in nuclear fuel, the mechanism of metallic phase formation is unknown.

This project takes a novel approach to this problem by combining atomistic simulations of radiation damage with targeted ion bombardment studies of a surrogate material to enhance our understanding of nuclear fuel. This project tests the following hypotheses: 1) defects produced by fission fragment damage enhance the migration of metals, and 2) metals diffuse to grain boundaries, cracks, and pores from surrounding regions. Our work will reveal the conditions under which metallic phases form in irradiated fuel.

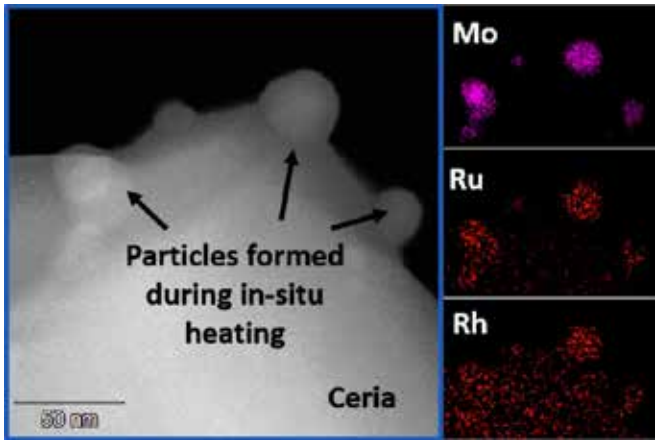
We used pulsed laser deposition to grow CeO_2 films doped with Mo, Re, Ru, Rh, and Pd; irradiated the films with energetic ions at 400ºC; subsequently annealed the material at temperatures up to 1,100ºC; characterized the microstructural changes using electron microscopy, X-ray diffraction, and secondary ion mass spectrometry; and integrated experimental data with molecular dynamics simulations of defect production and evolution under irradiation.

We showed that Pd particles of approximately 3 nm in size form near dislocation edges as a result of 90 keV He+ ion irradiation at 673 K. Subsequent thermal annealing at 1,073 K gave rise to larger nanoparticles containing Mo and Pd at grain boundaries. Further annealing at 1,373 K produced particles composed of all the doped elements with sizes ranging from 50 to 70 nm.

Our significant finding is that precipitation is mainly driven by irradiation-enhanced diffusion. Grain boundaries in CeO_2 serve as both faster diffusion paths and effective sinks for Mo and Pd to precipitate. Precipitation is preferred at locations such as voids, pores, open surfaces, and micro-fractures.

Our novel approach using synthesis of a doped ceramic, ion irradiation, chemical imaging, and atomistic simulation has provided valuable insights into the early stages of metallic phase formation in irradiated nuclear fuel surrogate.

The next steps are to characterize the role of iodine in metal particle formation, identify the metallic phase that forms as a function of temperature, and connect the findings of this work with the results of an ongoing project on UO_2 to develop new mechanistic understanding of phase formation in nuclear fuel.



Nanoparticles enriched in Mo, Ru, and Rh form in doped ceria after annealing at 1,350 K.

PN16055/2832

Low-Cost Rechargeable Aqueous Zinc Batteries

Xiaoliang Wei

This project aims to develop new cathode materials for zinc ion batteries to enable cost-effective energy storage systems, which will support the evolving global energy landscape with increased deployment of intermittent renewable energies.

Zinc ion batteries are attractive storage technologies because of the favorable properties of zinc anode, but very few cathode materials have been explored, because they are often plagued with limited cyclability or specific capacity.

To address these challenges, this project aims to develop Zn^{2+} intercalation cathode materials with two-dimensional (2D) layered lattice structures. The intercalation mechanism eliminates undesired electrolyte/electrode reactions and continuous consumption of electrolytes. 2D layered structures can provide lateral diffusion channels to enable fast Zn^{2+} ion transfer and accommodate large storage capacity.

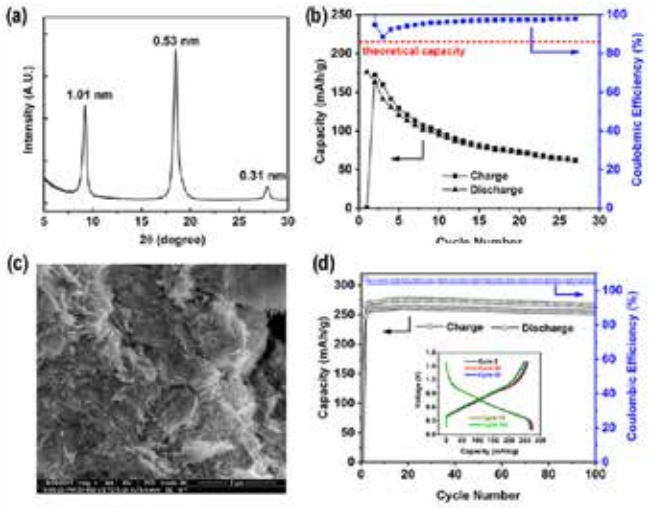
This research is expected to achieve promising materials and systems for zinc ion batteries with competitive performance merits of high power rate, high energy capacity, and long cycling stability.

This research developed new 2D layered transition metal oxide (TMO) materials, as well as the effects of interlayer spacing and morphology on their electrochemical performance in zinc ion batteries. Several TMO cathode candidates have been studied in zinc ion batteries.

The interlayer spacings of $\bar{o}$-MnO_2 were expanded from ~7 Å to ~10 Å by pre-inserting Zn^{2+} ions to decrease the energy barrier for Zn^{2+} intercalation. Improved Zn^{2+} diffusion and storage characteristics have been achieved, evidenced by a gravimetric specific capacity of 172 mAh/g, corresponding to a material accessibility of approximately 80%. However, the cells demonstrated continuous capacity fading, indicating the stability over extended charge/discharge cycling still needs to be enhanced. Detailed mechanisms of electrochemical reactions and capacity fading are being studied.

Another direction of research is to develop high-surface-area V_2O_5 aerogels to reduce Zn^{2+} ion diffusion paths and increase electrolyte accessibility. The cell using the V_2O_5 aerogel cathode and a nonaqueous $Zn(TFSI)_2$/MeCN electrolyte demonstrated a gravimetric specific capacity of approximately 270 mAh/g (i.e., 93% material utilization). More importantly, excellent cycling stability was achieved with no discernible capacity loss over 100 galvanic cycles, indicating the high structural stability of the aerogel V_2O_5.

The future plans will extend the research to other 2D layered materials, such as transition metal dichalcogenides, and carbides and nitrides. Fundamental understandings of structure-property relationships will be the focus to unveil the determining structural characteristics for Zn^{2+} intercalation and guide further materials design and engineering strategies.



(a) X-ray diffraction of Zn^{2+}-doped $\bar{o}$-MnO_2; (b) galvanostatic cycling of a zinc ion battery composed of Zn anode, Zn^{2+}-doped $\bar{o}$-MnO_2 cathode, and 2M $ZnSO_4$ electrolyte; (c) scanning electron microscopy of V_2O_5 aerogel; (d) galvanostatic cycling of a zinc ion battery composed of Zn anode, V_2O_5 aerogel cathode, and 0.6M $Zn(TFSI)_2$/MeCN electrolyte.

PN17052/2942

Mastering the Macromolecular-Materials Interface for Energy Science

James J. De Yoreo

The purpose of this project is to determine the mechanisms, pathways, and dynamics by which engineered, macromolecular building blocks, which are designed to interface with inorganic surfaces, interact and assemble into two-dimensional (2D) and 3D structures, and subsequently, direct nucleation of inorganic materials. This is achieved through the use of molecular-level imaging and spectroscopy methods.

Living systems create macromolecular-inorganic interfaces to guide the formation of functional, hierarchical protein-scaffold-inorganic materials through biomineralization. Through this process, organisms create unique structures to realize complex biological functions. Previous studies have demonstrated that the assembly of protein scaffold is controlled by the sequence dependent energy landscape that is highly sensitive to the surrounding environment.

We have also found that the nucleation of crystals on protein scaffolds are impacted by a free energy barrier associated with the interfacial match between macromolecular matrices and the crystal lattice, and kinetic barriers associated with atomistic processes. However, still lacking is the knowledge and capabilities to design a hierarchical protein scaffold to control the formation of functional materials with complexity and specificity. Our goal is to combine a protein-engineering and *in situ* imaging approach to define the energetic and kinetic controls on assembly of 2D and 3D scaffolds, and control the nucleation of inorganic materials on those scaffolds.

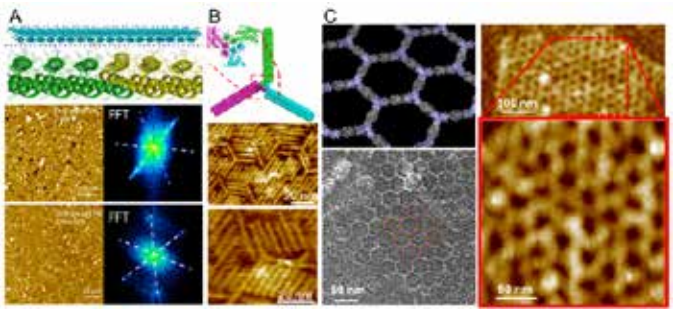
In FY 2017, we explored the nucleation and self-assembly behavior of the purpose-designed helical repeat protein (DHR10-micaX) via *in situ* high-resolution atomic force microscopy (AFM). Through the proper choice of background cations and the use of oriented single-crystal substrates, DHR10-micaX with dimensions in the nanometer range can self-assemble into 2D liquid crystals with uniform orientation over at least a million-times their length. This behavior and process can be altered by the concentration of DHR10-micaX and cations, with the uniformity of orientation controlled by the type and concentration of cations. From the dependence on cation concentration, evolution of order, and similarity of colloidal systems, we hypothesize that assembly is entropy driven.

We explored modifications to the sequence of DHR10-micaX (Mutant I) in order to vary the final self-assembled architecture. Upon increasing the protein head-head interaction (Mutant I), the proteins aligned into nanowire arrays, with uniform orientation across distances of up to hundreds of times the DHR-micaX length. In addition, we linked together three DHR10-micaX molecules (Mutant II) to form a three-fold symmetric building unit with three arms, with the goal of forming hexagonal networks. DHR10-micaX and its mutants have the potential to serve as artificial scaffolds for nucleation of inorganic materials.

We also succeeded in assembling a 2D hexagonal network with two distinct protein units—one trimer and one dimer—and characterizing the network via transmission electron microscopy (TEM) and AFM. The images demonstrate the success of this protein design and provide a path to further *in situ* studies. The preliminary *in situ* data imply the trimer state is stable, and the addition of the dimer is the slow step in the assembly process.

In FY 2018, we will use our *in situ*, high-speed AFM capability to quantify the self-assembly dynamics of DHR10-micaX and Mutant I. With the support from theoretical simulation, we seek to understand the mechanism of protein assembly at solid-liquid interfaces and to define the energetic and/or kinetic factors that control the process.

Based on our successful demonstration of self-assembly by DHR10-micaX, its mutants, and the two-component 2D protein system, we will use protein engineering to further mutate the sequences to drive the nucleation of inorganic materials, like ZnO and Au, with long-range order defined by the protein network.



Column A from top to bottom: side view of DHR10-micaX molecule, side view of head-head interaction of DHR10-micaX Mutant I, AFM images of the nanowires of Mutant I with fast Fourier transform pattern. Column B from top to bottom: structure of DHR-micaX Mutant II, and the AFM images. Column C: structure of the two-component 2D protein hexagonal network and TEM and AFM images.

PN17069/2959

Modeling the Interfacial Effects, Partitioning, and Production Routes of Epsilon Particles in Uranium Oxide

Richard A. Clark

The goal of this work is to elucidate the physical and chemical mechanisms that govern the formation of noble metal inclusions in irradiated fuel to understand the radiological, physical, and chemical conditions leading to their formation and unique properties.

The chemical fractionation leading to the formation of five-metal "epsilon phases," composed primarily of the elements Mo, Tc, Ru, Rh, and Pd in a reduced state, from nuclear fuel under irradiation is a key knowledge gap in our understanding of nuclear fuels. The development of these phases within uranium oxide matrices over time under irradiation has been linked to many important phenomena over the life cycle of nuclear fuel, from in-reactor operations to long-term spent fuel disposition. It has been hypothesized that epsilon phase formation can affect fuel performance, control cladding erosion, and cause stress corrosion through both the physical disruption of the fuel matrix due to growth and the alteration of the fuel's electrochemical behavior resulting from the presence of a unique chemical species.

This effort aims to understand the formation mechanisms of the epsilon metal phase during nuclear fuel irradiation and to understand the solid-solid-gas interfacial equilibrium effects that lead to variations in the properties of this unique phase. Developing a fundamental understanding of the chemical, radiological, and material structure conditions that leads to precipitation of these phases will significantly improve our knowledge of nuclear fuel under irradiation conditions.

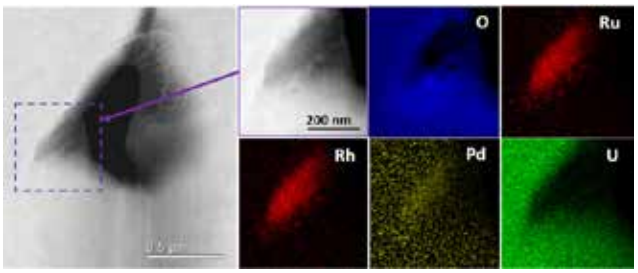
To understand the formation of epsilon particles, we utilized a varied analytical suite of techniques that combine experimental and modeling results to guide the research in a synergistic loop. Research was categorized into four key areas: Isotopic Analysis of Bulk Separated Particles, Advanced Microscopy of *In Situ* Particles, Microscale Irradiations, and Modeling.

Previously separated noble metal epsilon particles from three commercial irradiated fuels were used for complete characterization. The separated particles were dissolved and elements were separated for isotopic analysis. During the analysis, a sixth major metal (tellurium) was discovered as part of the epsilon phase. In FY 2017, the separations and isotopic analysis were completed. The isotopic ratios for the epsilon particle metals yielded information about the genesis of the particles in the fuel—particularly indicating whether the included elements are produced by fission, activation, decay, or a combination.

During FY 2017, this project reestablished a Radiochemical Processing Laboratory capability: a full cross section of an irradiated UO_2 fuel pellet was prepared and removed from the hot cells. Using the focused ion beam scanning electron microscopy (FIB-SEM), suitable samples were prepared for transfer to non-radiological facilities within PNNL to take advantage of the state-of-the-art analytical tools. These facilities are unable to handle conventional sub-samples of an irradiated fuel sample; thus, the FIB-SEM preparation demonstrated a method of generating samples of radioactive material that makes it safer and cheaper to use in studies and offers collaborative opportunities never before possible.

Analysis of samples via transmission electron microscopy and atom probe tomography is ongoing. We found that the concentration of each individual element was non-uniform, with some correlation between elements and independent variation between others.

In FY 2017, microscale samples (approximately 20 ng) of fresh UO_2 were irradiated in collaboration with the University



Formation of noble metal inclusions in a microscale irradiation of UO_2 fuel.

PN15100/2775

of Missouri. In this work, noble metal precipitate formation was observed at very low burn-up equivalence. This work provides the first viability of microscale irradiations.

Thermodynamic modeling was incorporated into this project in order to help determine the relative stability of different phases in the epsilon particle system. Early efforts used density functional theory (DFT) to simulate the free energy of formation of the Mo-Ru and Pd-Ru binary alloys. These calculations showed that the elemental composition observed in experimental studies is consistent with the most stable metal phase. More recent DFT work has included tellurium, which we have demonstrated is capable of forming a new, non-metallic phase that contains a high percentage of the total palladium and exists in close proximity to the epsilon metal phase. In this way, local differences in elemental composition can be observed, even on a scale of just a few nanometers, thereby enhancing the known complexity of epsilon particles even further.

Nucleation and Crystal Growth in 2-Dimensional Confinement

John S. Loring

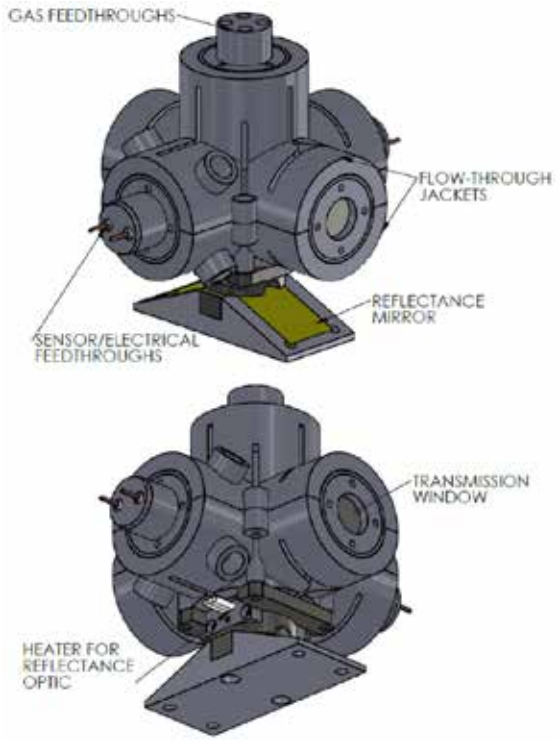
Adsorbed H_2O films exist whenever surfaces are in contact with humid atmospheres, environments that are ubiquitous across a wide range of disciplines, including geochemistry, materials science, catalysis, and atmospheric science. The goal of this project is to develop a novel infrared (IR) spectroscopic titration apparatus that will be used to investigate how the thickness of two-dimensionally (2D) confined H_2O films controls persistence of metastable precipitates and creates unique pathways to nucleation and crystal growth.

While previous studies have shown that nucleation under confinement versus from bulk solution can occur at dramatically different rates, these investigations have been limited to reactions in 3D space. There has been no systematic study to date concerning confinement effects on nucleation and growth in 2D (i.e., in Å- to nm-thick adsorbed H_2O films). For example, it is not understood why amorphous precipitates can persist below a threshold H_2O film thickness on mineral surfaces, yet beyond this threshold, large µm-sized crystalline carbonates form from nm-thick adsorbed H_2O films.

The overall aim of this research is to determine how thickness-dependent properties of adsorbed H_2O films can either stabilize metastable phases or facilitate crystal nucleation and growth. A key challenge is the development of instrumentation and methodology for systematically investigating 2D confinement effects. Hence, the specific goal of this project has been to design and assemble an IR titration cell for *in situ* and molecular-level measurements of H_2O sorption and thin film reactivity.

In FY 2017, we assembled an IR spectroscopic titration cell for studying metastable precipitates, nucleation, and crystal growth in 2D confined adsorbed H_2O films. Until now, no ambient-pressure IR capability existed that could quantitatively probe the partitioning of multiple gases onto solids while providing molecular-level insights about the structures of surface-sorbed species. This project has filled this need by developing a fully automated titration system that utilizes *in situ* IR spectroscopic monitoring of interfacial reactivity and sorbed species concentrations as a function of systematic changes in total gas phase reactants.

Innovative features of the system include: transmission IR optics for measuring concentrations of H_2O and other reactants in the gas phase and adsorbed on solids; attenuated total reflection IR optics for collecting spectra of adsorbed H_2O films on solids and monitoring reactivity within those films; water jacketing for variable temperature experiments from -40 to 200°C; gas feedthroughs for quantitative addition of reactants from autotitration instruments through a valve manifold, which also supports vacuum and continuous gas flow; and electrical feedthroughs for sensors that measure impedance of adsorbed H_2O films and mass uptake using a quartz crystal microbalance.



Schematic of the spectroscopic cell of the IR titration system for investigating reactivity in confined H_2O films.

PN17096/2986

Observing and Quantification of the Initial Stages of Nucleation and Growth in Liquids

Nigel D. Browning

This project will focus on a systematic analysis of a representative set of systems that exhibit specific processes important to energy technologies.

A fundamental atomistic understanding of precipitation phenomena, such as the nucleation and growth of nanostructures from an aqueous solution, can have far-reaching implications on new material synthesis and the design of cheap, stable, efficient batteries, which are research themes that align well with DOE goals. As part of our work, the unprecedented sub-μs temporal resolution of the dynamic transmission electron microscope (DTEM) will be developed to obtain a quantitative observation about the evolution of structures from individual atoms to mesoscale structures and complexes under varied environmental conditions that will be correlated with first principles modeling/simulations. These observations, including the fast transient intermediate states in each process, will provide unique insights to optimize material synthesis routes and operating parameters for advanced energy technologies.

In FY 2015, we focused on establishing DTEM as a working instrument capable of time-resolved imaging. In addition, preparations have been made for nanoscale experiments that can be performed when the DTEM becomes operational. These preparations involve two sets of experiments: one, a synthesis of nanoparticles involving bimetallic core shell architectures; the other, the identification of active catalytic sites in an oxide catalyst. A key result was the demonstration that DTEM could work as a pulsed microscope.

In FY 2016, we optimized the laser pulses and the alignment of the microscope column to achieve high spatial and temporal resolution images. Alignment improvements during the course of the year both identified several areas of the optical beam path that needed to be modified and improved the alignment of the column to push the spatial resolution of the pulsed images towards approximately 10 nanometers. During the course of this work, many of the critical lifetime limitations of the emission process were determined, and an optimal approach to increasing the operation time of the microscope was determined.

Work in FY 2017 focused on understanding the optics within the gun system for DTEM and establishing compressive sensing and inpainting methods for improving resolution in extremely low-dose electron images. In the case of the gun optics, aligning and improving the quality of the laser beam hitting the photocathode was improved.

Work is now underway to optimize the position of the photocathode relative to the first aperture to improve the throughput of electrons and increase the overall beam current. The inpainting approach of using mathematical methods to increase the information extracted from low-dose and noisy images has proven to be extremely advantageous for transmission electron microscopy imaging. Images can be shown to have an improved resolution and sensitivity from doses that can be one to two orders of magnitude reduced over standard conditions. This approach means that DTEM resolution can be enhanced at lower currents, thereby increasing the stability of the system and the ease of imaging.

PN15013/2688

Peptoid-Based Biomimetic Materials with Tunable Structures and Functions

Chunlong Chen

We are developing sequence-defined peptoids as designable building blocks for assembly of biomimetic materials with hierarchical structures. We also use* in situ *imaging techniques and molecular simulations to track and understand peptoid assembly pathway and mechanisms.

The exquisite self-assembly of proteins and peptides into highly ordered and dynamic functional materials has inspired innovative approaches to biomimetic materials design and synthesis. Although synthetic approaches may yield highly stable polymers with intricate architectural morphologies, the introduction of peptide-like side-chain diversity has been a long-standing challenge for synthetic polymers. In 1992, Zuckerman et al. developed the solid-phase synthesis of peptoids (poly-N-substituted glycines) that combined synthetic opportunity for side-chain diversity with high molecular stability.

Subsequent research on peptoids focused primarily on mimicking structure and functionality of peptides for drug discovery and molecular recognition. More recently, it has been demonstrated that peptoid side-chain interactions can be exploited for assembly. Unlike peptides, peptoids do not form backbone hydrogen bonds; therefore, the assembly of peptoids into nanostructures will be less complex to understand and predict than that of peptides, while still having the capability of tuning assembly and function through side-chain structure and sequence. In addition, the peptoid architecture will yield stable materials for future functional applications.

In this work, peptoids will be used as designable building blocks by exploiting the reversibility of supramolecular interactions to assemble them into biomimetic materials and understand their assembly mechanisms. We will also address the challenge of predicting ensemble outcomes of assembly driven by the molecular details of sequence and the consequent interactions. Fusion of such experimental efforts with extensive computational simulations will enable the rational design of sequence-defined polymers for which self-assembly is programmable, and the resulting function is predictable.

In FY 2017, we designed and synthesized a large number (over 30) of peptoids that assembled into membrane-mimetic two-dimensional (2D) nanomaterials, nanotubes, and pore-forming networks. We determined the role of peptoid hydrophobic side chains in controlling assembly morphologies and pathways, as well as the mechanism by which peptoid assemble. By designing a series of a series of amphiphilic peptoid oligomers (APOs), Nce6Nbpmn (n = 5, 6, 7) and their functionalized sequences, Nbpm = N-[(4-bromophenyl)methyl] glycines and [Nce = N-(2-carboxyethyl)glycine].

We demonstrated the assembly of APOs into a new family of highly designable, stiff, and dynamic single-walled peptoid nanotubes (SW-PNTs) through a unique "rolling-up and closure of nanosheet" mechanism. These SW-PNTs undergo a drastic contraction of approximately 46% in height upon solution pH decreases, and this pH-triggered response is reversible. We further demonstrated that SW-PNTs can be rationally engineered to tune their surface chemistry, wall thickness, PNT diameter, and mechanical properties. By precisely introducing β-cyclodextrins (CDs) or arginylglycylaspartic acid peptides within SW-PNTs, we demonstrated the applications of functional PNTs in purifying azo-contaminated water and in enhancing cancer cell adhesion and uptake. This study offers the first route to assembly of stiff and dynamic nanotubes from sequence-defined synthetic molecules.

To increase the transmission electron microscopy (TEM) contrast and mechanical properties of peptoid assemblies, we synthesized over ten peptoid sequences that contain polyhedral oligomeric silsesquioxane (POSS), a cage-like molecular structure that could be regarded as one of the smallest possible silica particles. In this design, POSS was introduced during the solid-phase peptoid synthesis as a side-chain. We demonstrate the assembly of POSS-containing peptoids into membrane-mimetic 2D nanomaterials and nanotubes. Atomic force microscopy (AFM) and TEM techniques were used to characterize these morphologies, and X-ray diffraction data show that POSS-containing peptoid membranes and nanotubes are crystalline. More importantly, as expected, these 2D nanomembranes and nanotubes can now be visualized under TEM without negative staining, which offer the potential of investigating the assembly pathways and kinetics of

PN17032/2922

peptoid assembly using *in situ* TEM technique. We also demonstrate that dye molecules, such as dansyl and rhodamine can be introduced into these POSS-containing hybrid peptoid assemblies to generate biomimetic materials with ordered structures and tunable fluorescence properties.

Two-step nucleation pathways in which disordered, amorphous, or dense liquid states precede appearance of crystalline phases have been reported for a wide range of materials, but the dynamics of such pathways are poorly understood. Moreover, whether these pathways are general features of crystallizing systems or a consequence of system-specific structural details that select for direct versus two-step processes is unknown.

Using *in situ* AFM to directly observe crystallization of peptoids, we showed that crystallization pathways are sequence dependent. When a short hydrophobic region is added to the sequence that directly forms crystalline particles, crystallization follows a two-step pathway that begins with creation of disordered clusters of 10 to 20 molecules and is characterized by highly non-linear crystallization kinetics in which clusters transform into ordered structures that then enter the growth phase. These results shed new light on non-classical crystallization mechanisms and have implications for design of self-assembling polymers.

The backbone parameterization of peptoids has been refined and is now consistent with high-level quantum simulations. Enhanced sampling methods showed the importance of peptoid sequence for the dimerization, highlighting that the significant effect of hydrophobic residues on peptoid assembly process. We found that poly sarcosine showed a much larger conformational flexibility and stronger interactions with water than the peptide analogue poly-alanine under bulk and interfacial solvation environments.

During FY 2018, we will use both *in situ* AFM and TEM to investigate the assembly pathways and dynamics, as well as the formation of mechanisms of peptoid-based nanomembranes and nanotubes. We will work on introduction of chromophore donors and acceptors into crystalline nanomembranes and nanotubes to build artificial light harvesting systems. Electron donor (D) and acceptor (A) will be also introduced as side-chains of membrane-forming peptoids with tunable D–A strengths to build long-range interactions with assembled peptoid membranes to achieve programmable and stimuli responsive transporting and optical properties. We will investigate the influences of 1) the ordering of peptoid assemblies, 2) the density of D and A, and 3) the local environments of chromophores on the optical and electronic properties of functionalized peptoid assemblies. The simulation effort will focus on the molecular details of the SW-PNTs and quantification of the driven forces that lead to formation of nanosheets and nanotubes.

Probing Collective Phenomena at Solid-Liquid Interfaces Under Reaction Conditions

Grant E. Johnson

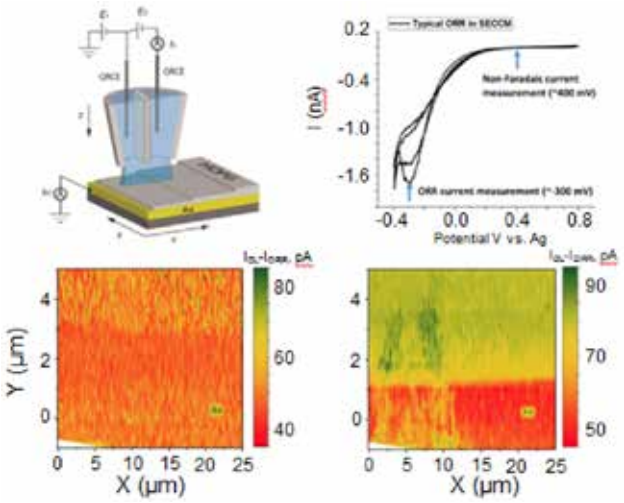
This project aims to understand collective phenomena and atomic-level processes occurring at solid-liquid interfaces and how they influence electrochemical performance at dynamic reaction conditions. New capabilities are being developed for synthesis and electrochemical imaging of supported cluster ensembles and* in situ *liquid X-ray photoelectron spectroscopy (XPS) that will have broad applications in energy storage and catalysis research.

Electrochemical processes that are critical to energy generation, storage, and catalysis take place at solid-liquid interfaces and may be promoted by supported clusters. Clusters have tunable properties that are not present in the bulk due to their large number of active surface sites, size-dependent redox reactivity, structural plasticity, and support-specific binding. While it has been demonstrated that certain intercluster interactions may improve the catalytic activity of clusters ensembles, it is not understood how collective cluster-support and cluster-cluster interactions change with coverage from isolated single clusters to dense assemblies and what effect this has on electrochemical phenomena.

Soft landing of mass-selected ions is ideally suited to preparing precisely defined ensembles of supported clusters with predetermined size and coverage for characterization by spatially resolved electrochemical techniques. These cutting-edge methods, when combined, will provide the insight required to design superior materials for energy generation, storage, and catalysis using models that incorporate collective phenomena.

It also is well established that the solid-liquid interface determines the rates of reactions and selectivity toward products. Consequently, it is important to determine the geometric and electronic structure at solid-liquid interfaces at dynamic reaction conditions. Characterizing the atomic-level processes involved in the formation of solid electrolyte interphase (SEI) layers in rechargeable batteries remains an outstanding challenge due to the vacuum requirements of traditional surface analytical techniques. Limited understanding of the process of SEI formation, the structural and chemical properties of the SEI, and how it evolves during charge/discharge cycles limits the advancement of technology for rechargeable batteries. To address this challenge, an *in situ* XPS capability is being developed to gain insight into the electronic structure and chemistry of solid-liquid interfaces at reaction conditions.

In FY 2017, we used magnetron sputtering and gas aggregation to generate clusters and nanoparticles with controlled composition and morphology for a variety of applications. We also explored how the presence of different functional groups in phosphine ligands influences the structures adopted by metal clusters that may subsequently be soft landed onto electrode surfaces.



Top left: Schematic diagram of the theta capillary used to form an electrochemical cell with the surface for imaging. Top right: Cyclic voltammograms of Pt_{30}- clusters soft landed on glassy carbon. Bottom left: two-dimensional scanning electrochemical cell microscopy (SECCM) image of 2 x 10^4 clusters/μm^2. Bottom right: SECCM image of 5 x 10^4 clusters/μm^2.

In FY 2018, we will control the size and spacing of supported clusters on conductive electrodes using soft landing of mass-selected ions onto templated substrates. Cluster formulations with specific size, spacing and coverage that exhibit enhanced oxygen reduction reactivity will be further characterized using atomic force microscopy and scanning electron microscopy, as well as *in situ* liquid transmission electron microscopy and XPS. This multimodal approach will allow us to understand how the reactivity, selectivity, and durability of supported cluster ensembles may be tailored for specific electrocatalytic processes that are ubiquitous in catalysis and energy storage.

PN17070/2960

Quantum Defects in Synthesized Diamond Aerogel and Diamond Nanoparticles

Ilke Arslan

Quantum defects in materials are atomic-scale defects in crystalline solids that can act as quantum sensors in a variety of ways across the physical and life sciences. Spin-dependent properties make diamond defects great for quantum computing or local sensing at the nanoscale, even within living cells. Diamond is chemically inert and has shown promise in drug delivery and promoting bone growth.

Diamond materials have received a substantial amount of both theoretical and experimental attention in recent years. In the physical sciences, diamond has been investigated for use in catalytic nitrogen and oxygen reduction reactions, optically addressable qubits, optical magnetometry, stable electrochemical double-layer capacitors, seeding the growth of polycrystalline diamond thin films, and also for monitoring the flux, position, and timing of intense X-ray beams from fourth-generation synchrotron light sources.

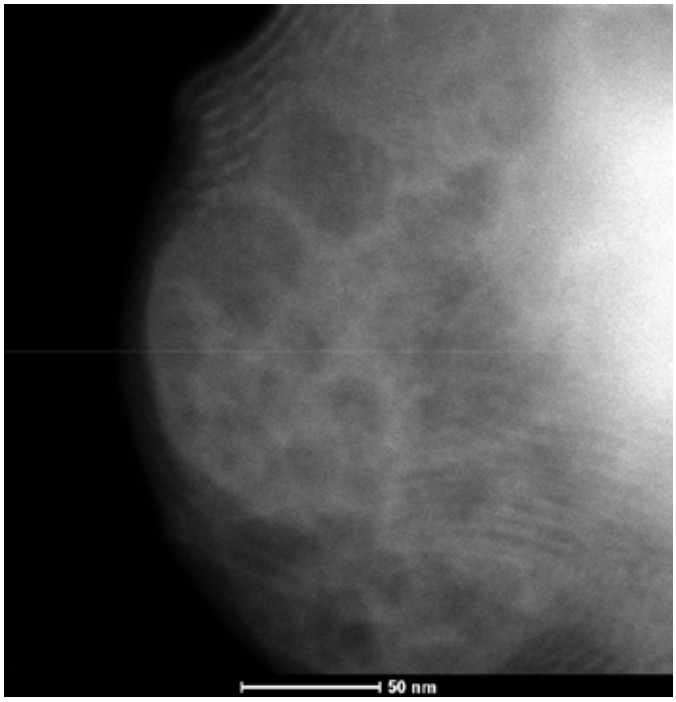
In the biological sciences, diamond materials have been investigated for non-cytotoxic targeted-delivery of therapeutic molecules and for long-term visible and near-infrared (NIR) *in vitro* cell labeling. However, to date there are no well-established methods for synthesizing nanocrystalline diamond materials with well-defined point-defect microstructure. This fact stands in stark contrast with the immense knowledge base and literature for controlling both the size and composition of other semiconductor nanomaterials such as quantum dots and porous silicon. The central goal underlying the basic research in this project is to uncover new approaches for the synthesis *and* characterization of nanocrystalline diamond materials with well-defined point-defect microstructure.

The central aim of this project is twofold: 1) demonstrate control over the concentration of nitrogen vacancy (NV)-centers within a nanocrystalline diamond aerogel synthesized at high-pressure, high-temperature conditions using a laser-heated diamond anvil cell, and 2) to develop a protocol to map the distribution of nitrogen dopant atoms in the final material using high-resolution imaging and spectroscopy in the scanning transmission electron microscope (STEM) to advance understanding of quantum defects in diamond materials. STEM tomography will be used to illuminate the three-dimensional (3D) morphology of the diamond aerogel and complement the high-resolution spectroscopic mapping. This combination of synthesis and characterization will provide a feedback loop to work towards a controlled distribution of nitrogen dopant atoms.

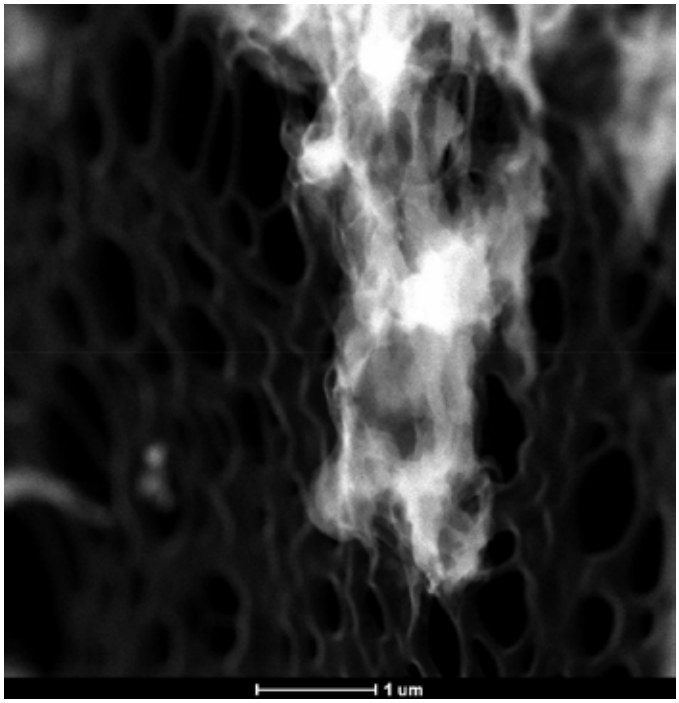
In the short 6 months of this proposal, research was performed on both tasks. The Pauzauskie group synthesized diamond aerogel specimens. Due to the difficulty of the synthesis process at high-pressure and high-temperature conditions, only a small amount of material is synthesized during each experiment. The first set of challenges is mounting an appropriate amount of material on a grid that will be well suited for electron tomography and microscopy. It must be in the center of a grid square and one that is wide enough to reach high angles as the specimen is tilted to acquire many different images. The results of our first experiments working together showed that the diamond material they synthesized was high quality, but some of the small pieces fell off during shipment, and others were not suitable for tomography because of the similarities between the specimen and the grid. However, we were able to obtain good information on the outcome of the synthesis.

Separately, electron energy loss spectroscopy was performed on test specimens to determine the parameters for spectral imaging. It was determined that the material must be extremely thin to be able to observe any NV centers, and as uniform as possible. The material should be below 50 nm in thickness, but closer to 20 or 30 nm would be better. Further, the density fluctuations of the carbon or thickness changes can affect the spectral results, so future spectral imaging to probe NV centers will likely need to be performed on specimens that have been put into a focused ion beam (FIB) and milled to a uniform thickness across of approximately 30 nm in thickness. Tomography will be performed on the materials in their native state, before the FIB sectioning, to ascertain their 3D morphology. Lastly, special care will be taken to place the small amounts of material on specific flat grids, and special shipping materials or hand delivering will be required to overcome any lost material in transition.

PN17080/2970



A nanoscale STEM image of the morphology of the synthesized diamond. The porous nanostructure and ordered carbon layers are visible.



A lower magnification image of the synthesized diamond taken at a tilt angle of 60 degrees. Due to the unwanted thickness of the support grid, and its similar structure to the diamond, interference would occur in the reconstruction of the 3D tomogram. Therefore, this particular specimen/grid could only be used for 2D imaging.

ShAPE Processing for Improved Corrosion and Creep Resistance Properties

Glenn J. Grant

In this project, we developed a magnesium alloy with enhanced corrosion resistance and enhanced the corrosion resistance of magnesium by developing a novel functionally/galvanically graded interface with aluminum cladding.

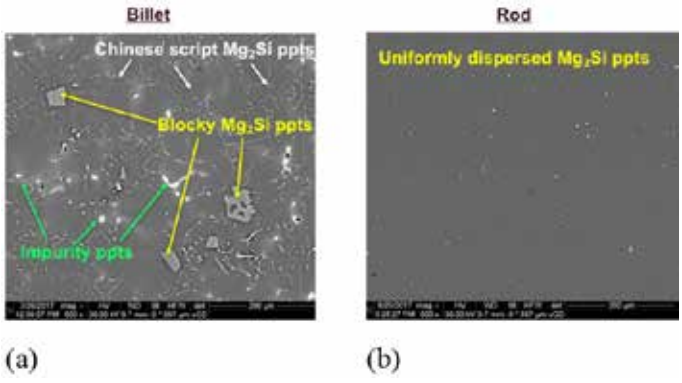
The low density and high specific strength of magnesium alloys makes them ideal candidates to replace many of the steel and aluminum alloys currently used in transportation systems. While cast components of magnesium and its alloys have a significant history of use as automotive components, the development of wrought components has been primarily hindered by magnesium's lack of formability at room temperature and by its corrosion resistance, especially in multimaterial systems. The enhancement in corrosion resistance can be brought about by alloying it with certain elements, such as silicon, or cladding it with aluminum or similar high corrosion resistance material.

The primary challenge associated with alloying is prevention of microgalvanic corrosion. This can be prevented by use of certain elements, such as silicon, which form Mg_2Si particles that have similar galvanic potential as that of Mg matrix. Despite their compatibility, these precipitates typically reduce the ductility and makes the alloy difficult to process or form into desirable shapes, thus limiting their use/application. Another way to improve the corrosion resistance of Mg is by cladding it with a corrosion resistance alloy or metal such as aluminum. Several technical challenges arise when forming such structure including controlling the texture of the magnesium, such that the asymmetry in mechanical properties under compression and tension is eliminated; the preferred grain size of the magnesium is less than 5 microns with an aluminum cladding bonded to the magnesium and, at the same time, forms a galvanically graded interface to minimize the corrosion rate in the system.

The proposed work will build upon existing PNNL research where the Shear Assisted Processing and Extrusion (ShAPE) process has been shown to improve material properties through novel microstructures that are possible only during far from equilibrium processing via ShAPE. Over the past several years, we have developed a novel ShAPE technique, which uses a rotating ram as opposed to the axially fed ram used in the conventional extrusion process. The ram face contains spiral scroll features. When the rotating ram face comes in contact with the material and a forging load is applied, significant heating occurs due to friction, thus softening the underlying billet material. The combined action of the forging load together with the rotating action of the ram face, force the underlying material to flow plastically. The scroll features on the ram face help in the material flow and help in controlling the texture. We have successfully demonstrated the scalability of this process, and we were able to alter and control the texture, grain size, and also uniformly disperse the secondary particles by changing a few process parameters and at loads/pressure several orders of magnitude lower than conventional extrusion.

With these recent discoveries and technical approaches serving as the foundation from which to build, new architectures and microstructures can be achieved that improve corrosion resistance in lightweight alloys. This work was intended to push the technical envelope into regimes (primarily temperature and microstructural control) where the scientific and engineering challenges are much more difficult and consequently much more impactful.

Magnesium alloyed with 2 and 7 wt% Si were cast into book molds. The as-cast microstructure yielded a Chinese script eutectic structure and Mg_2Si primary blocky phase. In order to enhance the mechanical properties of the already corrosion resistant alloy, these phases need to be uniformly

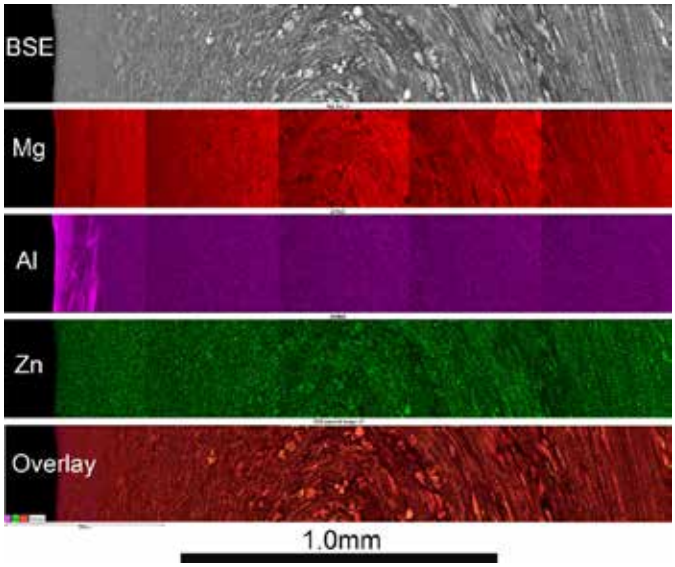


(a) Microstructure of the as-cast billet. (b) Microstructure of the formed rod after ShAPE processing. Notice the uniform dispersion of the Mg_2Si particles in the matrix.

PN17077/2967

dispersed in the matrix. The ShAPE process parameters, such as tooling design, extrusion ratio, and rotations per minute (RPM), were varied in order to uniformly disperse the Mg_2Si phases in the matrix. Using this ShAPE process, we were able to uniformly disperse these particles and, at the same time, improve the elongation of the alloys to over 12%.

The research goal of improving corrosion resistance was to demonstrate a method of forming extruded magnesium tubing or rods with an aluminum cladding. ShAPE process parameters, such as torque, RPM, feed rate, and temperature, were evaluated to achieve desirably aligned texture and grain within the magnesium core. Billet geometries, tooling, and die designs (along with process parameters) were explored to effectively bond the aluminum cladding to the magnesium core. Research was conducted to ensure adequate bonding between the core and cladding, while also achieving an interface microstructure that considers galvanic corrosion.



Al cladding on Mg with a galvanically graded interface formed on Mg alloy using ShAPE process.

In evaluating the efficacy of the process, we 1) utilized the ShAPE process to investigate the process parameter space (i.e., tool/die dimensions, tool rotational speed, external heat input, feedstock input rate, etc.) to determine the processing conditions required to produce extruded rod with aluminum cladding over a magnesium core and also developed a high elongation corrosion resistance Mg alloy; and 2) performed microstructural characterization to determine grain size and texture in the processed parts and evaluated the mechanical properties.

PN17077/2967

Solid Phase Processing

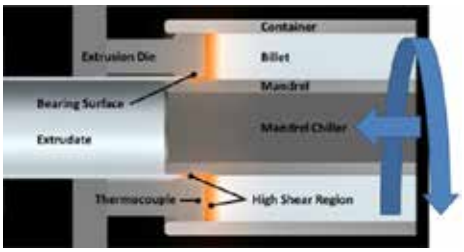
Douglas J. Waldron

PNNL researchers have discovered that processing materials entirely in a solid-phase state (as opposed to liquid/melt processing) can circumvent limitations that plague commercial practices and do not allow optimization of material microstructures to the fullest extent possible. By avoiding the melt and solidification regime entirely, materials difficult to form and microstructures that have never been producible before have now been demonstrated at PNNL.

As higher efficiency energy system designs emerge and continue to evolve, they require ever-increasing unprecedented materials performance. There is a growing problem, however, for designers and engineers to specify materials: there are none that adequately meet existing and emerging performance and durability requirements. As operating parameters become more extreme (e.g., ultra-high pressures, temperatures, and chemical compatibility), engineers are forced to compromise on performance for reliability and cost reasons. Material design requirements have quickly outpaced what the research community can deliver through incremental material property improvements. Further, nearly all emerging high-performance material production processing approaches are cost prohibitive and limited in scalability to commercial practice. A revolution in how structural materials are produced and assembled is needed across multiple markets. The United States needs an accelerated research program focused on superior, cost-effective materials, with performance far beyond that which is commercially available today.

This project's tasks focus on a specific use case, selected based on its relevance to sponsor priorities and user-inspired needs, but also to illustrate the diverse capabilities and applications of solid-phase processing (SPP). This first use case will be a high-strength, high-rate energy absorbent Mg-alloy for 1) vehicle frame/crush protection structures and 2) a low-cost, low-temperature formable Mg-alloy sheet for part stamping applications. Mg-alloys are being considered in vehicle technology frame applications based on their low density; however, application in front-end/engine box applications have been limited by low strength and crashworthiness. Similarly, the low density makes Mg-alloys attractive for sheet product and formed parts, but the high cost of high-temperature forming when accompanied by the loss in performance from exposure to the high temperatures of processing limit widespread commercial opportunities.

In the Shear Assisted Processing and Extrusion (ShAPE) process, a rotating billet is forced into a stationary extrusion die where significant heating occurs due to friction at the billet/die interface. The amount of heat generation is controlled by the applied torque, rotational speed, and chilled water flow rate through the mandrel. As temperature increases, the billet face softens and plastically flows through the spiral scroll features inward toward the extrusion orifice. Material streams within multiples scrolls are consolidated between the mandrel and bearing surface prior to exiting through the die relief. Due to the complex shear and thermal conditions present during ShAPE, crystallographic texture is thought to develop as a combination of recrystallization mechanisms due the interplay between rotational speed, extrusion rate, and thermal conditions. As such, the ShAPE process may offer unique control over grain refinement and crystallographic orientation, and hence, material property improvements beyond conventional extrusion techniques.



Cross section of the hardware showing a direct extrusion embodiment of the ShAPE

30-cm length of ZK60 tubing extruded using the ShAPE process along with a cross section showing the 1.52 mm

We have demonstrated an improvement (with respect to conventional rod and tube processing) in salient microstructural features known or determined to enhance properties and, thus, performance in end products. Key features include grain size, precipitate size and dispersion, and texture. We also demonstrated performance improvements of SPP-processed Mg-alloy tube and rod product over commercial, high-performance Mg-alloy counterparts by showing an increase in elongation and strength. In addition, we showed that the components can be formed at far higher extrusion ratios and pressures that were at least an order of magnitude lower.

PN17092/2982

Surface Modifications of Laminar Graphene Oxide Water Separation Membranes

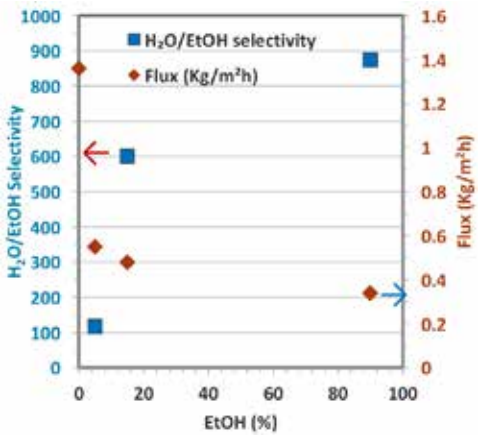
David Gotthold

Structured laminar membranes, consisting of hierarchically stacked, overlapping layers of two-dimensional (2D) graphene oxide, are a fascinating and promising class of materials with the potential for radically improved water separation with excellent selectivity and high permeability.

Laminar graphene oxide membranes have attracted a great deal of interest in the last few years as a potentially revolutionary method for separating water from other chemicals. However, early research has shown significant variation and instability in graphene oxide materials, which must be addressed. In previous work, we identified that the surface interactions likely play a dominant role in the performance of these membranes, especially when working with liquid separations. This work is focused on understanding the effects of both surface structure and chemistry on membrane behavior and developing surface modifications that enhance stability and performance for a range of applications including desalination, biofuel dewatering, industrial coolant recovery, and fracking water reclamation.

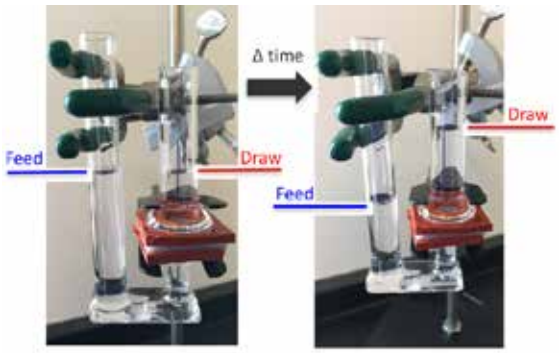
We have developed the fundamental test methodologies needed for understanding performance in liquid-liquid separation using graphene oxide membranes. We also developed stable and repeatable methods for measuring transport of organic and aqueous liquids through the membranes. Importantly, with water/alcohol mixtures, the presence of alcohol rendered the membrane water selective, as demonstrated by the generation of osmotic pressure. This first-time observation (along with X-ray diffraction data) indicates the formation of water channels within graphene sheets, presumably due to a high degree of ordering of the organics.

We have also demonstrated baseline chemical and structural modifications. For the chemical modifications, we incorporated diazonium benzene sulfonate and confirmed that it bonded to the graphene oxide flakes using optical absorption measurements. For structural modifications, we have evaluated the incorporation of two standard metal-organic frameworks (MOFs), ZIF-8 and UiO-66, into the graphene oxide films during the casting process and observed their incorporation with scanning electron microscopy.



Ethanol/water separation using graphene oxide membrane and showing rapid decrease in water flux with even small amounts of ethanol.

In FY 2018, we will use these baseline chemical and structural modification techniques to demonstrate energy-related chemical separations. We will then expand the most promising modifications to enable improved separation efficiency and membrane stability. We will also study modifications of the GO flakes within the membrane, including varying flake size through the thickness and selective oxidation/reduction of the surfaces to develop a better understanding of the fundamental transport behavior within graphene channels.



Draw	Feed	Result
1M NaCl	10mM NaCl	Initiation of liquid transport
MeOH+130mM NaCl	MeOH	Dramatic (~50x) enhancement of MeOH Permeance.
1:1 mixture of IPA+ DI water	DI water	Water selective membrane with very high Flux = 2.1 L/m²

Osmotic pressure in graphene oxide membranes showing high transport and selectivity.

PN17045/2935

Three-Dimensional Printing of Electrical Sensors for Biological and Chemical Detection

Rebecca L. Erikson

Three-dimensional (3D) printers have captured the imagination of Americans both young and old with the opportunity to create anything imaginable while in the comfort of your own home. This project provides a vision and a pathway to demonstrate 3D-printed parts capable of detecting chemicals that may someday be used as glucose monitors for diabetes monitoring or for vapor detection of hazardous chemicals such as gas leaks.

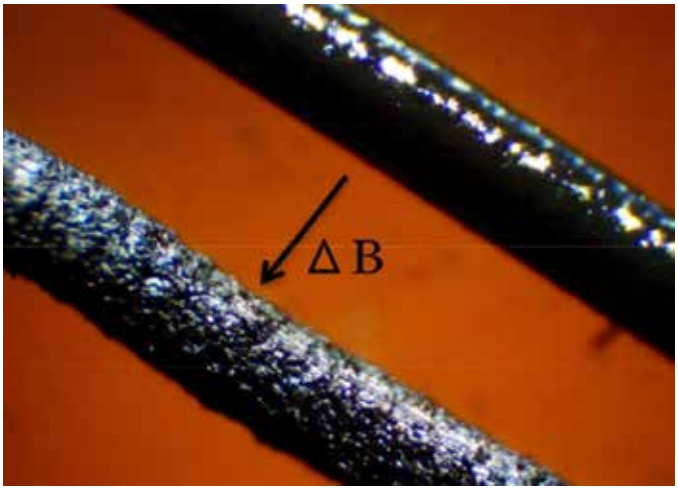
To date, 3D printers have been used primarily to create structural components of various systems, but limited efforts have been concentrated in the area of a functional sensing component. This project focuses on the development of methods in materials science, physics, and engineering to enable a fully functional, electrically active, biological, and/or chemical detector produced solely from a 3D printer. A large focus of this project is around custom formulations of materials that can be used with standard fused deposition modeling 3D printers.

Our multidisciplinary team has demonstrated significant science and engineering outcomes with our research over the past 2 years. In the first year, we developed and demonstrated techniques to produce both physical and electrical changes in material during a 3D print by altering the magnetic field around the extrusion nozzle during fabrication, which led to filing a formal patent application. By altering the magnetic field, we were able to show that a commercially available material can be made to change its conductive properties by threefold and that the surface area could also be increased or decreased.

We focused efforts on the development of polymers and polymer hybrids that could be used as feedstock to 3D printers with very specific chemical properties of interest. These materials included polymers that have very high conductivity (20 Ohms/cm), polymers that remain flexible after printing while maintaining conductivity, polymers that demonstrate piezoelectric properties after printing, and polymers that predictively change their conductivity when exposed to chemicals of interest.

In the second year of research, we focused on increasing the selectivity of chemical detection by attempting, for the first time, to incorporate metal-organic-frameworks (MOFs) into a polymer composite compatible with 3D printing.

Our research has yielded unexpectedly high utility with this composite. We have successfully demonstrated the ability to create these 3D printable composites with multiple MOF species, leading to active chemistry and capture of select chemicals in geometrically complex scaffolds. We also demonstrated the ability to create gradients of mechanical, electrical, or chemical utility on the fly through extrusion printing of polymer composites.



A change in the magnetic field (ΔB) significantly changes physical and electrical properties in this commercial polymer.

PN16015/2792

Towards a Better Understanding for Mineral Nanoparticle Assembly by Coupling Colloidal and Hydrodynamic Forces and its Application to Superlattice Formation of Nanocrystals

Jaehun Chun

This project will aim for a seamless connection of different length-scale frameworks, combining fundamental and detailed studies in theories, simulations, and experiments for mineral nanoparticle assembly. The project will establish a self-consistency between theories and simulations, validated by experiments. Such self-consistent framework, combined with a correct description of the coupling between energetics and dynamics, will build a rigorous understanding of the oriented attachment (OA) and the nanoparticle assembly in general.

OA has been observed and studied recently; it results from a combination of various short-/long-range forces between mineral nanoparticles, giving rise to the alignment of lattices of the mineral particles to their crystallographic axes. While orientation-dependent forces are beginning to be observed experimentally by using atomic force microscopy (AFM)-based dynamic force spectroscopy (DFS), a driving force for such preferential alignment is still unclear, especially as to a delicate balance/correlation between many existing forces that are influenced by detailed chemical physics at molecular scales.

This project is to develop a better understanding of the OA through a seamless connection between theories, simulations, and experiments, coupled with a correct description of dynamics of nanocrystals. An initial target system is ZnO(100). Through the superlattice formation of nanocrystals by using *in situ* transmission electron microscopy (TEM), the self-consistent framework can be generalized to understand the assembly of nanoparticles at large scales for various important applications such as energy storage.

AFM with the ZnO crystal tip was used to investigate force-distance. In order to stabilize the (100) tip, we used a thin layer carbon. Similar to AFM-based DFS measurements for other minerals, our results show 1) an approach process until the jump-to-contact, 2) a tip-substrate contact or solvent-separated stage, and 3) a tip retraction until a jump-from-contact. With the stable AFM tip, we will continue to measure the forces over different physicochemical parameters, such as pH and salt concentrations, along with a better experimental procedure to understand forces.

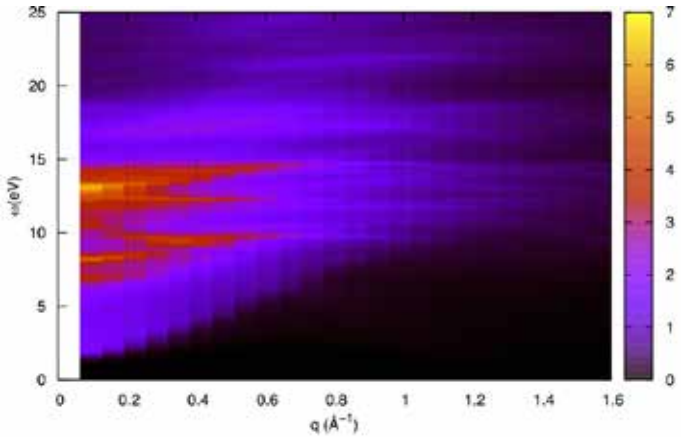
We used the density functional molecular dynamics (DFT-MD) method to study the surface hydroxylation of ZnO(100) contacting with bulk water, and we will continue using it to conduct constrained MD simulations using ReaxFF to investigate the dependence of hydroxylation on the particle-particle interactions.

We have used DFT to calculate the complex frequency- and wave-vector dependent microscopic inverse dielectric function of crystalline ZnO. We will construct a scheme and associated code based on real-space multiple scattering theory, combined with plane-wave DFT calculations of the microscopic dielectric properties, to compute Green's functions for the full electromagnetic fields (including both vector and scalar potentials) of cluster models of crystalline surfaces.

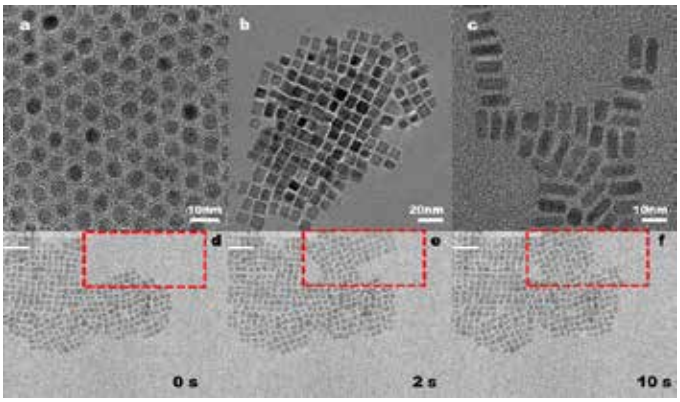
We successfully synthesized various shapes of monodisperse nanoparticles based on Pd, Ag, and Au. To investigate the self-assembly process of nanoparticles at non-equilibrium state, liquid cell TEM was used to monitor the trajectories of individual nanoparticles when dispersed in toluene. This facilitates an understanding of the mechanisms of solvent drying-mediated self-assembly. In addition, our study showed that Pd nanocube clusters, already self-assembled in toluene, combined and re-arranged to construct a superlattice structure.

We will study self-assembly mechanisms of Pd nanorods/nanospheres by using liquid cell TEM, focusing on the relation between the shape and force between nanoparticles. Further, we will synthesize MFe_2O_4 (M=Fe, Mn, Co, and Zn) nanoparticles, AO (A=Mn, Co and Zn) nanoparticles, and Te-based nanoparticles (PbTe, Ag_2Te, Bi_2Te_3) with various sizes and shapes. We will study particle forces through binary self-assembly processes at an equilibrium and non-equilibrium states.

PN17031/2921



A color map of the q-diagonal (along the 001 direction) energy-loss function of ZnO.



TEM overview images of a) Pd nanospheres (4 nm), b) Pd nanocubes (7×7 nm), and c) Pd nanorods (4×10 nm). Snapshots of *in situ* TEM observation of self-assembly process of Pd nanocubes in toluene: d) 0s, e) 2s, and f) 10s.

Towards Polarization-Switched Solid-State Molecular (POSSM) Pumps

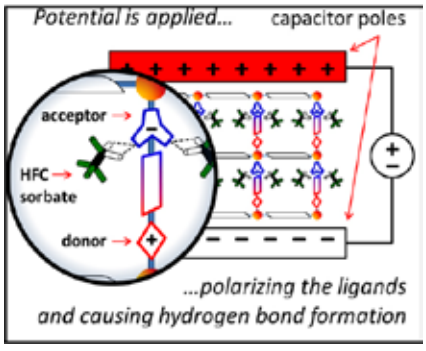
Carlos A. Fernandez

The goal is to demonstrate a highly efficient molecular pumping approach that could render obsolete mechanical compressors used in many applications today. By using a capacitor that can be switched "on" and "off," we will provide a simple basis for non-mechanical pumping that operates at the molecular level and involves zero moving parts.

Although methods vary widely, the way fluids are pumped against pressure gradients is by means of mechanical compressors. These compressors all operate with the same basic underlying principles and achieve compression by changing the velocity, enclosed volume, and/or temperature of the gas flowing through the compressor.

We seek to introduce a fundamentally new compression approach that operates at the molecular level and involves zero moving parts. The main goal at the end of the project is to demonstrate this principle through the fabrication of a polarization-switched solid-state molecular (POSSM) pump for automobile air conditioning (A/C).

POSSM will eliminate the related mechanical compressor inefficiencies due to energy conversions (combustion to mechanical rotation) and associated friction and dissipation (heat transfer). Instead, POSSM will operate by rapidly cycling the adsorption affinity of a thin layer of polarizable molecular sorbent arranged in a capacitor with similar volumetric pumping rates to traditional A/C and refrigeration/freezing systems, while consuming less than one-third the energy.



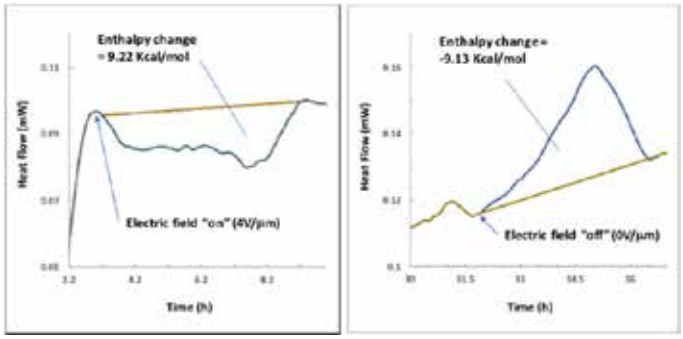
Polarization sorption pump operating principle.

First and foremost, we have demonstrated, in triplicate measurements, that a chromophore-coated capacitor exhibits an enthalpy change in presence of a refrigerant gas. No such effects were observed without the presence of both the chromophore and refrigerant. This demonstration is a key accomplishment that proves the feasibility of the principle that underlines the project.

Our first step was to demonstrate the functionalization of electrodes in a capacitor with a commercial chromophore. The application of an electric field to a double-layer capacitor where one of the electrodes is coated with the chromophore showed that a higher applied voltage is required for a breakdown of the capacitor dielectric compared to the same capacitor without the chromophore coating.

Next, we carried out calorimetry analysis on an operating capacitor to quantify enthalpy changes upon adsorption-desorption of refrigerant molecules at the capacitor.

Four experiments were performed, including three control experiments where a capacitor was subjected to an applied electric field ("on" position) followed by switching the electric field "off." The heat of adsorption and desorption was measured on four different capacitors as follow: 1) bare capacitor in the presence of nitrogen, 2) bare capacitor in the presence of refrigerant, 3) chromophore-functionalized capacitor (one electrode) in the presence of nitrogen, and 4) chromophore-functionalized capacitor (one electrode) in the presence of refrigerant. Only the chromophore-functionalized capacitor in the presence of refrigerant showed a reversible enthalpy change upon switching "on" and "off" the applied electric field.



Refrigerant adsorption and desorption on a double layer capacitor upon switching "on" and switching "off" an electric field, respectively.

PN17014/2904

Demonstration of the adsorption-desorption enthalpy change is a key concept-proving accomplishment. However, there are challenges that remain. For example, the exact amount of the chromophore that has been deposited onto the capacitor plate surface is difficult to quantify, which renders the calorimetry data semi-quantitative.

For precise determination of the adsorption enthalpy, a better method of chromophore coverage measurement will be required. So far, we have used a microbalance technique to determine the mass of the deposited chromophore, but it has its limitations outside vacuum systems and appears to produce mass values that are underestimated. We are currently performing additional experiments to determine more accurately the chromophore mass deposited at the electrode. Also, we are debating about the reasons why the processes of refrigerant adsorption and desorption, though occurring in a similar time interval, are slow. It may have to do with refrigerant transport being the rate-determining step. Work is ongoing.

Next fiscal year will focus on developing a capacitor with higher surface area to attain a chromophore capacity of +10 wt%. A porous metal foam electrode will be functionalized with the chromophore and the mass capacity will be obtained by gravimetric analysis.

The next goal is to quantify the refrigerant sorption capacity of the device with a target of 1 mol of refrigerant per litter of active material (chromophore). As described in the proposal, typical benchmark compressors pumping rate is 0.23–0.52 mol s^{-1}. Hence, for a capacitor containing 0.2–0.5 L of chromophore, and provided it can cycle at 1 Hz, it will require a volumetric capacity of 1 mol of refrigerant per liter of chromophore (active molecule).

Using *In Situ* Liquid Secondary Ion Mass Spectrometry and *In Situ* Transmission Electron Microscopy to Determine the Mechanism and Kinetics of Lithium Ion Mobility in Solid Electrolyte Interface Layers

Zihua Zhu

We are using the unique capabilities developed at PNNL, including* in situ *liquid secondary ion mass spectrometry (SIMS) and* in situ *transmission electron microscopy (TEM), to resolve critical questions regarding how lithium ion (Li-ion) migrates through the solid electrolyte interphase (SEI) under an electric field, as well as to provide a quantitative measurement of its mobility.

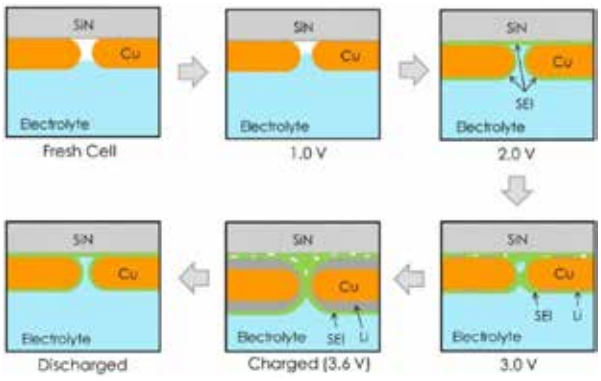
Li-ion batteries are now indispensably used as energy storage devices for portable electronics and electric vehicles and are starting to enter the market of renewable energies. The rechargeable capacity and battery life depend critically on the structural stability of the electrodes themselves, the electrolyte degradation rate, and the electrode-electrolyte interaction layer: the SEI. Processes that occur at the SEI in batteries are critical to battery lifetime and performance, but their details remain elusive because of the difficulty of examining the interface during battery operation. Although electron microscopy and magnetic resonance approaches have provided new insights, isotopic and time-resolved measurements using our novel *in situ* SIMS in combination with *in situ* TEM offer the potential to provide an unprecedented level of detailed information about mechanistic and dynamic processes.

For the SEI layer in rechargeable batteries, two sets of essential questions need to be answered. 1) How do Li^+ ions migrate through SEI? Does Li^+ need to pass through the SEI via tunnels? Or is their migration similar to proton migration in water (i.e., the "Grotthus transport" model)? 2) What is Li^+ diffusivity and conductivity through SEI? Finding answers to these critical questions will enable further advancement on the designing of rechargeable Li-ion batteries for enhanced performance.

To answer these questions, we need first to elucidate the formation mechanism of the SEI layer. Actually, many basic questions about SEI layer *in operando* have not been answered, despite previous efforts. For example, when does an SEI layer start to form? When does Li metal start to deposit on the anode surface? What is the thickness of the SEI layer at operational condition? What are major chemical components in the SEI layer?

During FY 2016, we focused on seeking answers to questions related to the formation mechanism of the SEI layer. The model system is composed of a Cu metal anode with lithium bis(fluorosulfonyl)imide in dimethoxyethane (LiFSI-DME) as an electrolyte, which is a very promising candidate of the next generation of Li-ion batteries.

Our results show that the SEI layer starts to form at about 1.5 V and almost fully forms at 2.0 V. At 3.0 V, a small amount of Li metal already deposits on the Cu electrode surface. Upon charging to 3.6 V, the Li metal layer continuously forms on the Cu electrode surface. After discharging, the Li metal is stripped from the electrode surface, but the SEI layer still stays on the Cu surface. A very interesting observation is that the components and thickness of SEI after discharging are very similar to those of the SEI layer upon charging to 2.0 V, indicating the SEI layer almost fully forms at 2.0 V.



Results showing the formation of the SEI layer at different levels of charge.

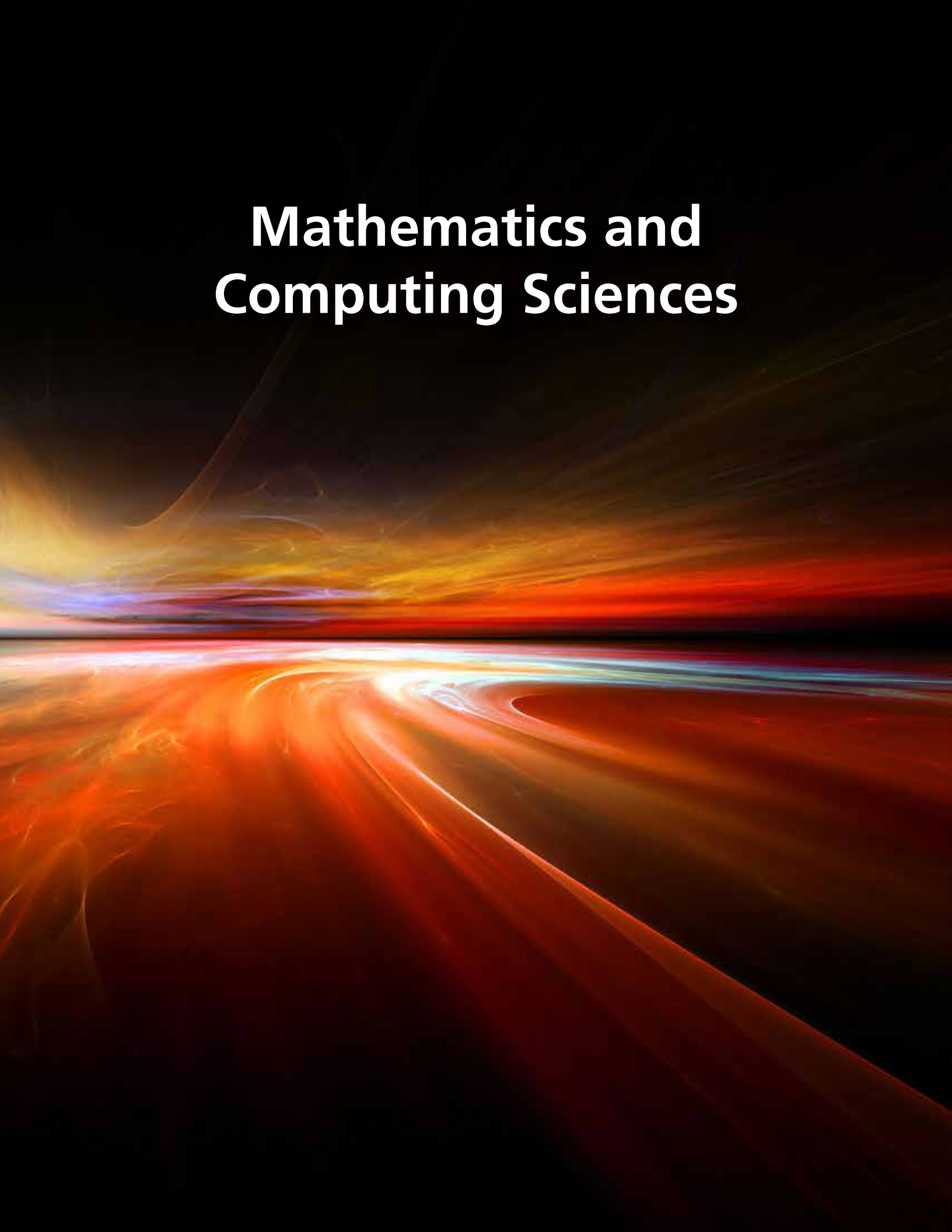
PN16009/2786

In addition, the thickness of the SEI layer is about 10 to 15 nm *in operando*. More interestingly, lithium oxide and lithium hydroxide are found to be the major components in the SEI layer, but the concentration of lithium fluoride is unexpectedly low. Our computer simulation results show that the low concentration of fluorine in the SEI layer can be attributed to the formation of an electrical double layer at the electrode-electrolyte interface. The above findings provide key insights for designing the next generation of Li-ion batteries.

In FY 2017, for the first time, we successfully examined Li^+ diffusion mechanism and determined the Li^+ diffusion coefficient in SEI. Our data show that 70–80% of Li^+ ions in SEI can be easily replaced, suggesting the "Grotthus transport" model may be more possible than the single Li^+ pass-through model. In addition, our data suggest that Li^+ diffusion in the SEI layer is much faster than Li^+ diffusion in glass and in lithium oxide. Based on this, the SEI is more like a solid than a liquid.

In FY 2018, we plan to complete the integration of experimental results with computational modeling. At the same time, we plan to extend the work with the development of new capabilities, featuring an innovative integration of the microfluidic cell in the *in situ* liquid SIMS to enable heating/cooling between 0–80ºC. This innovation will be of great fundamental and practical interest.

Mathematics and Computing Sciences



A Composable Interdependence Model for Cyber-Physical Systems

William B. Daniel

More and more, the monitoring and control of critical infrastructures is facilitated by network-based supervisory control and data acquisition systems that may fall victim to inadvertent network outages or intentional manipulation. This work will allow researchers and utilities to rapidly identify and assess systemic vulnerabilities in order to develop more robust systems and operational practices and serve critical public demands.

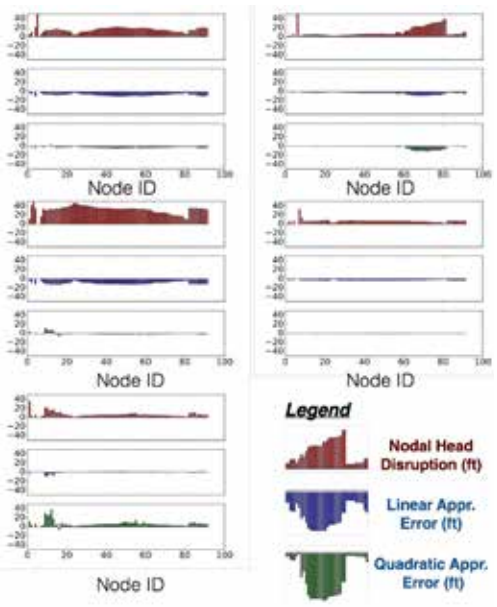
Past work in this field has tended to focus heavily on integrated cyber-physical system analysis. Detailed modeling has been performed of cyber networks, their software, packet-level network communications, and programmable logic controller vulnerabilities. Many commercial and research models also exist and have been used to characterize the consequences of specific contingencies. Less work has been done to effectively bridge the gap between the two areas of research, leveraging the potential for cyber-based observation and control to direct contingency analyses. The technical objective of the project is to provide researchers with novel tools to fill in gaps in infrastructure network datasets, rapidly assess the impact of individual contingency events to those networks, use notions of observability and controllability to bound the distribution of cyber threats, and identify, as a result, the most relevant contingencies and the distribution of their potential impacts.

The critical, water distribution infrastructure sector is being used as a concrete application area within which to test theoretical results and demonstrate capability. Two different lines of work are being pursued within this context: technical and theoretical. On the technical side, an industry-standard model of water distribution system performance, EPANET, has been wrapped in a Scala-based modeling framework to facilitate the analysis of contingency events. This distributed framework makes it easy to access and scale model runs across computational nodes as needed. It also makes it easy to integrate the EPANET model with other models.

On the theoretical side, a method has been developed to rapidly estimate a water distribution system's response to alterations in component states like tank or reservoir head levels, valve settings, and/or changes in pump operation.

Historically, the analysis of a single contingency (a set of coordinated changes to component states) has required the full solution of the network head and flow equations using an iterated solver. Each such solution requires tens of matrix inversions, and for each contingency to be analyzed, the process must be repeated in full.

The approach developed within this project, instead, first performs a single-step solution of the system state under normal operating conditions. A second-order approximation of the head sensitivity to changes in component state is then used to assess the potential impact that contingencies may have on network behavior. Each such analysis requires only simple matrix multiplications. No additional matrix inversions are necessary. This allows the time required to analyze a particular contingency to be reduced by orders of magnitude. An initial analysis comparing the accuracy of the second-order approximation to the full solution has been performed for a small, 92-node system.



Comparison of first and second-order forecasts for head disruption against EPANET simulation results, given 40% tank head-level disruptions. The five contingencies correspond to the individual manipulation of the head level in each of a 92-node system's five different storage tanks.

We have generalized the method to account for two critical features in nonlinear infrastructure models: 1) switching dynamics of time-varying models that trigger on or off states in critical control

PN16016/2793

devices and pumps, and 2) conditional control statements of pumps, valves, and tanks. The switching nature of the models was neglected in previous analysis methods. We have developed a new algorithm that accurately predicts the effect of a contingency scenario even when the underlying system states are subject to switching and control.

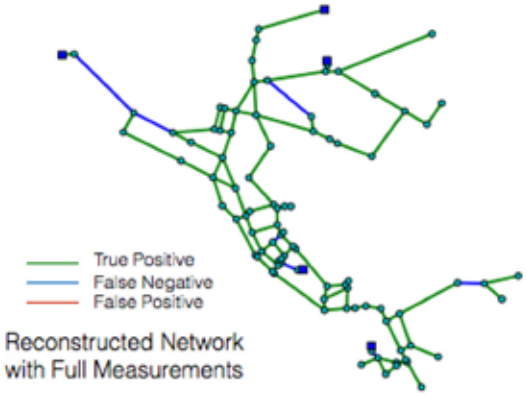
Building on this method, we developed an analysis pipeline integrating industry-standard EPANET solvers, implemented in C, with our one-shot analyses routines. We benchmarked the algorithm on a publicly distributed model of a water distribution system, Mesopolis, with 1,581 nodes, 2,133 pipes, 14 tanks or reservoirs, 69 pumps, and 22 valves. Prior efforts to analyze Mesopolis have been restricted to open-loop models where control is absent. Our new analysis framework allows us to directly analyze systems with switching and discrete control statements. Our algorithm takes approximately 300 minutes to generate predictions of the head disruption for all contingency scenarios—a computational speedup of more than 3 orders of magnitude over the 277 days required by exhaustive simulation.

A second technical achievement is the development of device models to articulate the interdependency between power and water systems. Power distribution (and transmission) systems deliver power to fixed and variable frequency-drive pumps. Previous efforts to model interdependency between power and water systems lacked explicit characterization of the physical dependence; we developed a model that captures the explicit effect of fluctuations in power delivery (as a source effect from the power grid) on realized rotational speed of frequency-drive pump devices in water distribution networks. This capability enables a rigorous interdependency analysis between power and water systems, enabling discovery of vulnerabilities that stem from system interdependency.

A third activity is the development of network inference methods for inferring topology, parameters, and dynamics of distribution network models. We considered regression methods, as well as an exhaustive topology search approach, enumerating all potential graph topologies. We found that regression methods ultimately require a minimal amount of data in order to ensure identifiability of the model; exhaustive topology search, while inefficient, requires far less data to infer a good estimate of the network topology.

We developed two algorithms for inferring a network model. First, an algorithm that ingests full measurements of head levels, flows, and head-loss terms and infers the connectivity of the network from the data. Second, an algorithm that ingests partial measurements of head levels, flows, and head-losses for a subset of the links. Both algorithms are capable of recovering a network model from data. The partial measurement algorithm infers a coarse-grained network, at the resolution of the data provided.

Over the coming year, this approach will be evaluated in the context of larger, more complex networks and via additional metrics, such as the spatial extent of impact areas. Additionally, the work will be leveraged to develop a network inference framework to allow researchers to fill in gaps in network topology data. This will allow researchers to better assess potential system behaviors for networks where complete information may not be available.



The network estimate for an EPANET model using nodal head and headloss measurements over 289 timepoints. Nodes are plotted as blue squares for tanks and cyan circles for nodes. Edges are plotted as true positive (green), false negative (blue), and false positive (red), as estimated by the least squares algorithm. Both the least squares and exhaustive optimization algorithm achieved a 96.2% edge sensitivity and 99.9% edge specificity.

PN16016/2793

Application-Hardware Codesign for Post-Moore's Era: A Study of Application-Specific Hardware Design for Computational Chemistry

Antonino Tumeo

In the context of the Beyond Moore Challenge (the increasing difficulty in scaling feature size of transistors and, consequently, the impossibility to obtain faster and more power efficient processors) we explored the possibility to automatically generate custom accelerators starting from specification in high-level languages such as C and Fortran. The objective is to demonstrate that we can continue increasing efficiency of circuits by employing a sea of application-specific accelerators, while at the same time reducing the development times of such accelerators.

Conventional fabrication technologies for integrated circuits are soon coming to an end. The path toward the 7 nm and 5 nm technology nodes has already slowed down. At the same time, however, less advanced technology nodes are becoming more affordable and reliable. In a post-Moore's law scenario, with technology research focused on searching for physical structures other than transistors, research is increasingly focused on the use of hardware designed for specific tasks in a target workload to maximize performance and power efficiency. For example, the performance improvements in Intel Kaby Lake primarily stem from the use of hardware decoders for video processing; cell phones and tablets consist of multiple application-specific integrated circuits (ASICs). While this is being extensively employed in other domains, use of custom architectures to address scientific problems is only recently being considered. However, the development of custom accelerators is difficult and time-consuming, so automated approaches to generate efficient custom accelerators starting from the high-level specification (i.e., code) of the applications are required. We propose to address the problem of application-hardware codesign in the context of computational chemistry. Our close collaboration with computational chemistry enables us to deeply understand the application needs and design effective hardware solutions. Specifically, we will prototype and demonstrate Field Programmable Gate Arrays (FPGA) solutions for key kernels, to evaluate the challenges faced and the potential performance and power efficiency benefits. FPGAs enable prototyping, evaluation, and validation at a significantly lower cost than custom ASICs. Hand-in-hand with the demonstration, we will develop models to estimate application performance under various architecture considerations. This will build on our in-house expertise in hardware design automation and performance modeling.

Quantum chemistry kernels are both compute (high floating-point intensity) and memory bound (high quantity of data to load from memory). The coupled cluster kernels, in particular, perform large dimensionality (N^7 for CCSD[T] – Non iterative) tensor contractions, leading to a significant number of Double precisions Generalized Matrix Multiplications (DGEMM). However, both the computational part and the data movement (memory access) part can be significantly optimized, depending on how the tiling (division of the tensors in tile for the execution of the DGEMM) is performed and the sparsity of the tiles themselves (sparse GEMM has significant impacts on the memory subsystem due to irregular data accesses).

Considering the fact that FPGAs are less effective for floating-point computation, we have focused our attention on the parallelism extraction and the optimization of irregular memory accesses. We have based our work on previous important additions we made to the Politecnico di Milano Open Source High Level Synthesis tool "Bambu". Bambu enables us to automatically generate the description in a hardware design language (such as Verilog or VHDL) of an accelerator, starting from the description in a high-level language such as C or C++. We have previously extended the tool to integrate architectural templates that better support functions and kernels with high task parallelism and irregular memory accesses. In particular, we have integrated an approach based on a parallel distributed controller to generate accelerators composed of parallel kernels. These could be, for example, tasks extracted from parallel iterations of a

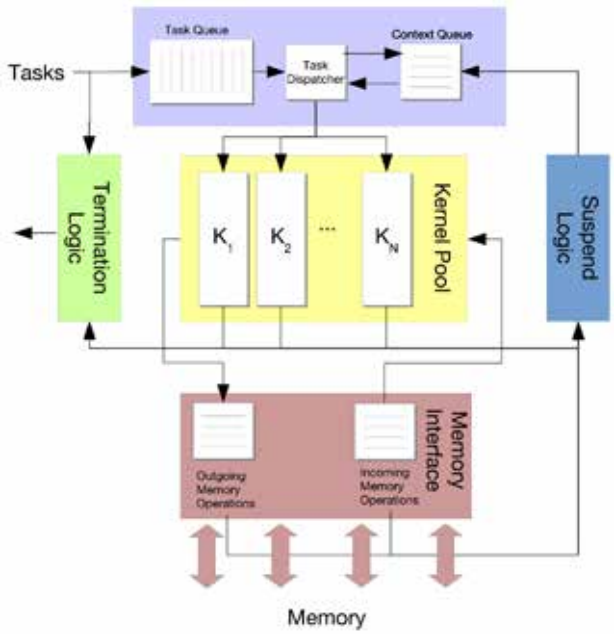
PN17066/2956

for loop. If iterations are data dependent, loop unrolling may not provide the same opportunities to extract memory parallelism.

On the other hand, task-level parallelism may provide more opportunities for optimizations. Activation of each kernel is performed with a data-flow approach through a token passing mechanism. To enable managing complex memory subsystems, we have also integrated a memory interface that dynamically resolves addresses, associating accesses from each hardware kernel to the memory banks where data are located. This relieves the synthesis flow from performing complex analysis of the data accesses. The interface also allows for implementing atomic memory operations. Finally, we have also integrated a dynamic task scheduler that, instead of block-based task scheduling, is able to start a new task as soon as one of the parallel hardware kernels is available.

During this project, we have further extended these designs, integrating support for task suspension. Our approach, previously, only supported spatial multithreading: the throughput, in terms of memory operations, was only increased through the instantiation of additional parallel hardware kernels. The revised approach, instead, also enables suspending tasks that are waiting for the completion of a memory operation and starting a new one to keep other resources busy. When the memory operation completes, the suspended task becomes again executable and can restart whenever a hardware kernel becomes ready. In practice, this approach allows supporting temporal multithreading on custom-generated data paths, providing the throughput required to keep busy a complex memory subsystem.

The approach required further decoupling of the data path from the memory subsystem, with queues for storing and associating pending memory operations, and enhancements to the dynamic task scheduler to store contexts of suspended tasks while other tasks are launched. While, given the memory latencies, the throughput is comparable, we can see that the overall design, with double the number of tasks, only requires 71% more slices (instead of double). In particular, the number of LUTs (used to implement the switching logic) is the current limiting factor, rather than the registers required to store the additional contexts.



High-Level overview of the new architectural template. Queues for suspended tasks' contexts outgoing and incoming memory operations are clearly visible.

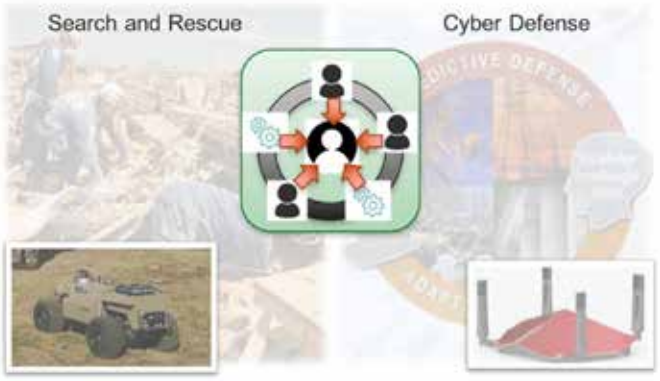
Channels – Kernels – Tasks	4 – 4 – 4	4 – 4 – 8	Ov.
LUT_FF_PAIR	3140	2994	0.95
SLICE_REGs	6869	4174	0.61
SLICE_LUTs	8568	17825	2.08
SLICES	3296	5643	1.71

The synthesis results of an architecture performing a parallel data structure walk with four memory channels, four parallel kernels, and eight tasks, compared to the old design able to concurrently manage only four tasks.

BIFROST: Bounded Informational FRamework to Optimize Streaming sysTems

Luke J. Gosink

Game theory is the study of mathematical models of conflict and cooperation between intelligent rational decision-makers. This work extends game theory to systems defined by both rational (machines) and irrational (humans) decision-makers to form expectations on winning strategies, game equilibria, and characterize fairness of game rules.



The synthesis results of an architecture performing a parallel data structure walk with four memory channels, four parallel kernels, and eight tasks, compared to the old design able to concurrently manage only four tasks.

Formally defining optimal performance for a streaming, complex system of systems is challenging. Operational components in these systems include a diverse array of roles, including human decision-makers and analysts tasked with maintaining situation awareness, machine learning methods that are critical for timely analysis tasks, and sensors that sample the environment based on different sensing and sampling modalities.

These components must complete their individual tasks under the constraints of shared, limited resources (e.g., compute resources, bandwidth, cache, available information, etc.). To ensure that such systems run optimally, control and decision frameworks must identify the best set of actions (from the space of all possible actions) for sharing resources and completing tasks.

This project will model the analysis processes and potential actions performed by such complex streaming systems as an Informational Theoretic (IT) problem. In this framework, each process is treated as an agent in a non-cooperative game. Bounded rational game theory asserts that each of these agents is fully "rational" and aware of the possible actions of every other agent, as well as the penalties and rewards of these actions. Based on each agents' assured rationality, this theory ensures the existence of an optimal, stable state for the system as a whole (i.e., the Nash equilibrium); any alternate state is guaranteed to result in degraded performance of the system.

BIFROST extends these notions to address the challenges posed by agents that may not act rationally (e.g., humans under stress). Specifically, this BIFROST addresses the theoretic difference between uncertainty and irrationality. In the first case, uncertainty about an agent's true goals, motivations, and strategies can make decisions appear irrational, when the error is actually in the way that the agent is modeled in the system. BIFROST is evaluating the hypothesis that machine learning methods can better account for these uncertainties and eliminate seemingly irrational behavior. If this hypothesis holds, many systems that cannot currently be modeled will be able to be evaluated in order to identify winning strategies and the set of optimal choices for system equilibria.

If this work is successful in casting the streaming analysis problem in an IT framework, we will be able to analyze many types of (currently unaddressed) systems and provably define the best course of action for each component in the system. More specifically, BIFROST's framework will be able to identify how humans and machines will be able to best work together in a streaming, complex system of systems.

During FY 2017, BIFROST completed the mathematical formulation of its theoretic framework. This formulation is fully coded as a MATLAB™ library and is being used to evaluate a suite of test systems. The empirical results of these tests, as well as the mathematical proofs supporting the BIFROST approach, will be released by the end of FY 2017.

PN17065/2955

Concurrent Design and Control of Complex Systems: Controllability, Observability, and Performance Metrics

Arun Visweswara Sathanur

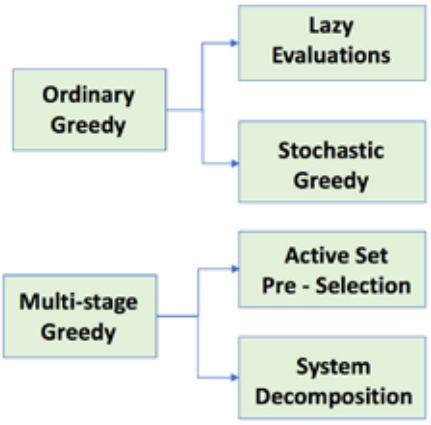
The principal goal of this project is to develop scalable methods for achieving design of complex systems by* emphasizing the discovery of key entities *in systems that have a greater than average effect on the system performance.

The smooth day-to-day living of the nation's inhabitants relies on continual and seamless operation of several critical systems such as the power grid, communication infrastructure, transportation systems, and others. Robust design of these systems with respect to observability, controllability, and other important design criteria is critical to keep the operation as close to the ideal scenarios as possible.

Example problems that we examined in this project are as follows: 1) sensor placement for maximal observability; 2) actuator placement for controllability; 3) mining for key entities in networked systems, such as cyber, social, and epidemic networks; 4) and defensive islanding in power grids to mitigate blackouts.

Since several of these problems can be cast as subset selection from a given set of entities, while maximizing an objective function, the naïve brute-force approach leads to exponential complexity. Given the combinatorial nature of the problem, non-linear and computationally expensive objective functions, and the limitations with existing methods, our focus was mainly on the development and implementation of scalable submodular optimization approaches to address large-scale real-world systems.

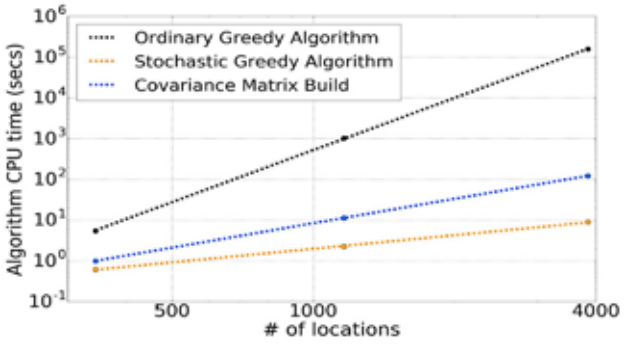
Submodular functions are a special kind of function operating on subsets that satisfy the property of diminishing returns (i.e., an element is worth more when it is associated with a smaller subset than with a larger subset). When a function is submodular, its maximum value can be found by the application of straightforward algorithms such as the greedy algorithm. For many problems, as the complexity of the computational oracle increases, greedy algorithms do not scale well. As a result, we began exploring a range of variants of the greedy algorithm to render this approach scalable.



Different variants of the greedy algorithm for maximizing a submodular function with cardinality constraint.

Sensors and actuators, along with wireless and wired communication infrastructure, will play a big role in the next phase of automation as facilitated by the paradigm of Internet of Things. Monitoring quantities of interest over large spatial regions and over large complex networked systems at a high enough resolution requires placement of sensors in an optimal manner so as to ensure maximal observability.

In general, sensor placement is a combinatorial optimization problem but has been solved in literature via submodular optimization using information-theoretic objective functions. Given that the objective function is computationally expensive, we applied several strategies to achieve scalability while maintaining sufficient accuracy.



Sub-sampling in both the greedy algorithm and the objective function computation can provide dramatic speedups without sacrificing the accuracy.

PN17042/2932

In many complex networked systems, such as online social networks, activity originates at certain nodes and subsequently spreads on the network through influence. In this work, we considered the problem of modeling the spread of influence and the identification of influential entities in a complex network when nodal activation can happen via two different mechanisms. The first mechanism of activation stems from factors that are intrinsic to the node. The second mechanism comes from the influence of connected neighbors. Our model can be considered as a variation of the independent cascade diffusion model. We sketched a proof of the submodularity of the influence function under the new formulation and demonstrated the same on large graphs, including a real-world Twitter sub-graph.

Our future work is expected to revolve around three main themes: 1) scalability of submodular optimization approaches with Matroid constraints that allow us to tackle a different class of problems, 2) focus on applications with respect to the power grid, and 3) influence maximization algorithms for coupled networks.

Co-Simulation Platform for Rapid Prototyping of Control Algorithms

Jeffrey A. Daily

This project is developing a unique tool/ infrastructure that will position PNNL as a leader in modeling complex systems that evaluate controls for the interaction of power grids, buildings, transportation, and related domains.

The ongoing modernization of power grids consists of integrating with communication networks to achieve robust, resilient control of grid operations. The typical applications of this emerging smart grid, such as demand response, dynamic reconfiguration, and distributed energy resources, employ control systems that operate on mission-critical devices.

Beyond the smart grid, control systems are used in additional domains, including building management and transportation systems. As such, these applications of control systems need to be tested rigorously before deployment. Developers usually do not have access to a large test bed; therefore, simulators are used for testing control system applications at scale.

There is a dearth of simulation software available for the intricate coordination and control of power grids, building management and automation, markets, and transportation infrastructure. Further, there are few (if any) tools for developing sophisticated control algorithms within one or across two or more specialized domain simulators. Cooperative simulation of power grid, communication network, and building simulators, coupled with advanced control algorithm implementations, promises to reproduce more accurately the interactions of these components. Taken together, these domain components exhibit a macroscopic behavior not predictable by examining the individual entities alone. This is a complex system, and there is a need for new methods to control such systems. Co-simulation is, thus, the most cost-effective means of rapid prototyping and testing new control approaches.

The objective of this project is to deliver a general, scalable co-simulation platform for use in various domains. In addition to the direct support of power grid, communication network, building simulators, and advanced control algorithm implementations, our platform supports hardware-in-the-loop capabilities. We are, thus, striving to advance the state of the art in co-simulation environments, focusing on the needs to support developing novel control algorithms using diverse simulator software packages and hardware.

The culmination of our FY 2015 efforts resulted in the release of our new co-simulation software framework: the Framework for Network Co-Simulation (FNCS), version 2.0. FNCS 2.0 features a new publish/subscribe messaging paradigm that greatly simplifies the experimental setup of a co-simulation. Data streams can be abstractly represented and fulfilled by a simple playback of data, a sophisticated simulator software package, or a hardware signal. The entire system is extensible, connecting arbitrary data sources and sinks. These data streams are fully synchronized by the FNCS 2.0 broker application and can be represented at different time scales, as needed.

Many simulation software packages have been modified to use FNCS 2.0. These simulators include MATLAB™ and, by extension, its many packages: MATPOWER, power distribution simulator GridLAB-D™, network communication simulator ns-3, and whole building energy simulator, EnergyPlus. In addition, FNCS 2.0 is fully cross-platform capable, running on Windows, OSX, or Linux systems, or any combination thereof. As a result, we are now well positioned to target the integration of any simulator software on any operating system platform.

In FY 2016, we focused on implementing the first hardware-in-the-loop use case, in addition to developing larger-scale use cases in preparation for FY 2017. Supporting hardware-in-the-loop was part of the FNCS 2.0 design since its inception. To that end, we began investigating the incorporation of OPAL-RT in FY 2015. Specifically, we added the Python and C programming language interfaces to aid in the integration. FY 2017 saw continued usability improvements by adding FORTRAN language bindings.

In FY 2017, we focused on scalability. We were able to combine tens of thousands of simulators together while exchanging tens of millions of data values during each time step. This scale was never before achieved using any co-simulation platform. We achieved this goal by carefully benchmarking FNCS, as well as the simulators, to identify and correct important bottlenecks, such as improving the memory efficiency of GridLAB-D 10-fold. The FNCS software is mission-ready and already facilitating ongoing work to enable a resilient, secure, and controlled power grid.

PN15009/2684

CyberCore

Thomas W. Edgar

We are enabling high-fidelity experiments on complex systems by designing and implementing tools for federating testbed resources from research collaborators for modeling full complexity across large computational platforms and with high quantities of physical equipment.

Industrial control systems (ICS) supporting critical infrastructure and Department of Defense platforms are not resilient to cyber threats. We envision a future where these systems are secure against broad categories of cyber threats through state awareness, resilient design, and automated response. Research and development (R&D) specific to ICS is necessary to improve their resilience. While there are many open questions that need to be answered specific to ICS, the most pressing relate to detection and response. Specifically, it is important to determine highly effective indicators for ICS and the best set of feasible, appropriate, and beneficial reactions for defense.

Modern ICS is quite complex in design, which increases the burden on human operators to comprehend the myriad alerts and indicators representing the activity within the controlled physical process. In similar fashion, current application of ICS cyber security may be overly complex and overwhelming. Even with a centralized interface, constituent components of the ICS are dependent on humans to monitor and make corrections in response to security events.

The recent focus on cyber-physical systems resilience show the need for security awareness based on cyber-related performance data. Ultimately, a comprehensive cyber security strategy may achieve fault detection and diagnosis by characterizing their effects to the overall system control envelope. Alternatively, malicious intrusion can be detected in a more decentralized fashion by monitoring the behavior of individual ICS assets (or small groups of them); potential indicators include their memory usage, packet transmission, temporal performance, etc. Suitable reactions to detected malicious cyber activity might include applying restrictions on traffic flow or disabling advanced (but not strictly necessary) functionality. False positives are an important challenge, as they may limit performance of an ICS for no valid reason if used by an attacker to cause denial of service. A key goal for this R&D is formalizing a detection/reaction framework and prototyping case studies to confirm extensibility.

The overall research objective for the tri-lab team from Idaho National Laboratory, PNNL, and Sandia National Laboratories is to collaboratively develop a generalized cyber feedback loop, focusing on a use case based on thorough analysis of a bounded, yet significant subset of field computing devices (FCDs; the subset is based on critical applications of FCDs in the field). This research approach intends to move from design- and vendor-specific techniques of collecting FCD indicators and executing FCD reactions, to a generalized approach widely applicable to achieving resilience on all FCDs. A methodology of defining and developing a high level abstraction layer that standardizes a detection and response interface from comprehensive FCD capabilities will follow. This is dependent on analyzing specific models of FCDs to determine their unique characteristics for collecting indicators and executing reactions, and subsequently developing a generalized approach for collecting indicators and executing reactions that is more broadly suited for other FCDs. These three main objectives, broken into the tasks below, will govern the overall work of the tri-lab team, which will have regular exchanges to discuss collaborative efforts, examine technical problems, and compare techniques and results.

In the first year of the project, we developed a general list of possible indicators and reactions. From this list we performed a targeted study determining which are possible for an SEL 451 relay device. For each indicator and reaction we developed a plan for how to collect or control it. From this we developed a set of tools to collect measures for indicators and to control reactions. These tools include a Python-based software library for interfacing with the configuration management console of the SEL device, an Arduino-based sensor platform for side channel sensor for monitoring thermal, auditory, and visual signals, and some JTAG tools to interface directly with the hardware.

PN17079/2969

Data-Driven Cities and Regions: Towards a Computational Framework for Urban Science

Robert J. Rallo Moya

This project will develop and prototype a machine learning (ML)-powered computational framework for urban science. We will initiate the exploratory work to design, assemble, and evaluate (as proof-of-concept) a software stack for reproducible urban data science.

Cities are complex "systems of systems" where people interact with structural elements at multiple time and space scales. Data analytics is key to understanding city dynamics and to contribute to more livable and data-driven cities.

Information and Communication Technologies (ICT) enable cities and communities to enhance existing and develop new services to increase their efficiency, reduce costs, and improve citizens' quality of life. The availability of city data, together with current data-intensive processing capabilities, provides new tools to understand and study cities. However, this new scenario requires the development of a data-intensive urban science that, by leveraging advances in big data, data science, and machine learning, will impact and transform urbanism and future cities. We will advance the emerging discipline of *Science of Cities* by exploring new capabilities in Computational Urban Science to support data-driven city management and operation in the context of DOE missions.

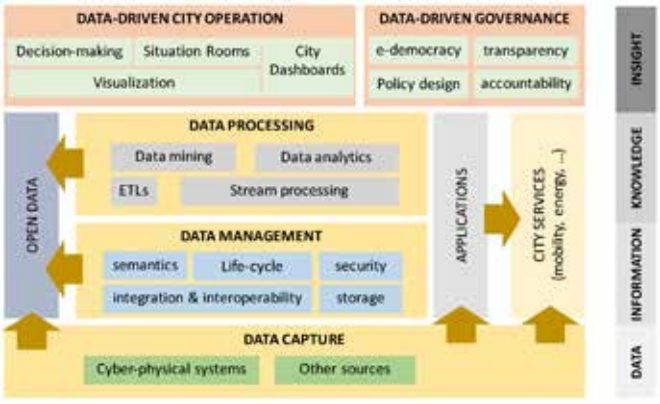
The path toward data-intensive urban science is paved by technological and social challenges, such as: 1) the lack of common standards and reference models to provide a holistic vision of the city as a *system of systems*; 2) the need for scalable urban data management systems; 3) data life-cycle and governance models for cities; 4) policies and technologies to ensure data safety and trust, citizens' privacy preservation, and city resilience; and 5) city-specific data analytics capabilities.

The concept of data-driven cities is founded on the efficient and effective use of urban data and opens a wide range of opportunities, including: 1) new data analytics paradigms for limited resource devices (e.g., edge and fog computing); 2) development of a *CityOS* as a common technological architecture for data-driven cities; 3) platforms for large-scale provenance tracking in urban systems; 4) the use of semantic metadata annotations for data sharing, complex analytics, and decision support; 5) development of tools for knowledge management and decision support; and 6) data-driven tools for open government, e-democracy, and citizens' participation.

A globally accepted definition of a smart city, including their basic components and interrelationships, is still lacking. We have contributed to the International Organization for Standardization (ISO) Technical Committee 268 WG3 effort to develop a *Descriptive Framework for Smart Cities and Communities* (ISO 37105). The framework is a reference model for a smart city based on a systems of systems approach. We have also contributed to the development of a foundation ontology that formalizes the conceptual framework.

Ontological description of the infrastructure components of a smart city.

We have developed a conceptual framework for urban science to support data-driven research in the areas of livability (e.g., sustainability, energy efficiency, environmental quality,



Elements and architecture of the proposed urban science framework.

economic growth, and social equity) and resiliency (e.g., robustness, redundancy, and flexibility).

Data-driven cities are the next step in the evolution of the smart city vision. By leveraging big urban data and existing data science capabilities, data-intensive urban science emerges as a new paradigm to study, understand, and operate cities.

From a big data analytics viewpoint, cities should be studied from a holistic perspective where the city is perceived as a set of interacting and interdependent complex systems (e.g., environment, infrastructures, built domain, and society). The integrated analysis of all city subsystems will capture their interdependencies and interactions, resulting in enhanced insight on city dynamics that subsequently can be used to optimize city operation. The use of big data in cities is fundamental to advance toward cognitive cities where information and ICT resources can be leveraged to develop a new generation of adaptive and self-learning urban systems that will go beyond current city operation capabilities and will contribute to make future cities more livable, sustainable, and resilient.

PN17078/2968

Deception Detection, Tracking, and Factuality Assessment in Social and News Media

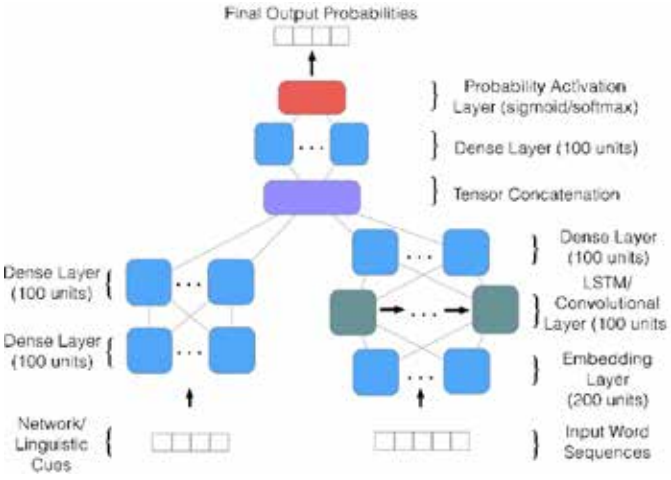
Svitlana Volkova

Unverified news stories in social and online media, ranging from deliberate propaganda to hoaxes, cause a great deal of confusion about the facts of current events. Developing models and tools to classify news posts as suspicious or verified, and predicting different types of unverified posts, is essential to lower the influence and spread of deceptive information in online media. Novel models and key findings of this work will advance understanding of factuality assessment and information propagation online.

The goal of this project is to develop novel deep learning models capable of predicting trusted and deceptive news, as well as types of deceptive news in mixed social streams. Unlike any existing approaches, our models rely on combined representations learned from content, dynamic social network structure, and linguistic markers of uncertainty and subjectivity. Moreover, we are the first to develop models capable of inferring a variety of deceptive news types in many languages, from text and image representations, ranging from satire to disinformation.

In FY 2017, we built and evaluated models to classify social media posts as deceptive or verified, and types of deceptive news. It relies on three types of signals, including tweet content encoded as word embeddings for English and character or byte embeddings for the multilingual case; social network interactions encoded as one-hot or Doc2Vec user similarity vectors, or graph embeddings (e.g., deepwalk); and linguistic cues (e.g., hedging, biased, and subjective language markers. We found that tweet content and linguistic cues are the most effective for classifying suspicious versus verified news posts. However, content and social network signals are effective in differentiating types of deceptive news.

We also performed large-scale analysis of suspicious news spread on Twitter. We found that different types of deceptive news are being shared differently. For example, retweets from accounts that spread disinformation are shared 15 times more by the top 10% of users than by the bottom 40% of users, compared to conspiracy (5 times), propaganda (3 times), and clickbait (2 times), measured using Palma ratio. Gini coefficients that measure the degree of inequality (0 – equality, 1 – inequality) are 0.8 for disinformation, 0.6 for verified news shares, 0.7 for conspiracy, and 0.5 for clickbaits and propaganda.



Elements and architecture of the proposed urban science framework.

In FY 2018, we will further analyze deceptive news propagation in social networks. We will evaluate model generalizability by validating our models on Google clickbait challenge and Reddit news datasets. In addition, we will analyze connotations behind images across deceptive news types and study framing in social media news posts to improve factuality assessment. Finally, we will develop an end-to-end framework capable of making predictions about deceptive news types in real-world social streams.

Predictive results for English models.

English, F1score	Words	Words + network	Words + network + ling. cues
2-way, 130K tweets: suspicious trusted	0.76	0.81	**0.95**
4-way, 65K tweets: clickbait hoax satire propaganda	0.63	0.7	**0.64**
4-way, 2.3M tweets: propaganda conspiracy clickbait hoax	0.67	**0.84**	-
5-way, 2.5M tweets: disinformation propaganda conspiracy clickbait hoax	0.65	**0.85**	-

Predictive results for multilingual models.

Multilingual, F1score	Chars	Chars + network	Chars + deepwalk
4-way, 2.9M tweets: propaganda conspiracy clickbait hoax	0.57	0.71	**0.76**
5-way, 7M tweets: disinformation propaganda conspiracy clickbait hoax	0.55	0.69	**0.76**

PN17038/2928

Deep Learning for Scientific Discovery

Nathan O Hodas

Contemporary deep learning has enabled the next generation of artificial intelligence applications, opening the door to potential breakthroughs in many aspects of our lives. Understanding the capability (and limitations) and improving contemporary artificial intelligence (AI) with application to scientific problems, will enable PNNL to advance the frontiers of scientific research and national security.

There is a revolution happening in AI, with tremendous investments by industry, academia, and government. If we do not invest in maturing our capability now, we will be left behind. With the advent of deep learning, problems previously not considered tractable (e.g., computer vision, automated feature selection, the game of Go, etc.) have achieved unprecedented success, in some cases exceeding human performance. These techniques have just begun to be applied to scientific research. Understanding and improving contemporary AI with application to scientific problems, its limits, and its weaknesses (i.e., how it may be defeated), is of great importance for PNNL's leadership position.

PNNL will apply deep learning across mission sciences, enabling groundbreaking discoveries and making transformational impacts to accelerate innovation and scientific discovery by focusing on scientific achievement, capability development, and mission relevance. This will allow us to better leverage our data, computational resources, and scientific talent to increase our research impact.

We leveraged deep learning in the following five areas in FY 2017: 1) AI adversaries in cyber systems, 2) neutrino physics, 3) integrating scientific knowledge into neural networks, 4) fault tolerance and soft error detection in high performance computing, and 5) integrated 'omics and imaging. Deep learning is accelerating innovation in these five areas of domain expertise at PNNL and methodological advances in deep learning research. We are rapidly maturing our understanding of where and when deep learning is applicable.

We developed AI adversaries using generalized adversarial networks to generate novel attack patterns and more effective defenses, greatly reducing the limitations of existing approaches. We also are improving neutrino event and proton decay characterization in Liquid Argon Time Projection Chambers, leveraging deep learning and the MicroBooNE consortium.

We increased predictive power and efficiency of computational models by incorporating scientific knowledge into deep neural networks by design, specifically the laws of conservation.

In partnership with the Fred Hutch Cancer Research Center, we are using deep learning to predict a patient's outcome through abbreviated breast MRI and patient records, while increasing specificity and maintaining sensitivity.

In addition, we are detecting if an application is sensitive to soft errors and approximation strategies, and understand the performance and accuracy trade-offs. This year we deployed a multi-layer perceptron and engineered features to identify soft errors.

In FY 2018, we will pursue numerous additional seedling efforts, and we will increase the pace and impact of discovery through applying deep learning across many Laboratory projects. These combined efforts will help make PNNL one of the premier laboratories for leveraging deep learning to advance the frontiers of science and security.

PN17086/2976

Development of Computer Techniques for Understanding, Modeling, and Application of Multiscale Phenomena

John T. Feo

The Northwest Institute for Advanced Computing, a joint research institute established by PNNL and the University of Washington (UW), is a center of excellence for understanding, modeling, and applying multiscale phenomena. In this project, PNNL and UW have developed novel computer techniques in biology, energy-aware computing, advanced manufacturing, chemistry, and atmospheric sciences.

This project validated synthetic cyclic peptide designs by comparing experimental measurements of mass and collision cross section (CCS) with predicted values calculated from the designed structure.

Various synthetic cyclic peptides containing 6-10 D- or L-amino acids and designed to adopt a specific structure were obtained from UW. These were analyzed using ion mobility interfaced with mass spectrometry (IM-MS), a method that rapidly (< 1 s) measures the drift time and mass of one or more ions. The IM-MS measurements proved to be invaluable for characterizing mixtures of designed cyclic peptides. In particular, minor conformations revealed themselves, as they have a different CCS (due to different structure) but the same mass. The CCS calculations were within a few percent of the measured CCS for all of the peptides, suggesting they can be used to validate designs.

This project conducted an in-depth characterization of a big data/data analytics software stack to determine the salient features and possible synergistic parameters that can be used across different big data and high-performance workloads. The information gleaned from these workloads can be used to inform allocation of resources within or across the environments.

Using previous experience in runtime systems, the research team instrumented and developed a big data/data analytics solution integrated with the Open Community Runtime (OCR), a high-performance runtime. The group's knowledge and expertise in this area afforded the creation, collection, and understanding of the runtime internals, including how its interactions affect system performance. The team enhanced OCR with several introspective and adaptive capabilities, together with high-performance versions of big data operators based on the MapReduce paradigm. The latter served as an exemplar to help understand workload behaviors on high-performance computing systems.

Using the exemplar with TeraSort and WordCount workloads as the testing framework, the team found that network and data ingestion are prevalent bottlenecks for both workloads. Moreover, when using the adaption on these workloads, network impact was alleviated, but the input/output (I/O) pressure increased (or vice versa) and became the largest bottleneck. As it currently lacks a global I/O view, this adaptive framework remains a work in progress.

The advent of additive manufacturing technologies has introduced the potential to customize designs with different material properties in a single part. This project explored these possibilities by finding the "optimal" configuration of material distribution in a structure. The potential implication is that future designs will not be driven solely by geometry changes, but also by material property distributions. Ideally, if gradients in strain energy are absent throughout a volume, the state of stress will be uniform everywhere with no possibility for stress concentrations.

For this study, we modeled a plate with a hole under uniaxial tension using linear elasticity theory for a two-dimensional (2D) plane stress problem. A prescribed displacement was applied to the right edge, and stress, strain, and strain energy results were compared to those of a model with uniform Young's modulus. Work is in progress to use vat polymerization to enforce changes in the Young's modulus field for a 3D-printed "optimal" distribution of materials properties and to test and validate the optimized design with finite element results.

Our research results and observations support the conclusion that selectively changing E to that of the optimized configuration will lead to significant reduction in high gradients.

PN17074/2964

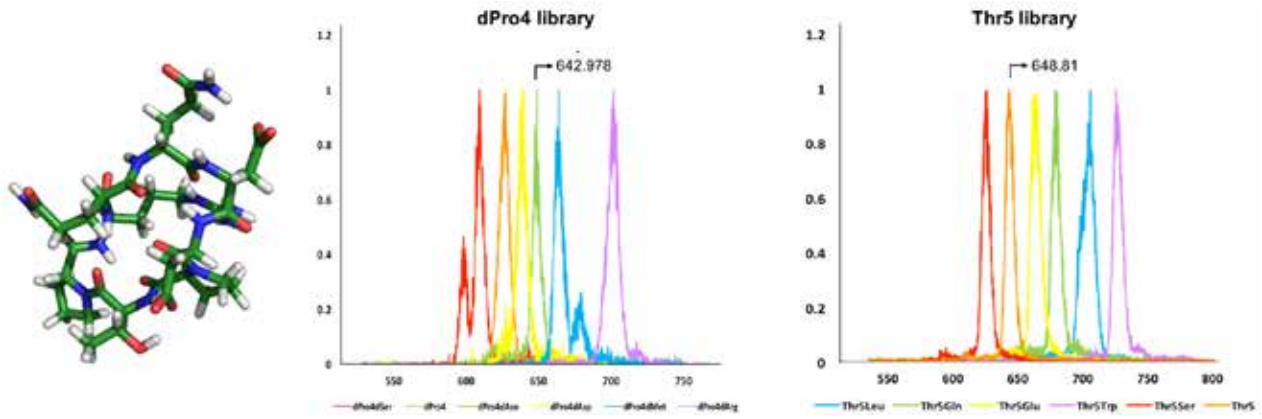
We are developing new methods of 3D printing that allow for variation of material within a solid, such as methods developed by Filisko et al. and Peterson et al. Currently, we are collaborating with UW to print and test preliminary results.

This project also developed methods for using deep neural networks in the "small data" regime to predict complex chemical properties that typically are difficult to model using first-principles simulation-based methods. The SmallNet architecture is a hybrid network design that uses both engineered features and raw data. It has demonstrated superior performance to current state-of-the-art machine learning (ML) algorithms used in biodegradation prediction.

In many domain sciences, the availability of large quantities of high-quality data is missing, owing to time and cost considerations. One example is predicting chemical biodegradability, which has pertinent impact on environmental regulations and planning. Because of the long timescale involved, biodegradability is computationally difficult to predict using simulation-based approaches. This project is examining the effectiveness of deep neural networks to model such complex chemical properties in the "small data" regime.

The "hybrid" SmallNet neural network architecture used > 2,000 biodegradability measurements to train the model. SmallNet differs from contemporary neural network design because it uses both engineered features (computed molecular descriptors) and raw data (2D images) to make its prediction. Results demonstrate that the SmallNet model outperforms current ML algorithms used to model biodegradability in the literature, including a consensus ML model.

METHOD	SENSITIVITY	SPECIFICITY	ERROR
kNN	0.75	0.91	0.17
PLSDA	0.80	0.86	0.17
SVM	0.74	0.91	0.18
Consensus ML	0.76	0.91	0.17
SmallNet	0.81	0.91	0.14



Left, representative cyclic peptide structure. Center and right, ion mobility results. Intensity of observed peaks as a function of arrival time for two crude peptide libraries, dPro4 (center) and Thr5 (right) with point mutations. The original designs are indicated. The dPro4 (dPro4 to dSer, dPro, dAsn, dAsp, dMet, dArg) library clearly has members with more than one conformation.

Development of Physics-Compatible Stochastic Models for Multiphysics Systems with Nonlinear Field Variables

Huan Lei

Our research develops a rigorous mathematical framework for constructing high-fidelity reduced models of large-size, nonlinear stochastic systems, and it has the potential to lead to a deep understanding and control of various challenging multiphysics multiscale engineering and science applications.

The fundamental challenge in quantitatively modeling a multiphysics multiscale system is rooted in the multifaceted nature of these systems. Traditional computational approaches are often developed based on certain physical models on isolated scales or coupled between the phenomenological descriptions through ad hoc scale separation and parameterization. The developed frameworks may show limitations, because no universal field variables and order parameters exist that characterize the states of the emergent phenomena.

To construct an accurate model for such systems, computational frameworks based on mathematically rigorous projection methods have been proposed. However, such frameworks often rely on the assumption of linearized memory term and field variables. In particular, the applicability of the above computational model for nonlinear field variables is unknown in general.

In this work, we aim to fill the gap by developing a novel computational framework for constructing an accurate reduced model with respect to general nonlinear field variables with broad application to modeling of a non-equilibrium multiscale multiphysics system (e.g., nanoscale material synthesis, biomolecule assembly, transport of complex fluid, etc.).

This project began in late FY 2017. Thus far, we have developed the theoretical foundation for constructing the mathematically rigorous reduced model with respect to nonlinear field variables. The major difference between our approach and the previous studies is that the present approach takes advantage of both the mathematically rigorous projection formalism and the cutting edge data-driven parameterization algorithms to construct the reduced model. In particular, we constructed a novel projection operator in terms of free energy function that implicitly incorporates the nonlinear field variables. This operator is different from the traditional linear projection operators defined by Mori and nonlinear projection operators defined by Zwanzig.

The novel projection operator defined above yields nonlocal and nonlinear correlation term, in general. Direct evaluation of the correlation and fluctuation terms involves the evaluation of the orthogonal dynamics, which becomes computationally intractable beyond short times. Alternatively, we developed new numerical methods for constructing the model through data-driven parameterization of the nonlocal correlation term by utilizing the data set collected from the fully resolved systems.

Furthermore, we have also developed the mathematical framework for the constructed stochastic models with nonlocal fluctuations. Due to the non-Markovian nature of the energy-dissipation process, the stochastic modeling equations involve, in general, a colored noise contribution. For our method, we have rigorously proved that the constructed stochastic model shares the same invariant measure with the fully resolved system, which sets the theoretical foundation for studying the non-equilibrium dynamics of such multiphysics multiscale systems.

In FY 2018, we will extend the present mathematical framework to study challenging biomolecule systems with nonlinear field variables such as bond angles and dihedral angles. We will develop the theoretical foundation for studying the transition dynamics for stochastic system with nonlinear field variables and nonlocal fluctuation. This research direction is expected to have broad impact on predictive modeling of multiscale multiphysics systems such as nanoscale hierarchical material synthesis, transport process in porous media, biomolecule and soft matter assembly, and so forth.

PN17098/2988

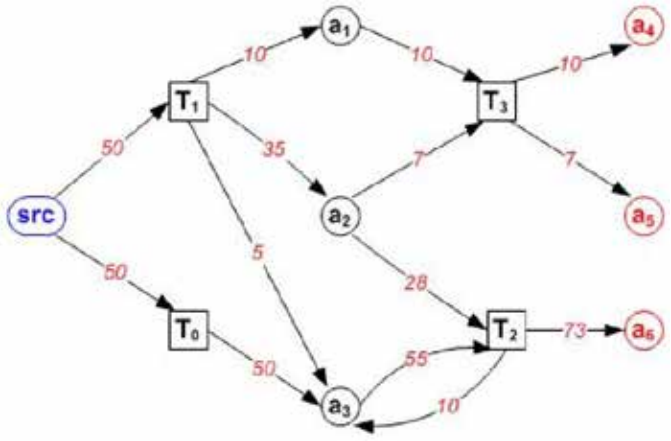
Digital Currency Graph Forensics to Detect Proliferation Finance Patterns

Cliff A. Joslyn

We are providing methodologies for identifying patterns of illicit activity derived from large pseudonymous data networks and establishing PNNL's leadership in modeling global-scale, decentralized ledger systems.

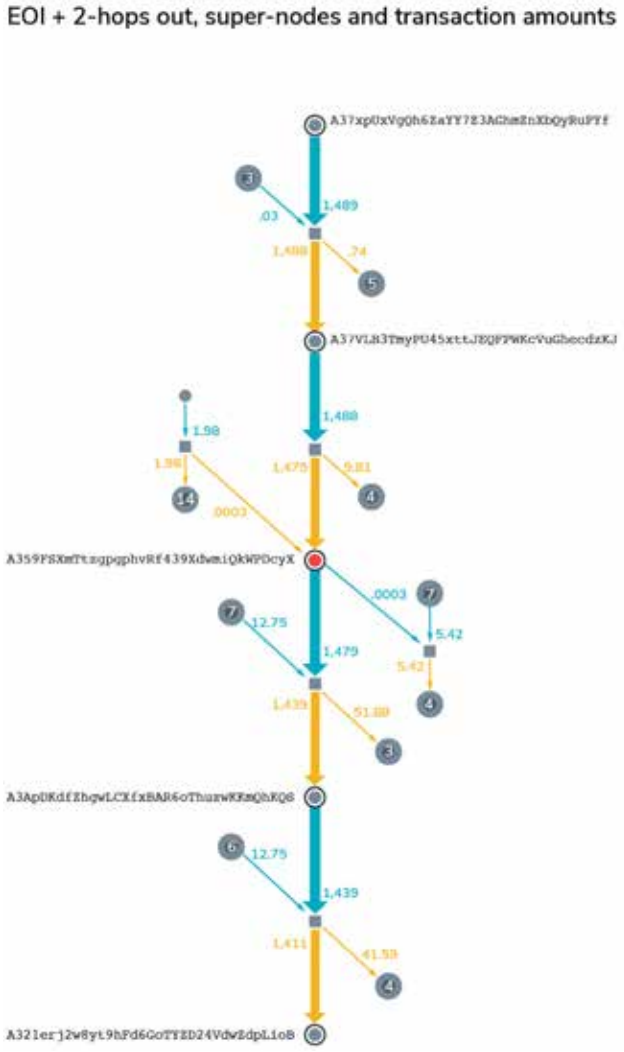
Financial systems have matured through digitization with the rise of digital and crypto-currencies (CCs), the most prominent of which is Bitcoin. The attraction of digital currencies to launder or transfer funds for purposes like trafficking nuclear materials poses severe technical challenges to both technologists and analysts. The ability to identify patterns of behavior in publicly available information derived from large anonymous data networks requires novel technologies for graph data management, pattern mining, and modeling. In this project we examine the open and publically available historical data for CCs transactions (e.g., the Bitcoin Blockchain) to identify patterns of illicit activity.

In our first year, the team initiated a process to identify relevant available financial datasets, including both traditional (e.g., bank or credit card networks) and CCs. We conducted research into various money laundering and money transfer methods and developed semi-formal, conceptual, and/or process models for each. We also started to develop formal mathematical models of Bitcoin transactions.



Bitcoin transaction graphs modeled as a weighted, directed hypergraph (DH); transactions as squares, addresses as circles.

This led us to focus on the structural properties of the network of Bitcoin transactions. While structured mathematically as a graph, where quantities of Bitcoin stored at pseudonymic addresses are connected as inputs and outputs of transactions, reuse of addresses means that a normal graphical structure cannot capture the level of complexity needed. To accommodate this need, we model transactions as participating in the mathematical structure of a DH. Whereas in graphs, transactions must have single inputs and outputs, in DHs, multiple addresses can serve as inputs and outputs, and addresses can be inputs or outputs to multiple transactions.



Forensic Bitcoin exploration.

PN15017/2692

Of particular interest is the identification of motifs in DHs, small DH patterns with particular identifying characteristics, usually statistical. Small motifs of sizes three, four, or five are commonly used in network science, but not yet identified in the more complex DH patterns. We have identified the mathematical properties of DH motifs and some candidate corresponding patterns of potential laundering.

We have more recently focused on Blockchain applications outside of the purely financial. In particular, the Ethereum system provides methods for distributed, secure, smart contracts involving not only financial exchange, but also arbitrary digital work. The potential for systems like Ethereum to provide disruption to a range of legal and marketing applications involving surety and distributed trust is great, and our experience in distributed cryptographic networks is positioning us to be at the forefront of technical developments.

Distributed Deep Learning and System Identification for Community Detection and Classification

Enoch H. Yeung

This project aims to lay the foundation for a revolution in distributed computing using deep learning. This project will develop new methods for learning distributed representations of models for complex and noisy systems.

How do we learn predictive models of dynamic systems directly from data? Real world systems are extremely high dimensional (e.g., social networks, biological and chemical systems, infrastructure, and networked control systems), making it difficult to develop learning algorithms that scale with their complexity. We are developing learning algorithms that generate distributed representations of complex nonlinear dynamical system models from noisy time-series and discrete data.

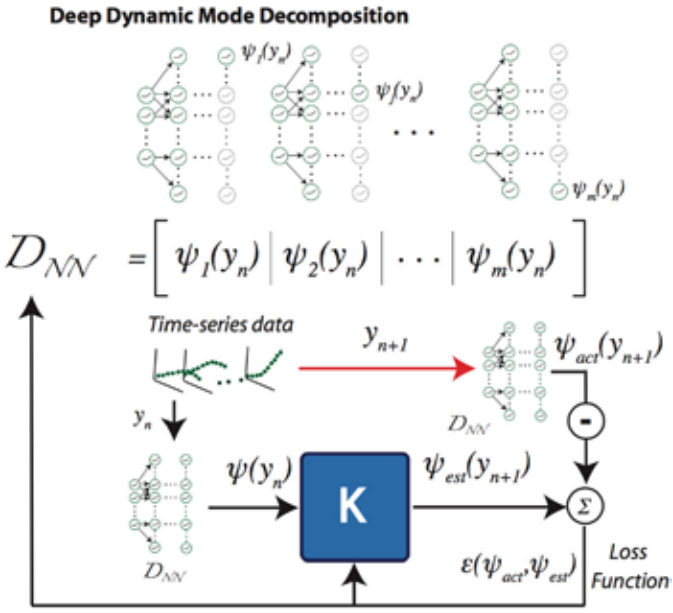
One of the challenges of learning predictive models for nonlinear networks from data is that not all of the variables are measured simultaneously. Indeed, many systems retain hidden variables that encode latent dynamics along network edges. Previous methods for learning dynamical network models assumed linearity of the underlying dynamics. Moreover, these methods struggle to scale for extremely large systems.

We will develop new methods for learning dynamical network models of systems that are both nonlinear and stochastic. In particular, we will focus on learning algorithms that generate distributed deep representations for continuous and discrete state systems that enable faster and scalable training.

In FY 2017, we developed a new method for learning distributed neural network representations for fine-grained classifiers. Fine-grained classifiers are discrete state predictors; they predict whether a data point belongs to one of many classes (e.g., whether a Netflow log is normal, suspicious, or anomalous).

We showed that existing deep learning algorithms fail to preserve statistics relating data features to class. We developed a new training algorithm that identifies which features determine class membership, enabling interpretability of a deep neural network's decision process.

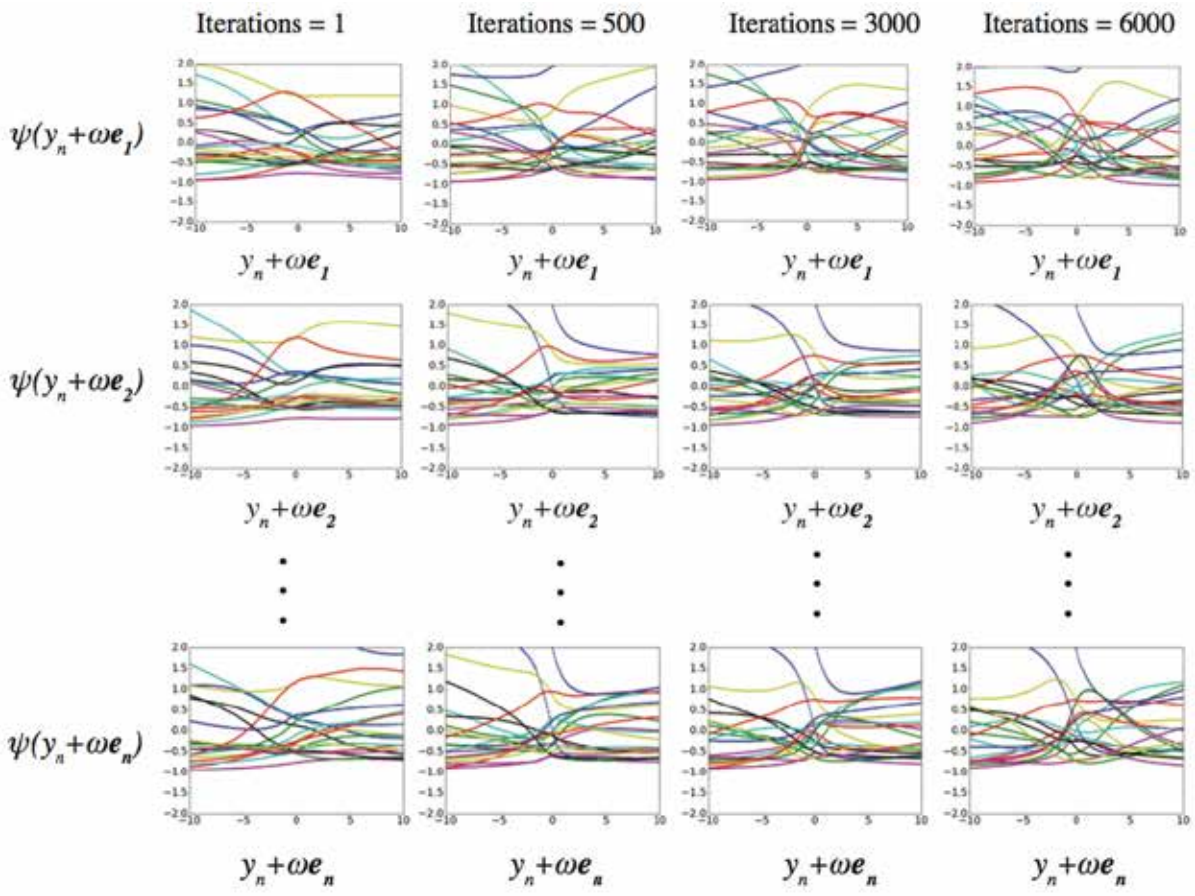
Additionally, we developed a new algorithm for learning dynamical system models directly from data using deep learning. The algorithm learns a dictionary of functions to describe a complex nonlinear process (usually unknown) and automatically updates the parameters and structure of a neural network until the dynamical model is predictive.



A schematic showing our new approach for learning deep neural network based representations of dictionaries for a nonlinear dynamic system. Dictionary functions are randomly initialized and refined automatically using Tensorflow loss functions during training until a predictive model is found.

Finally, we developed a constructive proof-of-concept algorithm that identifies deep neural networks directly from data. The idea is to define a geometric representation of the data that uniquely prescribes parameters, layer by layer, of a deep neural network.

PN17016/2906



The dictionary functions for a 10 layer drop-out exponential linear unit neural network. The x-axis shows the size of the perturbation to a particular dimension of the measured state variables, and the y-axis shows the response profile of the dictionary functions. Each line represents one of 20 dictionary functions.

Going forward, our research will investigate both the scalability and broad applicability of these methods in nonlinear inverse modeling problems, transfer learning problems, lifelong learning, and design of novel computing architectures. We will explore application areas in data mining for news alerts, monitoring and community detection in social systems, inference of natural and engineered biological networks, analysis of infrastructure systems, and anomaly detection.

PN17016/2906

Dorci – The Defender's Role in Resilient Cyber Security

Roderick M. Riensche

Through Dorci, we are building a foundational understanding of the needs and roles of next-generation cyber network defenders who use resilient infrastructures for a more secure network.

Current research has explored defender roles and responsibilities; however, it has been conducted without the introduction of resilient infrastructures. A foundation of understanding roles, responsibilities, and needs of cyber network defenders can be leveraged by new and existing research in resilient infrastructures to increase utility to defenders, minimize complexity, and assist with the adoption of the new technologies. The object of this work is to advance knowledge about enabling defenders to manage resilient cyber infrastructures without increasing their cost (resources) to do so. As architectures, algorithms, and methods are developed to make networks resilient to attack, the network becomes more complex for the defender to manage. If resilient cyber security research is to tip the asymmetric advantage, it must not increase the burden on the defenders who are understanding and managing the network.

In FY 2014, we started by evaluating the importance of a "current state" study in order to provide context for future work. Preliminary questions to assist with the correlation of previously completed research were leveraged.

Initially, we had planned on a scenario-based, semi-structured interview to help users imagine their environment. Next, we explored use of a tabletop excursive within which users would be given roles and responsibilities, activities will occur, and various options for actions taken. We determined that our study should explore resilient infrastructures but needs to be believable and reasonable to defenders today. While the implementation and adoption of resilient infrastructures is not widespread (which makes technologies difficult to map), we had widespread support from other projects to help ground our research in the plausible domain.

During FY 2015, we concentrated primarily on developing a traditional technology study. Key accomplishments for FY 2015 involved design development of human subjects study and game play framework, creation of a study tracker application to track the study during participation, and additional studies, including two industry, one research, and one government study, along with several pilot studies in order to refine gameplay, as well as study effectiveness. All studies were well-received and generated results for traditional technologies with interesting findings. We also brought in a researcher from Middlesex University to assist our team with research questions regarding the next year's resiliency work. With this collaboration, we were able to establish much of the groundwork necessary to build out our planned experiment with resilient technologies.

In FY 2016, we conducted detailed surveys to gauge expert opinion regarding expected costs and effectiveness of resilient technologies. We used this information to update the game definition that was created during FY 2015 and conducted an additional study. From this study, we were able to observe how participants compared new and potentially unfamiliar technologies versus traditional technologies and approaches. By completing the game definition, we have built a convenient and engaging platform that can be used to facilitate structured brainstorming exercises when considering potential impacts of future technologies.

In FY 2017, we focused our efforts on refining the Dorci exercise for use as a training tool. We conducted pilot training sessions with both internal and external users and now have a training capability that can be deployed for a wide range of user expertise levels. Dorci was used as part of a curriculum developed for the National Nuclear Security Agency and stands ready to be used by any other government agencies.

PN14087/2671

DYnamic Network Analysis via MOtifs (DYNAMO)

George Chin Jr.

***The DYNAMO project is developing a graph mining-based approach and framework that will allow humans to discover and detect important or critical graph patterns in data streams through the analysis of local patterns of interactions and behaviors of actors, entities, and/or features. We refer to these local graph patterns, which are small, directed, attributed subgraphs, as* network motifs.**

Dynamic network motif analysis represents a graph-theoretic approach to conduct "pattern-of-life" analysis of agents or entities. Conceptually, this approach equates to identifying small, local subgraphs in a massive virtual dynamic network (generated from data streams) that have specific meaning and relevance to the user and that are indicators of a particular activity or event. Detecting motifs in streaming data amounts to more than looking for the occurrence of specific entities or features in particular states, but also their relationships, interactions, and collaborative behaviors.

Dynamic network motif analysis allows an analyst or scientist to:

- Expose relationships and interactions of an agent with other various articles or entities such as other people, computing resources, places or locations, and subject areas
- Learn normal relationship and interaction patterns of different classes of agents from monitored activities
- Detect when anomalous patterns occur and offer potential explanations
- Facilitate encoding of hypothetical behavioral patterns or indicators for an agent to use in future detection
- Provide a framework for evaluating and confirming multiple indicators
- Analyze temporal evolution of interactions and behaviors of an agent
- Support all the above analytics in the presence of streaming data.

The two key development efforts on this project were to create algorithms to detect approximate motif patterns in a data stream and to construct an interactive framework for defining motifs for querying and monitoring. Network motif mining falls into the class of NP-hard graph algorithms known as subgraph isomorphism or subgraph pattern matching. A variety of subgraph isomorphism algorithms have been published over the last several decades. Graph mining through the prism of small graph structures has also been studied for decades, but has experienced a recent renaissance in research interests in areas and problem spaces such as triadic analysis, triangle counting, and biological network analysis. Subgraph and motif mining in the presence of streaming data adds yet another level of complexity and uniqueness to the subgraph isomorphism problem. The project explored existing research and development in this area and implemented new algorithms to enable dynamic motif analysis.

The dynamic network analysis algorithms and framework were applied to an insider cyber threat use case. One may characterize the type of an insider agent, computer user, or organizational role by examining the modes and frequencies of its local interactions or network motifs within the cloud. For example, a cloud interaction graph may be generated from cloud telemetry data that shows a user's particular interactions with or access to specific tenants, data stores, and applications. In examining such a graph, one can look closely for small, local graph patterns or network motifs that are indicative of an insider cyber threat. Additional motif analysis of proximity card access data may identify changes in an actor's work patterns (time and place) and the people with whom she interacts. Motif analysis of email and phone communications records may additionally identify an actor's interactions with other people and their associated statuses and roles.

The concepts of the network motif and the motif census have some very compelling characteristics for threat, event, or feature detection. They provide a simple and intuitive representation for users to encode critical patterns. They provide a structure for assessing multiple indicators. They enable a mechanism for throttling the accuracy of findings and limiting false positives by adjusting the threshold associated with the computed similarity distance measure (e.g., lower threshold equals higher accuracy and less false positives). Indicators in a motif census may be active or inactive depending on whether data is available to assess the indicator. In the case where active indicators may be alerting a potential insider or critical event, the inactive indicators identify additional data to gather and assess, in order to further confirm or reject the alert. This enables an approach for directing humans or systems to specific areas to look for additional information or data.

PN17056/2946

EvoGraph: Highly Efficient Large-Scale Graph Processing on Accelerator-Based Supercomputers

Leon Song

Efficient large-scale graph processing is a very essential and challenging problem, covering many areas of interest that span from social network traffic analysis and power grid development to biological/environmental research and physics simulation. In this work, we are developing an accelerator-based graph processing framework to address this issue.

High-performance machines are increasingly using graphics processing units (GPUs) to leverage their scalability and low dollar to FLOPS ratios. As a result, GPUs have become the main compute engines for today's high-performance computing supercomputers. This trend continues with the move toward exascale machines that are expected to be composed of millions of accelerators and general purpose cores, whether packaged as "thin" or "fat" nodes. Another recent trend is the gain in popularity of GPU processing in many domains such as social networks, e-commerce, advertising, and genomics. This has motivated the growing interest in large-scale, real-world graph processing for both scientific and commercial applications, as well as the recent efforts in accelerator-based graph processing frameworks.

An important aspect of real-world graphs, like Facebook friend lists or Twitter follower graphs, is that they are massive and evolving. Thus there is a huge need to quickly analyze this high velocity stream of graph data. However, state-of-the-art graph analytics for dynamic graphs follow a store-and-static-compute model that involves batching these updates into discrete time intervals, applying all of the updates to the total graph, and then rerunning the static analysis. There is considerable redundancy and inefficiency in this approach to analyzing this evolving graph sequence.

Static graph analytics on a single version of the evolving graph, even when leveraging a massive amount of parallelism offered by thousands of cores in a GPU, can be very slow due to the large scale of many real-world graphs Additionally, data movement of the entire input graph repeatedly between the host and the GPU over the slow peripheral component interconnect express link can result in substantial overhead, which in turn can overshadow the benefits from the massive parallelism offered by a GPU. To tackle these challenges, we are developing an accelerator-based incremental graph processing framework named EvoGraph.

For FY 2017, we have completed designing EvoGraph, an accelerator-based, high-performance incremental graph processing framework on GPU, processing evolving graphs by avoiding naive static graph recomputation on snapshots. To address the inefficiency of snapshots recomputation approach, we propose the property-guard feature, a runtime heuristic to dynamically execute graphs based on user-defined or built-in graph properties, resulting in high speedup. We also now understand different graph algorithm's merge patterns (e.g., stateful, partially stateless, or fully stateless) for efficient targeted processing in coordination with property-guard.

Specific achievements for FY 2017 include: 1) designing and implementing an initial version of the EvoGraph framework and its components; 2) designing three common graph algorithms based on the Gather-Apply-Scatter model—Breadth-First Search, Connected-Components, and Triangle Counting—which included finalizing design choices for the framework: programming model, data structure, static versus dynamic runtime, and GPU-related optimizations; 3) distinguishing merge patterns of different graph algorithms for more efficient processing strategy—stateful, partially stateless, and stateless—of which initial results are promising, significantly outperforming the state-of-the-art CPU-based STINGER by avoiding recomputation.

We also made several technical contributions. EvoGraph, the accelerator-based, high-performance incremental graph processing framework built on top of GraphReduce, processes evolving graphs by avoiding the naive static graph recomputation approach. It seamlessly maps users' sequential codes for incremental graph algorithms onto GPU for acceleration.

We improved GPU core utilization via dynamic merging of GPU contexts (or context packing) from different graph applications and additional hardware parallelism extracted using deep copy operations on separate CUDA streams to leverage multiple GPU hardware queues.

PN17067/2957

An extensive evaluation of three general classes of graph algorithms (i.e., stateful, partially stateless and fully stateless) on real-world and synthetic graph datasets demonstrates that EvoGraph can significantly outperform the existing static recomputation approaches. Compared to competitive frameworks like STINGER, EvoGraph achieves a speedup of up to 232x and overall throughput of 429 million updates/sec.

We also developed graph-property-based performance optimization called property-guard to dynamically decide between static and dynamic graph execution based on user-defined and built-in graph properties, resulting in a speedup of up to 18.4x over a common streaming approach.

During FY 2018, we plan to explore deep learning strategies on predicting future graph properties to enable effective current computation management strategies.

Fundamental Mathematical Models for Human Interactions

Jennifer Webster

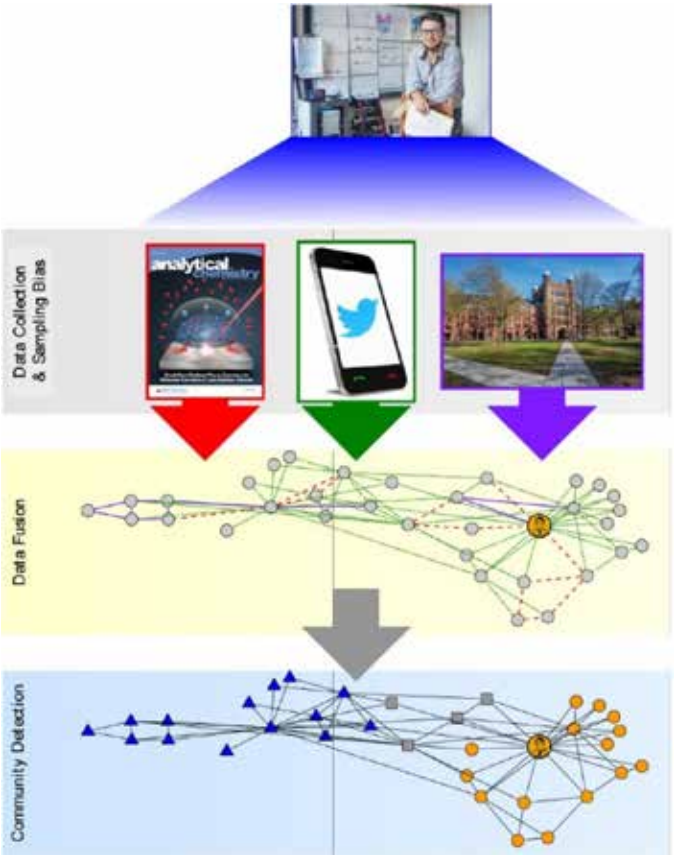
Human community behavior is notoriously difficult to model. There has yet to be a discovery of fundamental principles that govern how humans interact and form communities. This project is looking at mathematical models that address observational and data limitations in discrete systems while also incorporating expert knowledge.

Community detection methods are typically developed on either a theoretical mathematical graph construct with specific properties or a very specific data set. Transferring these methodologies to other real-world data sets can be problematic, as sampling schemes and data collection obscure the true underlying network and introduce a measure of uncertainty into the mathematical formalisms. The goal of this project is to bridge the gap between theoretical development and the application of network analysis techniques. This will involve development of new community detection methods and the expansion of existing algorithms.

This year, we have focused on extending our understanding of two community detection methods. The first method is a continuation of our work from last year, which incorporates expert knowledge of the relationships between communities into the algorithm and facilitates a more precise clustering of individual members. For example, it is common in politics to describe the communities as "left, right, and center," with the understanding that the political identifiers for the left and right are more different than comparisons of either of these groups to the center. Using this knowledge of the communities, we can better identify members of the center group.

Community detection methods like Min-Cut treat relationships between all communities equally and, as a result, frequently misclassify community members of the center as part of one of the other two communities. We have extended our theoretical understanding of this and have also worked on a parallel software implementation that allows us to scale up our analysis and operate on real-world data sets, including a protein structure data set and voting records from the U.S. Congress.

The second methodology, Djikstra Modularity, is a variation on the common understanding of modularity as a goodness of fit metric in network science and combines it with subject matter intuition in that all communities should be connected via message passing algorithms, especially those messages passed by influential persons. Our implementation not only highlights the communities, but also their "drivers" in terms of these influential persons.



Bridging the gap between the theoretical development and the application of network analysis techniques.

PN16008/2785

Geopolitical Discourse Characterization through Deep Learning in Diverse Data Modalities

Michael J. Henry

This project will leverage and extend the latest state of the art in machine learning and deep learning research to develop algorithms that deepen our understanding of complex narratives in social media data streams.

Data analytics, particularly image analytics, has traditionally focused on leveraging subject matter expertise to hand-engineer features to be extracted from data in order to perform the analysis. This methodology has been very successful in a number of domains, but in recent years, the emergence of deep neural networks (deep learning) has produced phenomenal and revolutionary results in a number of tasks, including: object detection, automatic image captioning, sentiment analysis, and topic analysis.

While deep learning has stretched the bounds of what machine learning and data science can do, much remains unstudied about how these algorithms can be leveraged to solve national security, energy and environment, and other DOE-relevant problem sets. In this work, we leveraged deep learning, utilizing and extending the current state of the art, to focus on addressing the complex challenge of analyzing points of view and narratives present in open source data streams.

The first year of this project focused on the following tasks: 1) development of a use case for data analysis, 2) identification and collection of social media data, 3) a study of sentiment and emotion classification within social media as it relates to both users and groups, and 4) an evaluation and integration with automatic image captioning models.

The development of a use case centered on the social media conversation about global climate change. Specifically, we targeted the global climate change conversation on Twitter during December 2015, which coincided with the Paris Agreement, an agreement established within the United Nations Framework Convention on Climate Change that sought to formalize greenhouse gas mitigations within the member states, starting in 2020. This use case was selected because it deals with the public perception of climate change and energy policy. Additionally, a number of unique and interesting narratives could be drawn from the data supporting this use case.

The data curated to support this use case was primarily from two social media sources: Twitter and Flickr. Flickr was selected because the data on Flickr is almost exclusively visual (some text data is available via tags, image metadata [when available], or comments). Twitter was selected as a data source, as it is a popular vehicle for social conversations. Twitter is an open platform and an attractive mechanism by which an individual or an organization can share their point of view with a global audience. In total, 300 thousand images were curated from Flickr, and 8.5 million tweets were curated from Twitter.

We developed a mechanism for determining whether an account belonged to an individual or an organization. We then analyzed demographics, emotion, opinion dynamics over time (1 month), and location to draw correlations between each element, as well as behavioral differences between person and non-person account types. This work provided an in-depth insight into the behavior and opinions around a topic where multiple distinct narratives are present in the data.

Finally, work was conducted on leveraging and extending existing image captioning models. Whereas a typical image classification model will provide a ranked list of labels for an image (e.g., “dog”) describing its content, a captioning model is capable of generating a human-readable, rich description of an image (e.g., “a dog catching a ball in a park”).

Year two focused on the following tasks: 1) improving the image caption model, 2) developing a model for labeling images with a sentiment score, and 3) developing a system by which a user could describe, using text, the content of images that they would like to see, which would then be used to retrieve images that match that description.

Improving the image captioning model relied on leveraging the lessons learned in year one in what makes a captioning model useful. Building upon this, the project team designed a novel encoder-decoder image captioning model that relies on training two deep neural networks: an image feature

PN16051/2828

extractor built from the state of the art in convolutional neural networks and a text generation network built from a fully unitary recurrent neural network. This model was able to train an expressive image captioning model.

The field of text sentiment and emotion classification is well studied, and while that remains an open area of research—as explored in the first year of the project—nonetheless, text sentiment analysis coupled with the multi-modal nature of social multimedia provided the impetus for developing an image sentiment model.

Image sentiment is a difficult challenge, as the concept of "sentiment" in images is less well-defined. To work towards this goal, an attribute classifier was trained. These attributes are adjectives used to describing a given noun. In the model, there are 212 different adjectives across approximately 500,000 images. Those adjectives could then be used to develop a sentiment score using traditional text sentiment methods. The project team used an image convolutional neural network as an encoder, which was then fine-tuned to produce the final attribute classifier. The model trained by the project team was able to achieve a 40.3% top-5 accuracy on a hold-out validation data set.

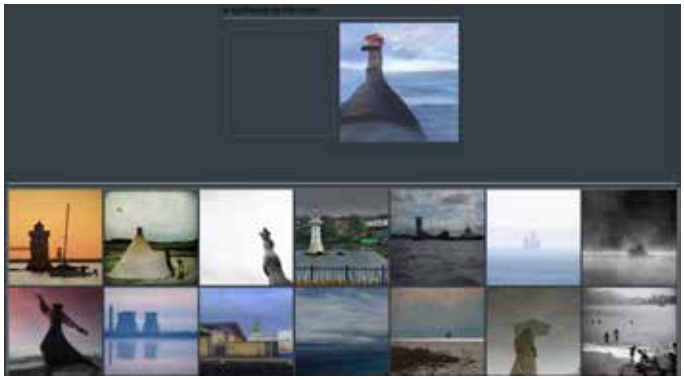
The last thrust of the final year of this project focused on developing an image query-by-example system, but instead of the typical image queries as seen in a traditional image query-by-example system, this work intended to fuse the text and image domain to provide a more advanced interface for image query-by example.

Above are two images that were fed through the attribute mode. This model produces adjectives to describe the contents of the image. The top scoring attribute label for the image on the left is "freezing," and the top scoring attribute for the image on the right is "dark."

In order to accomplish such a task, this work required a model that could "translate" between image and text features, and another model that could produce a suitable query for the query-by-example system.

To interface with the system, the user would enter a text query (for example, "a lighthouse by the ocean"). The model, much like the captioning model described previously (essentially, the reverse of the previously discussed workflow), would extract a feature vector from the text query and use that as a conditioning vector into a model—a generative adversarial network—that could produce a two-dimensional (2D) image illustrating the concept described by the text query. That image would then form the basis for the query-by-example system, as a visual representation of the underlying concept described by the text query.

This work intended to explore the boundaries of what deep learning could accomplish within the context of social data. While the methods described herein have produced novel results, there is much to be yet explored in this rapidly changing field.



Above is an image showing an example of the novel query-by-example system. At the top, the user has entered a query for "a lighthouse by the ocean." The image immediately below it is an image generated by the network, representing the visual concept that the network thinks "a lighthouse by the ocean represents. At the bottom of the image grid is an example of the images that are visually similar to the generated image.

High-Throughput Genome-to-Metabolome Computational Methods for Microbiome Metabolomics and Modeling

Ryan S. Renslow

How do we identify the preponderance (99%) of metabolites if standards do not exist or, critically, if molecular structures are not even known? This project's goal is to develop a high-throughput metabolite-identification approach based on a quantum chemistry, super-computer-driven software engine for application in complex microbiome metabolomics.

Currently metabolomic tools are limited by the ability to build libraries of compounds, which are required to identify and quantify metabolites. The premier nuclear magnetic resonance (NMR) metabolomics tool, Chenomx, has a library of approximately 900 compounds with about 11 new compounds added per year and one of the largest open access databases, the Biological Magnetic Resonance Data Bank, has data for approximately 1,200 compounds. In both of these libraries, most compounds are derived from urine and human health-related samples.

Currently, there is no robust environmental- and microbial community-focused NMR metabolomics library, and well over 90% of the known metabolites in BioCyc are not found in an accessible library. To achieve DOE's goal of understanding, predicting and controlling complex microbial communities, and to understand specific microbiomes undergoing transition, the capability to comprehensively identify and quantify metabolites is necessary.

Quantum chemical calculations have been used to predict NMR chemical shifts and ion mobility spectrometry-mass spectrometry (IMS-MS) collision cross sections (CCSs) of metabolites; however, to date, these calculations have been limited to small sets of molecules. Datasets are typically arbitrarily chosen, particularly in the case of NMR, and most simulations do not emphasize compounds common in environmental sources or microbial communities. Furthermore, most simulations and current databases for metabolomics do not consider molecular conformers or custom solvent conditions.

Our novel computational research will expand the current sets of simulated molecules by one to two orders of magnitude and provide a streamlined process for generating relevant metabolite libraries from microbial genomes (or community metagenomes). Currently there is not a unified method for generating reaction/pathway lists with corresponding metabolites to build a library for identification and quantification.

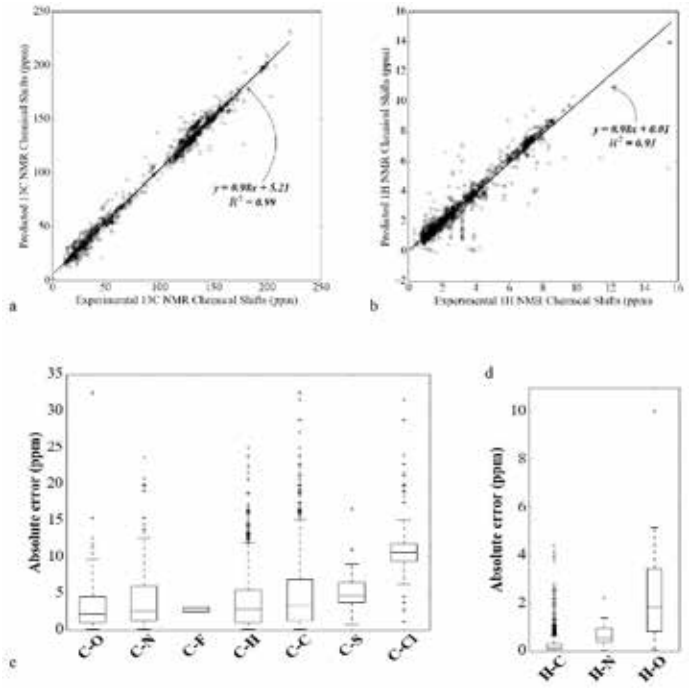
This past year, we developed a supercomputer-driven pipeline to predict NMR chemical shifts and IMS-MS CCS values for any metabolite automatically, rooted in quantum chemical calculations. Solvation effects are accounted for via computationally efficient implicit models for a reasonable description of the solvent behavior. The performance of the software has been tested successfully on a set of metabolites via a benchmark study for quantum chemical theories.

To the best of our knowledge, our metabolite set is the largest studied so far in the literature. The outperforming method gives 4.48 ppm and 0.30 ppm of error compared to experimental results for ^{13}C and ^{1}H NMR chemical shifts, respectively. The level of confidence in the computed NMR chemical shifts is currently being studied to understand the accuracy required for small molecule identification in complex mixtures.

Additionally, computationally efficient explicit solvent models are being developed for the water-solvated compounds common in environmental microbial communities to account for the behavior of water-solvated condensed phases with more accurate predictions and improved understanding of the physical processes. Another program to handle conformers automatically is being developed by our team, which is intended to be a part of our NMR and IMS-MS metabolite prediction pipeline.

During FY 2018, an algorithm will be developed to determine the minimum level of accuracy required to identify the metabolites. Once the algorithm is developed, our automated program will provide calculation of NMR chemical shifts of all the Human Metabolome Database (HMDB) metabolites for a unified method, which will be determined via a benchmark

PN16018/2795



Performance of the quantum chemical calculation pipeline. a & b) Plots of calculated ^{13}C NMR chemical shifts versus experimental ^{13}C and ^{1}H NMR chemical shifts, respectively. These figures illustrate how the calculated chemical shifts are close to experimental data. c & d) Plots of absolute errors found in different functional groups bonded to ^{13}C and ^{1}H NMR chemical shifts, respectively.

study for quantum chemical theories. Conformers will also be considered, in order to decrease the error of our calculations compared to experimental results. Ultimately, our proposed research is expected to identify currently unknown metabolites and, overall, to build a metabolome library specifically for microbiomes and environmental samples.

Impediments

David O. Manz

Cyber security testing and evaluation is currently a manual and expensive process. This effort will develop capabilities to automate and improve the reproducibility of cyber security testing.

This project will characterize and reproduce threats and vulnerabilities on the basis of a network with real-world traffic to increase accuracy and resiliency. This effort will be used in testbeds and other realistic sandbox environments that are safely contained from operations.

Traditional network analysis projects or penetration tests have been limited to non-realistic networks composed of static servers, workstations, firewalls, etc. The testbed used for this project comprised powerful Red/Blue Team tools and network traffic generating workstations to rationalize a real network. Eventually, this project will result in a design framework with resiliency that detects and resolves unknown attack patterns and unpredictable threats.

This effort moves beyond the static networks of point solution testbeds and adds the dynamic and co-evolutionary nature of sophisticated adversaries' engagements. There are no one-offs in the real world, and in a small way, this project introduces that complexity to the laboratory environment.

FY 2016 was spent in the preparation stage by researching state-of-the-art Red/Blue Team practices, real network traffic generating agents, and testbed configuration. Not only using the given tool itself, we also modified and implemented some applications (e.g., Metasploit attack automation tool).

For Red Team simulation, we installed exploitation tools running on the Xenial server and optimized the cloud network. Security Onion, Etherape, TNV, Netgrok, Kali, Core Security, Mutillidae, Virtual Hacking Lab, WebGoat, and Metasploitable are running on the testbed. Furthermore, attack automation application written in Python code will be used to increase complexity and efficiently send malicious packets to the testbed. Intrusion detection systems, such as Snort and Bro, are installed and successfully detect malicious packets.

During FY 2017, we configured all applications and tools that we can use and implement under legitimate license into the testbed and simulated Red/Blue Team scenarios with real traffic.

In the final year of this project, the tools developed, in coordination with the CyberNET testbed, were leveraged to execute an experiment studying the effectiveness of deception as a cyber defense strategy. Utilizing a commercial-off-the-shelf implementation of deception, we defined a general network including deceptive systems. Using this experimental configuration, we measured how long it took penetration testers to find tokens representing valuable data. Our hypothesis was that deceptive systems would represent easier targets and slow down the attack by diverting attention.

The results of the experiment were inconclusive, because the study population did not behave as expected. We learned that our experimental design was flawed in both how we presented the study to the test subjects, as well as including a team of penetrators to capture the full range of skills in a study.

PN16040/2817

MinT-Net: Novel and Scalable Network-Enabled Comparative Tools for Stress Studies of Microbiomes in Transition

Mahantesh Halappanavar

The overarching goal of MinT-Net is to lay the foundational building blocks for a novel, network-driven approach to environmental microbial community analyses and, in the process, provide scalable computational tools that domain scientists can use to predict, control, and, ultimately, engineer microbiomes in a broad range of environment conditions.

There are several key challenges and limitations that underlie computational techniques that are currently used to compare microbial communities and to study how a specific community has evolved under environmental stress conditions. Most tools limit themselves to computing a composition-based comparison of how a community has evolved, using primarily the 16s RNA data. With high throughput technologies, different samples obtained from different environmental states of a community can be sequenced at depth. Consequently, a whole metagenome (or metatranscriptome) inter-sample comparison has the potential to provide novel insights into community dynamics, including tracking intimate interactions that persist within a community (such as horizontal gene transfers, gene addition/deletion events, or metabolic exchanges) as it evolves and adapts dynamically to the environment. However, current tools are limiting both in scope and in scale to cater to this need. MinT-Net addresses this problem.

We accomplished the following outcomes in FY 2017 for MinT-Net: 1) *Scalable Network construction and benchmark datasets*: Developed tools for network construction and creation of benchmark datasets in collaboration with domain scientists. 2) *Algorithms and scalable implementations*: Developed novel mathematical framework, algorithms, and scalable implementations for static and dynamic community detection and influence maximization. 3) *Application use cases*: Developed several use cases in collaboration with domain experts and designed validation methods spanning a wide range of applications involving soil microbiomes, cancer research, bacterial co-cultures, and climate change.

Scalable implementations: We developed a distributed-memory implementation of a popular agglomerative clustering algorithm, the Louvain algorithm. We also optimized the performance of our shared memory implementation of this algorithm and demonstrated its quality and performance on a large set of inputs.

Novel algorithms: We developed a Riemmanian framework for efficient dynamic community detection. Using principles from linear algebra, Riemmanian geometry and spectral graph theory, we developed a tool to interpolate and extrapolate graph snapshots and to accurately predict community merges, splits, births, and deaths.

Application studies: In collaboration with domain scientists, we developed a bipartite model for representing heterogeneous multi-omics data consisting of transcriptomics and polar metabolites from an unpublished bacterial co-culture. We performed bipartite clustering, a specialized variant of clustering, to identify co-clusters of metabolites and transcripts.

An illustration of bipartite clustering using bipartite clustering (Bi-Louvain algorithm).

PN17033/2923

Modeling Continuous Human Information Processing

Leslie M. Blaha

This project seeks to identify models and related metrics for dynamic decision-making to support the development and evaluation of adaptive interfaces for interactive streaming analytics.

Analyst behavior in a streaming visualization environment demands online, continuous decision-making. Modeling human decision-making, in general, requires formal mechanisms to characterize information accumulation and choice biases.

Models of decision-making, particularly from the human information processing perspective, are capable of teasing apart the mechanisms involved in the decision-making process based on non-invasive behavioral data, namely response choice and response time. Existing models capture discrete decision events in static environments.

To capture continuous decision-making, we seek to extend information processing models of decision-making to capture both the discrete decision events and the cognitive dynamics for waiting and monitoring in between explicit decision events. This requires theoretical extension of response time modeling approaches, incorporating exploration of the correlations between sequential decisions.

Additionally, we want to test novel measures of performance based on both total response time and motor dynamic of executing the response. The full dynamic profiles provide insights about the decision-making as the process unfolds. We will develop appropriate metrics to enable use of these dynamics to inform and evaluate the design of analytics interfaces.

The first step of the research was to evaluate models of decision preparation dynamics as candidates for modeling the cognitive behaviors between specific decision-making events. We identified a successful candidate model of decision anticipation, which predicts the likelihood that an event requiring a decision and response will occur at different points in the foreperiod (the time prior to the onset of the stimulus or event cue). The model captures probabilities of cuing events occurring as a function of the time since a response or an alerting event indicating the start of a decision-making trial.

Model simulations show that the model predicts a series of important response time effects, including the foreperiod effect, in which the duration of waiting time affects the speed of responding. Preliminary model fits have been conducted against the average data values from a continuous detection task requiring sustained attention, the psychomotor vigilance task.

We have made additional, parallel strides in the use of Gaussian process models for functional analysis of hand and eye movement behaviors. A new functional version of Fitts' Law is the result of modeling hand movement data, which we have demonstrated on touchscreen interaction data for two-dimensional (2D) movements. We developed a data-driven approach to classifying eye movements according to the movement dynamics. This approach extracts estimates of gaze position, velocity, and acceleration and uses a beta-process vector auto-regressive hidden Markov model to group those dynamics into unique classes of gaze behavior. We can map those to known movement types, like saccades, fixations, and smooth pursuits, in a way that adapts to variability in the data characteristics.

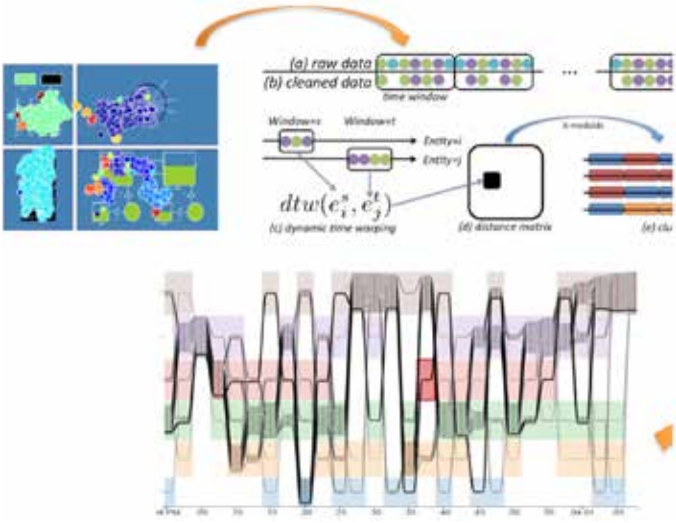
To support the study of dynamic time series of data simulated or captured by these models, we collaborated with the User-Centered Hypothesis Design program to develop a storyline visualization technique for eye movement data. This technique leveraged dynamic time warping for aligning data from multiple human observers viewing movie clips. We are able to find patterns of similar behavior, which likely reflect people who have similar task strategies.

We have made strides in establishing foundational methods for assessing strategies and workload efficiency in continuous multitasking. We have established that the capacity coefficient of response times can be used for a more nuanced analysis of workload efficiency, instead of or in addition to standard dual task assessment measures. As a testbed for this work, we developed the Modifiable Multitasking Environment (ModME) software to support flexible, dynamic multitasking.

PN16096/2873

Over the coming year, we will use this combination of modeling tools and visualizations of strategies to assess human decision-making efficiency in two continuous tasks. We use the ModME to measure multitasking strategies and workload efficiency as people respond to streams of alerts and events in four simultaneously running tasks. We will extend the workload assessment approaches to a multi-unmanned-aerial-vehicle control task to test both how reliance on automation influences decisions and how machine transparency adds to measurable cognitive load.

The results of this work will provide a new set of models, metrics, and visualizations that can be used to characterize user performance in interactive streaming analytics tasks. This will advance the science of usability evaluation, while also enabling the study of the cognitive mechanisms of interactive streaming analytics decision-making.



The data analysis flow to generate a storyline visualization of gaze position data. We start with a set of gaze points recorded with an eye tracker while a person completes a task, like the modified multi-attribute task battery depicted. Data are gathered from a set of multiple viewers. Data errors (e.g., blinks) are removed, and the data time series are chunked into windows. Dynamic time warping is used to measure the similarity between all viewers in each time window. *k*-medoids clustering is used to identify similar behavior patterns. The results are plotted such that each viewer has a line reflecting his/her gaze over time, and similar behaviors are clustered into the differently colored groups.

Module Integration Interface for Resilient Cyber Systems (MiiRCS)

Jeff L. Jensen

This project created an enterprise integration software architecture based on United States Air Force Colonel John Boyd's Observe, Orient, Decide, Act (OODA) loop. Using the OODA loop as a strategic decision cycle allowed for the integration of research algorithms, utilities, and open source software to create a resilient cyber response against illicit activity on the network.

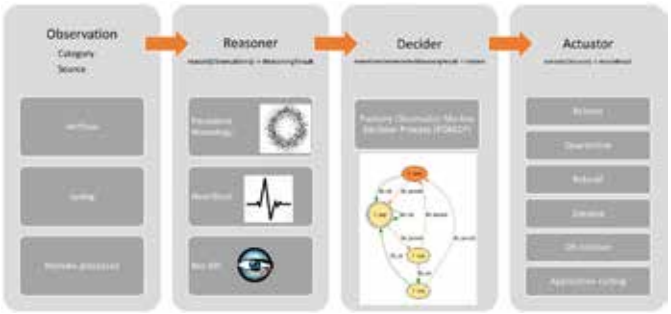
MiiRCS is a specialized software integration architecture that provides secure data transport, sharing, and routing between cyber security applications that act in a larger integrated solution to achieve greater resilience in the enterprise while facing attack.

Using the products of PNNL's research as implementations of OODA components, we have demonstrated resiliency in the face of attack. Applications distributed throughout the enterprise perform various duties, within stages of the decision cycle, and use MiiRCS to communicate in an iterative fashion to enact change in the enterprise.

The MiiRCS team has created an Application Programming Interface (API) based on the OODA decision cycle. By implementing these interfaces, using existing research and software applications and deploying in a Java Enterprise Edition (EE) environment, we have demonstrated a resilient response mechanism. Research algorithms based on Persistent Homology implement the *Reasoner* interface and consume *Observations* in the form of netflow. The Bro Network Security Monitor has a software component that also implements the Reasoner interface. A Partially Observable Markov Decision Process (POMDP) model, derived from PNNL's Rendezvous research, implements the *Decider* interface. Various utilities implement the *Actuator* interface and are executed based on decisions from the POMDP model. Using technologies from the Java EE specification, we have integrated applications in a loosely coupled and distributed fashion.

During FY 2017, the MiiRCS team created the OODA API and implemented it with research algorithms, software applications, and various software utilities to form an integrated and extensible solution. Our software is based on common tools and design patterns that can grow to meet requirements as they are defined. This architecture allows for the integration of research applications to operate in the OODA decision cycle and perform various functions that support resiliency and mission preservation. An actual demonstration of this technique was performed where an attack was executed in a small-scale enterprise environment.

A later demonstration of MiiRCS was conducted where an autonomous response mechanism prevented illicit activities from executing in their entirety. A user (insider) performed a login, installed a network mapping utility, and began scanning machines. MiiRCS was able to recognize this activity as abnormal for the machines involved and initiated a restoration of the machine performing the scanning activity. During this restore action, the user's connection to the machine was reset, requiring another login. After doing so, the user realized the network scanning utility was no longer installed. The user then performed a lateral move to another machine and began searching and viewing content stored there. This activity also was recognized as abnormal and MiiRCS rebooted the target machine, causing the user to lose their connection. The last activity performed was a large copy operation from a network server. Again, this action was flagged as abnormal, and the user lost their connection before the data could be copied. At the conclusion of this demonstration, the user gave up, and the illicit activity stopped.



Software API showing decision interfaces and implementations deployed in MiiRCS.

PN15076/2751

Nuclear Trafficking Objective: Materials Processing Characterization

Richard A. May II

Facility operations require collection and monitoring of many systems to determine if the facility is running properly. This project is attempting to determine if material and other processing activities conducted by humans within a facility can be identified and characterized through only the analysis of routine facility operational data.

There is a lack of research looking at how to use existing facility monitoring information for something other than assessing and monitoring facility performance. Attempting to provide detailed characterization of work being done by humans in the facility is not what the monitoring is designed to assess. In this project we expect to be able to evaluate which facility monitoring signals are of the most value (for instance power metering) and how best to use those signals under varying conditions for characterizing different human-conducted activities. If successful, we will have established a new base line capability to serve as the foundation for continued research.

We identified three monitoring sources already being used to assess facility operational status. These three sources resulted in approximately 30 variables (different sensor measurements) collected at sampling rates between 1 and 15 minutes over a 2-year period.

A challenge working with the data is the limited ground truth about the human activities. There is limited documentation as to the exact time that processing activities occur in the facility and only sparse information on specific sensor responses to a human-run activity. For example, turning on a light draws more power, but does it produce enough heat to turn on the air conditioning? We needed to keep in mind that the monitoring process was not designed for the purpose of detecting or characterizing the human activities we are trying to detect. The sensors were put in place only to help maintain an operational facility. We have begun to address these limitations by instituting new activity documentation processes and working with facility subject matter experts (SMEs) to model activity representations in the collected signals.

Two distinct technical approaches are being taken to model the problem space: statistical model fitting and deep learning model fitting. Both approaches worked this year to develop models for detecting various outliers. For the statistics approach, polynomial models and Probabilistic Principal Component Analysis were used to explore multivariate (all signals together) signals providing a score that can serve as a metric of similarity. The deep learning approach used Nonlinear Principal Components Analysis implemented as deep neural net autoencoders. These approaches were selected to overcome the lack of ground truth information. In the case of the deep neural net approach, the lack of labeled data was a significant factor in the selection of our approach.

By working with the facility SMEs, we are identifying which "signatures" in the signals are relevant and which are not. We have identified four broad classes of possible causes for signatures. First are the human-run activities of interest themselves. Another relevant signature can come from accidents associated with activities of interest (for instance, causing a power surge). While these are unexpected events, they are still providing indications of relevant activities. Then there are two classes of non-relevant signatures (routine signatures around normal building functions but that are not related to activities of interest). Finally, there are facility outages, either planned (maintenance) or unplanned (equipment failure). We are working to understand ways of modeling each class of signature.

Classifying the signatures requires working with the facility SMEs to label detected signatures and to identify salient characteristics, we can use to tune the models. This is the current state of our work. Detecting outliers in single and multivariate signals, working with SMEs to manually classify potential signatures, and starting to develop models for identified signatures.

Work for next year will be to continue to refine our knowledge of what are signatures of interest and how best to model them. Key to this effort will be developing better ground truth. This will be done by capturing relevant information during human-conducted activities, so we can develop a reasonable time period of known events. We will also label datasets to facilitate automated learning for signature classification and to generate prediction models.

The current models are not fully exploiting the temporal nature of the data. This is true for raw signal analysis but also for process flow. We will explore higher order modeling that takes into account the fact that many of the human-conducted activities have inherent order; that is, A is required before B can happen. Leveraging these dependencies is critical to reducing the operational space.

PN17087/2977

Observational Study on Understanding the Security and Efficacy of First Responder Communication

David O. Manz

This project will focus on studying and identifying opportunities to improve first responder performance using Internet of Things (IoT) technology.

IoT is the rapidly expanding network of devices embedded with electronics, software, and sensors that enable them to exchange and analyze data with other device and systems via the Internet. Because of the tremendous opportunities IoT offers for obtaining actionable information more quickly and efficiently, many industrial and government sectors are exploring ways to use IoT to enhance their services and products.

For first responders—including police, fire, and emergency medical services—near-real-time information about emergency situations is crucial for securing the nation's communities and protecting lives and property.

First responders are already using early versions of IoT; for example, sensors embedded in firefighter suits, body-worn video cameras for law enforcement officials, and drones and robots used in hazardous situations. However, many of these devices are limited in what they do, and many do not connect with each other or with existing networks. Among the known IoT challenges are inefficient processes, data speed and volume, data privacy and security, device and platform integration, scalability, and ease of use in the field.

In the future, the new IoT will create additional infrastructure and communication networks. Could first responders leverage this connectivity and use it to communicate and support vital functionality? Additionally, from time to time, there will be constraints to connectivity and energy when faults isolate and disable portions of the communication and power grid. Could an alternative infrastructure be co-opted by first responders to provide consistent situational awareness, a robust means of communication, and an additional method for information gathering?

This research builds a foundation for improving first responder communication effectiveness, security, and reliability by understanding the current role technology plays and exploring improvements in communication and situation awareness, including IoT. To understand the current state and needs of first responders, 250 people in the U.S. first responder community were surveyed. They were asked how they currently use technology in their jobs, what technology inadequacies or gaps they see, and how they might use future IoT technologies. This project went through Institutional Review Board approval.

First, we wanted to understand how familiar first responders were with various technologies; we asked them about four categories of technology that they typically use on the job: 1) their privately owned communication devices (such as their own cell phones), 2) government-issued devices (such as handheld police radios), 3) vehicular communication devices (such as truck radios), and 4) situational awareness technology (such as oxygen sensors).

Then, we wanted to understand the most commonly used features in technology devices; we asked whether they enabled real-time communication, location identification, were multifunctional, and enabled risk detection. We asked questions aimed at understanding where first responders thought technology was headed. We also asked about the potential usefulness of two IoT-related scenarios: 1) getting situational awareness data from smart devices in homes, and 2) piggybacking onto other residential smart devices as an alternative communication system in emergencies.

Our goal was to gain an understanding of current practices and satisfaction levels associated with current communication methods and devices. In addition, we wanted to understand sentiments, expectations, and perceptions around using connected devices to enhance and improve the ability of first responders to perform their duties.

The overriding finding is that first responders are primarily focused on whether their equipment and processes work as needed to do their jobs. Most important to them are connectivity (e.g., location-finding, no dead zones), reliability, interoperability, and affordability. If these are lacking, the technology will not be used.

PN17060/2950

Predicting the Predictions: A Visual Analytic Workflow for Data-Driven Reasoning about Climate Model Predictions

Aritra Dasgupta

A current major challenge in climate science is the establishment of appropriate methods and tools to assess the ability of climate models to faithfully simulate observed past and present climate. Novel data analysis tools have the potential to dramatically speed up the calibration of scientific models and lead to improved understanding of, and new scientific insights into, complex model behaviors and the physical processes they represent. We are developing and evaluating a visual analytic workflow for supporting and distilling expert understanding of the agreement of model predictions with observations (commonly termed model fidelity), and exposing their dependencies on complex combinations of model parameters.

Climate model fidelity is measured by the degree of consistency between models and observations for specific model output variables or features. Scientists frequently summarize aspects of the model's fidelity using statistical metrics such as (but not limited to) the root-mean-square error, correlation, and relative variance of the model output variables compared to observations of the same variable. Because a credible climate simulation requires reasonable simulations of many different physical processes and state variables, it is typically not sufficient to evaluate models on a single metric; instead, models must be compared across a suite of such metrics, leading to a more complex analysis situation.

In addition, identifying models that exhibit similar or dissimilar patterns of performance across a suite of metrics may lead to important insights into model behaviors. However, for such multi-model, multi-variable, and multi-metric analysis, two key problems exist. First, there is a lack of analytical solutions that let climate scientists compare across hundreds of combinations of models, variables, and metrics for developing alternative hypotheses about good or bad model behavior. Second, there is a lack of consensus in the climate science community about the relative importance of different variables and metrics in classifying model performance as good or bad.

To address these problems, we are developing and evaluating visualization techniques and statistical methods that will better enable climate scientists to elicit their expert judgment in the analysis process. We hypothesize that these analytical techniques will lead to a much more reliable and efficient climate modeling workflow, as compared to the state of the art.

To date, we have accomplished the following results: 1) developed novel model fidelity visualization techniques that are more preferred by experts and, in some cases, more accurate than existing techniques; 2) conducted a broad survey for eliciting expert judgment about relative importance of model outputs; and 3) developed an interactive visual analytic tool for scaling up the analysis process to hundreds of models and for incorporating expert knowledge in ranking models based on their performance.

In FY 2018, we will focus on continuing to develop the visual analytic tool for interactive parameter analysis, as well as statistical methods to come up with model rankings that best express the consensus of the climate science community.

PN17013/2903

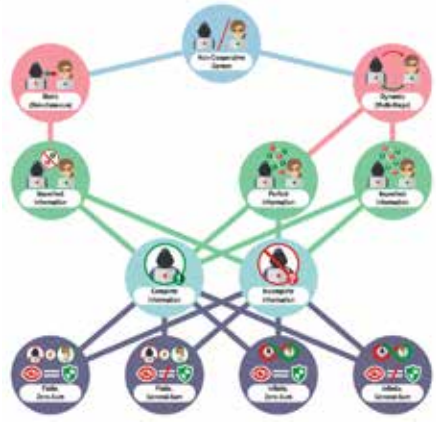
Rendezvous: Optimization and Stochastic Algorithms for Asymmetric Resilient Infrastructure

Sam Chatterjee

By developing a flexible decision-support engine that generates optimal cyber defense action recommendations under uncertainty in this project, we enable the application of game-theoretic and uncertainty quantification methods for improving the resilience of cyber systems.

Cyber security is a multifaceted, challenging research problem. Defenders typically function within resource constraints, while attackers operate at relatively low costs. Design and development of resilient cyber systems that support mission goals under attack must also account for dynamics between attackers and defenders. Game theory provides a mathematical framework to model and analyze dynamic interactions between multiple strategic decision-making agents (e.g., attackers and defenders). Game theory has, therefore, been applied to solve several cyber security problems such as intrusion detection and network interdiction within non-cooperative settings. However, the practical application and scalability of game-theoretic methods is significantly limited by uncertainties in attacker payoffs and accurate identification of system states.

Game theory is the mathematical study of the interaction between two or more entities acting to achieve their individual goals. These methods have been successfully applied to solve several challenging problems related to cyber security. Various taxonomies for classifying game-based modeling approaches exist. These game formulations contain assumptions about rounds of game plays, past player actions, types of players, number of cyber system states, number of player actions in a given system state, and payoff (reward or penalty) functions associated with player actions.



Types of non-cooperative game models for cybersecurity.

Probabilistic attacker payoff framework.

However, a significant limitation of current approaches is the lack of rigorous treatment of uncertainties in payoffs and system states. A thorough investigation of uncertainty quantification methods applied to cyber security games will, therefore, benefit the cyber security research community and help advance the state-of-the-art in this domain.

In a realistic setting, a defender cannot assume that all necessary information—both about the attackers and their own system—will be available. Since a cyber attacker's payoff generation mechanism is largely unknown, appropriate representation and uncertainty propagation is a critical task. One must also account for the lack or absence of perfect cyber system state information; such uncertainties may arise due to inherent randomness or incomplete knowledge of the behavior or events affecting the system. For example, partial observability may make a cyber system's state uncertain over time. Moreover, multiple types of attackers could potentially target a system at any given point in time.

Advances in state-space modeling of cyber systems and reinforcement learning approaches for Markov decision processes have inspired the development of partially observable stochastic games (POSGs) and their potential applications for cybersecurity. Solving such problems involves iteratively finding policies (that map system states to actions) that achieve high rewards, on average, over the long run. However, POSGs are very general formulations and become intractable.

We previously developed a probabilistic framework for quantifying attacker payoff uncertainty within a stochastic game setup that accounts for dependencies among a cyber system's state, attacker type, player actions, and state transitions. This approach adopts conditional probabilistic reasoning to characterize dependencies among these modeling elements. The application of probabilistic theories (such as

PN15031/2706

total probability theorem) and functions (such as marginal and conditional) may then lead to simulation of attacker payoff probability distributions under various system states and operational actions. The framework is flexible and accounts for multiple types of uncertainties—such as *aleatory* (statistical variability) and *epistemic* (insufficient information)—in attacker payoffs within an integrated probabilistic framework. Statistical probability distributions typically address aleatory uncertainty, while mathematical intervals address epistemic uncertainty. Depending on these representations, uncertainty propagation methods may include Monte Carlo sampling analysis, interval analysis, and/or probability bounds analysis. Application of uncertainty propagation techniques generate probability distributions, intervals, or intervals of distributions associated with attacker payoffs that serve as critical inputs within stochastic cybersecurity games. These probabilities may be informed and updated based on empirical event and system data, simulation experiments, and/or informed judgments of subject matter experts.

The game-theoretic and uncertainty quantification methods outlined above model the dynamics between cyber attackers and defenders, and have real-world potential to address proactive resource allocation challenges within resilient cyber systems. However, challenges to their implementation exist, including real-time, data-driven system state determination, "realistic" payoff uncertainty representations, and scalability of uncertainty propagation and stochastic game algorithms.

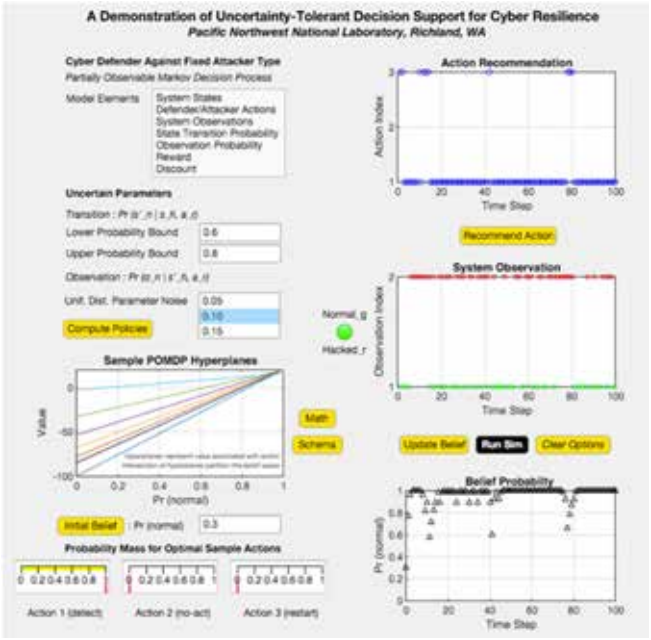
In FY 2016, we focused on four research areas that led to the development of rigorous mathematical formulations of the optimal defender strategy problem and potential solution approaches. The first focus area involved computational experiments (using a notional cyber system) with the characterization and propagation of attacker payoff uncertainties (using two-phase Monte Carlo sampling and probability bounds analysis) as statistical probability distributions, mathematical intervals, and distributional bounds.

The second focus area was the development and application of approximation algorithms for real and synthetic attack graphs to characterize cyber system vulnerability and criticality.

The third focus area involved the development of a stochastic network interdiction problem with probabilistic attack graphs as a two-stage stochastic mixed integer linear program. The goal of this formulation was to generate optimal interdiction plans that minimize the expected maximum risk from a cyberattack. The sample average approximation scheme, along with Benders decomposition technique, was implemented to solve the mixed integer program.

Finally, we developed a formal approach for resilient reachability based on end system route agility in the context of network security. This involved formalization of an efficient and resilient End to End (E2E) reachability problem as a constraint satisfaction problem that identifies potential end-hosts to reach a destination while satisfying resilience and quality of service constraints.

In FY 2017, we focused on three research areas and developed a prototype demonstration of uncertainty-tolerant decision support for cyber resilience. The first focus area was the design and development of an agent-centric decision-making approach to formulate and solve the cyber defense problem as an ensemble of partially observable Markov decision process problems against particular attacker types. This resulted in the generation of optimal defense action recommendations that enable continuation of mission critical operations.



Prototype demonstration of uncertainty-tolerant decision support for cyber resilience.

The second research focus was network vulnerability with dynamic simulation. Finally, we worked on a proactive routing mutation against distributed denial of service attacks.

PN15031/2706

Robust Statistical Data Exploration and Analysis for Microbiome Metabolomics

Lisa M. Bramer

As capabilities to detect and quantify abundances of small molecules grow, the need for integrative and robust statistical methods for exploratory data analysis, data normalization and quality control, and biomarker discovery are needed for impactful and interpretable biological research. This project aims to improve analysis methodologies, develop a robust and reproducible biomarker discovery process, and improve the quality and reduce the time to reach meaningful results integrating multiple heterogeneous data types.

The microbiome response to perturbation is first observed as changes in the metabolome, followed by changes in the metatranscriptome and metaproteome and, ultimately, to the metagenome in the case of a chronic perturbation. However, the analysis of metabolic data from complex microbiome communities poses several challenges: multiple analytic platforms are available, each with unique capabilities and differences in run-time variability and, therefore, varying data analysis requirements.

The ability to perform exploratory data analysis (EDA) is the critical first step in analyzing such data. EDA is of paramount importance for the purposes of assessing data quality, determining relationships among explanatory variables, assessing relationships between explanatory and outcome variables, and making a preliminary selection of appropriate models. Metabolomics data from a microbiome has unique challenges for the EDA process. The biochemical relationships between variables (metabolites) are often complex and, in the presence of other confounding variables, such as censored or missing data, traditional statistical metrics (e.g., correlation) are insufficient and data quality procedures become extremely difficult, sometimes resulting in large amounts of data going unused or being underused.

This project aims to modify existing methodologies and develop new ones for EDA, data quality control, and normalization to identify important biological relationships based on multiple sources of 'omics data. These methods and algorithms developed will be deployed in an interactive framework, allowing researchers to actively investigate data in a manner flexible to multiple data types. Finally, this project aims to be scalable to large amounts of data that are produced in biological experiments. It will do so by using a divide and recombine paradigm that allows for high-performance statistical computing and integration into high-powered visualization tools, which can be combined with statistical metrics.

In FY 2017, we further developed a streamlined framework for data management that is generalizable across data types. This structure was developed in the form of an R package, providing the framework for reproducible research and automated reporting and tracking of data filtering, normalization, etc. We developed functionality for analyzing and visually exploring data generated by PNNL's Fourier transform ion cyclotron resonance (FTICR) instrument, which is utilized across many biological experiments. These functions were implemented in an R package that deploys the data in an interactive environment, allowing the researcher to effectively explore and interact with large amounts of data. We further mapped this data to databases of functional pathways, thus allowing biological researchers to interactively explore statistical results that are interpretable in the biological context. These packages were used in collaboration with several biological experiments and are now deployed on FTICR data generated for any Environmental and Molecular Sciences Laboratory user proposal.

In FY 2018, we will finish the development of two R packages currently in development for the visualization of multiple data types and robust statistical tests and models. Additionally, we will finish software allowing researchers to actively explore the results of a sequencing pipeline implemented on their biological samples. We also will focus on models for integrating heterogeneous data types generated by biological experiments, while still providing statistical results that are interpretable by biologists. Finally, we will further pursue initial developments in detecting interactions observed in heterogeneous datasets, in the presence of missing data.

PN16017/2794

Scalable Hierarchical Validation & Calibration for Robust Distributed Control of Large-Scale Complex Systems under Uncertainty

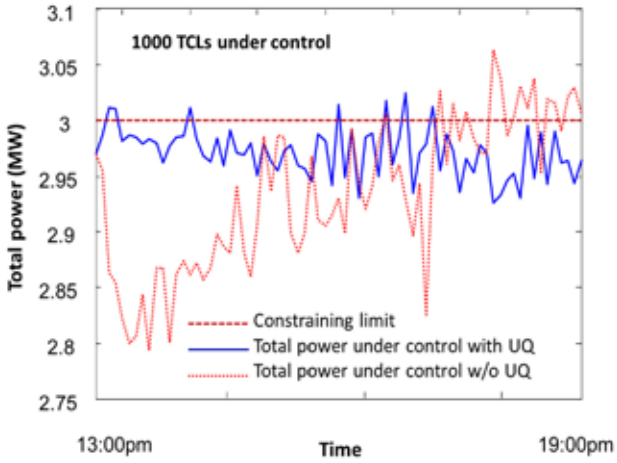
David W. Engel

Optimal, efficient control of the nation's complex systems—such as the power grid, buildings, and transportation—requires a strong predictive capability that enables operators to see into the future and efficiently manage the system to meet user needs. This capability is hindered by uncertainties, or factors not readily evident, that impact prediction and, ultimately, system operations. Being able to quantify uncertainties informs development of more sophisticated models and simulations, leading to better system designs and robust control decisions.

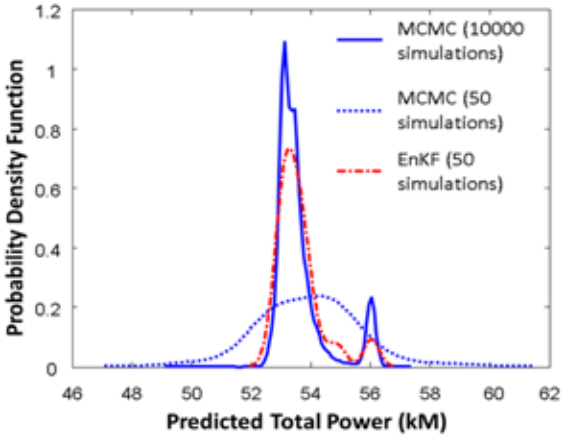
Approaches for validation, calibration, and uncertainty quantification (UQ) that are data-driven may be capable of providing estimates of uncertainty that are time-varying as the quantities being measured vary with time. Such a capability provides the option of adjusting acceptance criteria and, potentially, set points in a time-varying fashion to meet the needs of the next generation of complex systems.

The majority of control systems do not adequately quantify the uncertainties that occur within complex systems, such as the power grid. The identification and quantification of different sources of uncertainty is needed, since it leads to an estimate of confidence about the results of a control strategy. To help with the design of robust control strategies, our goal has been to incorporate UQ techniques to accomplish the following: 1) develop a general toolkit in MATLAB™ for uncertainty quantification, optimization, and model calibration and validation; 2) discover the major sources of uncertainty and quantify how much error/uncertainty we currently have in our control feedback; and 3) determine efficient addition of sensors to reduce uncertainties and achieve more accurate control.

Despite the relatively simple mathematical formulations involved in a UQ analysis, the UQ problems are generally difficult to solve. One cause of this difficulty is that the simulation and measurement models in many real applications are nonlinear and often lack analytical representations. Consequently, the solutions to the UQ problems also need to be sought with some numerical approximations. Another difficulty is seen in high-dimensional UQ problems (i.e., problems



Results showing how incorporating UQ methods better aligns to our control objectives for 1,000 houses with thermostatically controlled loads (TCLs): minimal uncertainty and total power close to the constraint (3 MW). Among the 72 market cycles in the 6-hour simulation period, the aggregated power exceeds the constraint in 7 market cycles, which coincides with our setting of the threshold (0.1).



Because the convergence of the Monte Carlo (MCMC) algorithm with the true posterior probability density function is guaranteed, we use the result obtained from the first method as the benchmark for accuracy comparison (blue line). The ensemble-based method (ensemble Kalman filter [EnKF]) achieves similar estimation accuracy compared with the benchmark while using only a fraction of the computational cost (red dashed line). Meanwhile, we see that the Monte Carlo algorithm with the same reduced computational cost (50 simulations) gives an estimation that is far different from the benchmark (blue dotted line).

PN15002/2677

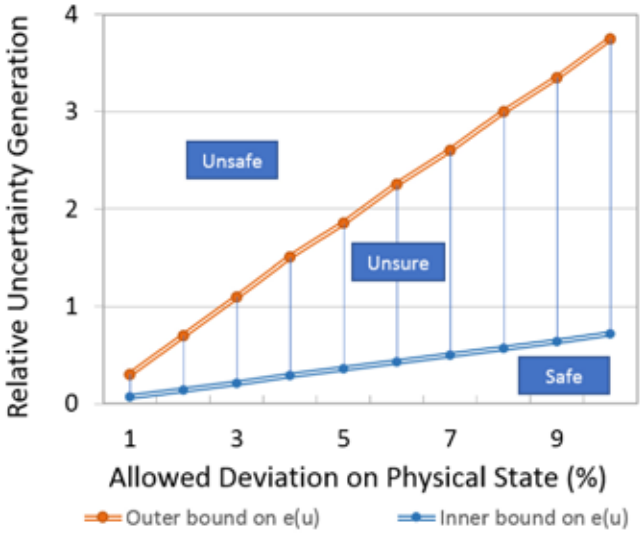
involving a large number of uncertain parameters). This results from the computational cost required by many UQ approaches growing exponentially as the dimensionality of the problem increases and, hence, can become unaffordable.

To accomplish our goals, we have developed a general framework and toolkit that provides a systematic treatment of the uncertainties involved in the control of complex systems. The four major UQ problems in this framework are the identification, the propagation, and the reduction of uncertainties, as well as making control decisions under quantified uncertainties. A specific group of UQ methods (i.e., the ensemble-based methods) were adopted and implemented to solve these problems.

Scalable Verification of Controllers for Complex Infrastructure Networks

Krishnamurthy Dvijotham

Safe and reliable operation of infrastructure networks like the power grid and gas pipeline network are critical to most aspects of the U.S. economy. These networks are modernizing with the large-scale deployment of sensors and controls and active end-user participation. While these developments enhance efficiency and sustainability, they also create challenges by adding uncertainty leading to difficult-to-predict operating conditions. This project develops computationally efficient algorithms to verify that infrastructure networks will work correctly (within specified engineering limits) under uncertainty in demand and supply.



Inner and outer bounds on maximum allowed uncertainty e(u) for the IEEE 30-bus power system benchmark.

Several researchers and engineers have developed solutions (e.g., contingency analysis, etc.) that have even been integrated into commercial software deployed in control centers today. However, these solutions usually involve a heuristic approximation (for example, assuming that the worst case contingency has all power generation at a maximum or minimum value) and, thus, do not provide rigorous guarantees that the network will be safe under all possible situations.

Mathematicians have developed solutions for these problems in great generality (applicable beyond infrastructure networks), but these solutions are usually only theoretical and do not have efficient algorithmic implementations that can scale to networks with thousands of nodes. In this project, we leverage two fields of applied mathematics, computational topology and convex optimization, to obtain computationally efficient algorithms to verify that an infrastructure network will operate safely (voltages/pressures in the network will remain within acceptable limits) given bounds on uncertain inputs (generation or consumption of power/gas, for example).

Further, we are able to compute lower/upper bounds on the maximum permissible uncertainty, so the network can operate safely. This provides system operators a rigorous quantitative estimate of the amount of uncertainty the network can tolerate, enabling them to take appropriate actions (i.e., shed load, curtail excess wind generation, adjust gas compressor settings) to mitigate potential dangerous effects.

We have developed novel algorithms capable of obtaining tight lower and upper bounds on the maximum allowable uncertainty (denoted e[u]), so the infrastructure network can be operated safely. The algorithms have been demonstrated to be practical on several Institute of Electrical and Electronics Engineers (IEEE) benchmark power system models.

Further, we have developed algorithms for robust optimization of infrastructure systems under uncertainty—for example, to optimally choose the set points of conventional thermal generators when there is significant uncertainty in wind or solar generation so as to minimize costs while ensuring that the network operates within limits for all realizations of uncertain inputs within a known uncertainty set (forecasts plus error bars, for example). This is an example of nonconvex robust optimization with recourse, which is considered extremely challenging. Our algorithms are able to deal with this challenge successfully and solve large-scale robust optimization problems in power systems.

Future work will focus on developing scalable, open-source implementations of the algorithms that can be used by researchers and industry practitioners. It will also investigate applications of our algorithms to aid in infrastructure network design and planning (e.g., ensuring that adequate resources are available in the network to deal with the expected growth in uncertain generation and demand).

PN17041/2931

Sparsity-Based Data-Driven Learning Method for Complex Systems

Xiu Yang

Our research aims to develop new, data-driven machine learning (ML) methods for scientific computing. It will provide alternatives to the first principle rule approaches to studying complex systems, and it will lead to an explosion of new discoveries in science and engineering applications.

We are developing new, accurate, and efficient predictive capabilities for complex systems using data-driven approaches. The project's goals include: 1) the integration and development of the emergent and novel state-of-the-art data methods of compressive sensing (CS), ML, sensitivity analysis, and uncertainty quantification, with more traditional methods for dimensionality-reduction and data-assimilation; and 2) the application of this approach to the discovery of fundamental relationships, model equations, and improvement of existing models for complex systems.

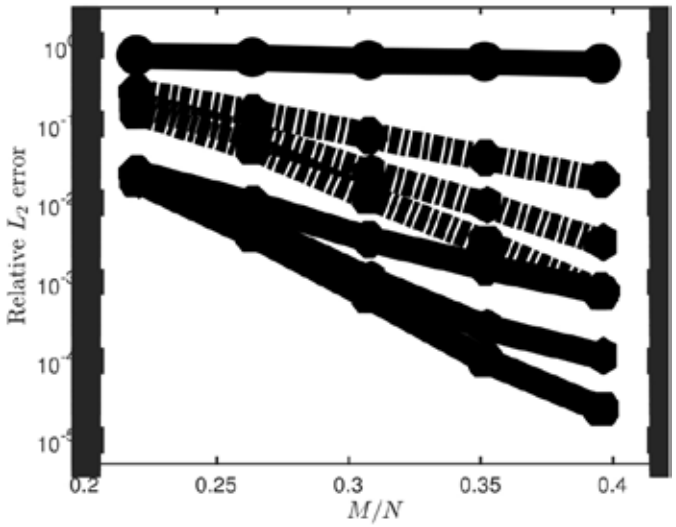
This proposed research is a significant improvement over the existing efforts in this area, because they mainly rely on the proper orthogonal decomposition (or singular value decomposition) analysis of data, which does not capture the dynamical information. Also, the existing techniques using "sparsity" structure of the complex systems do not explore the low-dimensional property of the system sufficiently; hence, they cannot exploit information from limited data efficiently. Moreover, most techniques developed in this area only concentrate on the simple prototype systems and the efficacy for complex systems is not verified.

We developed a new framework of ML that uses sparse measurement. Based on research conducted in the first year, we propose a general framework that uses an alternating direction method to enhance the sparsity of the representation of the quantity of interest in the complex system. The improvement of the sparsity facilitates the CS method to recover the full-state space accurately with limited data. We also incorporate the sliced inverse regression method as the data preprocessing step in the new framework, which further increases the accuracy of detecting the low-dimensional structure in the parameter space. Consequently, it requires even fewer data to construct a regression model of the same accuracy level than the original method we proposed in the first year. This is a breakthrough within the sparse recovery techniques of the past 2 years, as it intrinsically explores the low-dimensional structure of the systems and yields more accurate results than the existing CS method. Moreover, it also helps to identify optimal sensor locations, in order to maximize information acquisition for reconstructing the entire fields in temporal-spatial space.

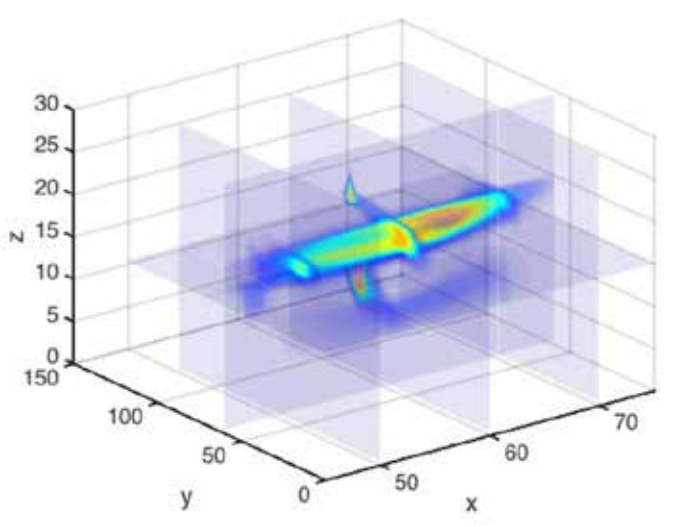
Our collaborator at the University of Washington used the dynamic mode decomposition (DMD) method to investigate the dynamical properties of a stochastic subsurface flow. This system is complicated, in that it has a much smaller ratio of snapshot over the degree of freedom in physical space, which is common in the systems we specified studying in the proposal. The dynamic modes were used to decide the locations for sensors that would allow maximum information acquisition. Different from the existing cases where DMD are used to study deterministic flow, in this study, our collaborator used the technique for a stochastic flow, which is common in hydrology problems, and provided the probabilistic description of the selection of the sensor locations. Namely, the probability of optimality of putting sensors at each location was computed, which provided more robust results for decision-making.

In this project, we proposed in the first year a novel CS method to construct the surrogate model based on very limited data. The general framework proposed in the second year further improved the accuracy (i.e., the general framework requires fewer data to reach the same accuracy level). We also partially integrated this method with the existing DMD method to obtain a sparse representation of the dynamics of the system. As such, it can identify optimal sensor locations to maximize the information acquisition and can yield better reconstruction of the full state of the system. The method facilitates a family of local, reduced-order models for the various physical regimes measured, thus allowing for an accurate reconstruction of the full state of the system, along with a low-dimensional and accurate prediction of its future state, even under noisy measurements. We used subsurface flow as the test case and obtained better results than the existing method, as it provided a more robust description of the selection of the sensor locations and yielded better full-state recovery.

PN16064/2841



Comparison of the accuracy of different CS methods for a benchmark test function; circles are the results by the standard CS, dash lines are results by the method proposed in the first year, solid lines are results by the general framework proposed this year.



Recovery of a subsurface flow from "sensor" data based on DMD method.

SQUINT: Streaming Query User Interface

Svitlana Volkova

This project will develop a stream summarization architecture that facilitates real-time summarization and visual representation of large collections of temporal event sequences. The outcome will summarize the stream within the appropriate historical context to improve decision-making (for example, in operating the electric grid).

Often analysts are handed huge piles of data, such as news articles, images, social media posts, and cyber events data that need to be summarized, organized, or triaged. Machine learning can help with this task, but experienced data scientists know that real world systems require a human-in-the-loop to correct obvious mistakes and improve performance. Current approaches have problems with scalability for the machine (e.g., cannot re-train in milliseconds for large datasets) or for the human (e.g., they overwhelm the user by showing all the data, often in abstract statistical visualizations).

The existing tools are frequently developed for one extreme end of users or the other—either data scientists to debug particular models, or for crowd-sourcing labels for common knowledge tasks (e.g., image annotation), sometimes using active learning. Our contributions are performance and design improvements to CHISSL—a human-machine interface that utilizes interactive supervision to help the user with grouping unlabeled instances into their own mental model—in support of a use case between the two extremes. Our intended user is an analyst (i.e., a domain expert), not necessarily a data scientist, for which active learning would be inappropriate.

In FY 2017, we focused on two research hypotheses to summarize large volumes of unstructured event sequences: event sequences can be represented using text analytics techniques, and the analyst can rapidly build "just in time" models to summarize sequences. We first demonstrated that text analytics is applicable to encode cyber event sequences. For that, we learned event representations using Word2Vec embeddings (i.e., dense vector representations in a low-dimensional space). Embeddings lower the dimensionality of the feature space, reduce sparsity, and can be used as an input layer to neural network models. We then evaluated the effectiveness of event representations for prediction tasks not known *a priori*. For that, we relied on supervised (e.g., Log Linear, Random Forest, and Long Short-Term Memory) and unsupervised (e.g., KMeans, Affinity Propagation, HDBSCAN) models to infer user characteristics and cluster user behavior, respectively. Next, we found appropriate visual representation, level of abstraction, and level of aggregation for cyber event sequences using Computer Emergency Response Team (CERT) Glyph design. Finally, we built CHISSL to allow users to rapidly build and refine models to summarize large quantities of temporal event sequences.

Our approach allows us to build and refine summaries, generate real-time predictions per analysts' intent, learn models to organize and summarize the stream, and refine machine learning models with the human-in-the-loop (using CHISSL). Our system is capable of incorporating user feedback incrementally and immediately without a structured or predefined prediction task. Computation is partitioned between a lightweight web-client and a heavyweight server. The server relies on representation learning and agglomerative clustering to learn a dendrogram (a hierarchical approximation of a representation space). The client uses only this dendrogram to incorporate user feedback into the model via transduction. Distances and predictions for each unlabeled instance are updated incrementally and deterministically, with O(n) space and time complexity. Our algorithm is implemented in a functional prototype, designed to be easy to use by non-experts. The prototype organizes the large amounts of data into recommendations. This allows the user to interact with actual instances by dragging and dropping to provide feedback in an intuitive manner. We applied CHISSL to several domains, including cyber, social media, and geo-temporal analysis.

In FY 2018, we plan to refine and evaluate existing analysis technologies, specifically CHISSLv4 and CERT Glyphs. We will develop compelling application areas for CHISSLv4 to go beyond cyber, social media, and vision domains.

PN17054/2944

Stream Adaptive Foraging for Evidence (SAFE): Human-Computer Co-Assisted Signature Discovery and Evidence Generation for Streaming Data with Deep Learning

Nicole M. Nichols

Some of the most pressing machine learning applications, such as cyber security and multi-sensor data monitoring, lack sufficient ground-truth data to build a supervised classifier. We solve this problem by using unsupervised deep learning techniques to determine when data were anomalous from a baseline norm and provide interpretable event identification to the system's user.

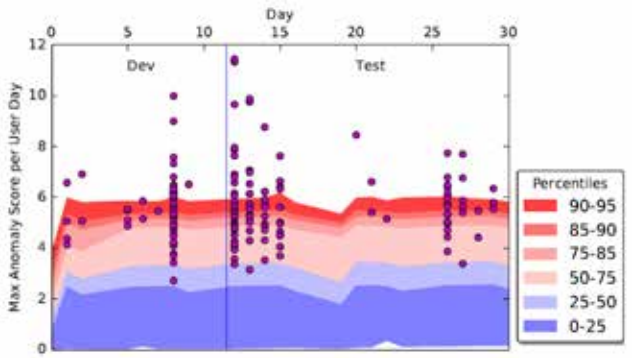
Traditional modeling methods require either expert derived features to construct models, or large volumes of prior example data. Cutting-edge, deep learning approaches are now been used to characterize complex events, but they have generally not been applied to the streaming environment. However, these techniques require a great deal of data, which is typical of large-scale streaming data applications. We demonstrated the use of an autoencoder and relevant featurization techniques to both learn a feature space and then identify anomalous portions without the aid of labeled data or domain knowledge. We have additionally developed capability to work directly from raw data streams of arbitrary format. In this way, we demonstrate the detection of anomalous features without domain knowledge or tagged examples and delivered these features to users.

In FY 2016, the SAFE project developed initial autoencoders and other deep learning techniques for finding the most relevant data from a stream. Models were dynamically trained on numeric or event-based data streams and generated estimates for reconstruction of these same streams. The difference between the input and reconstructed data at every point was compared to determine where rare or unexpected behavior happened in the data set. This method was applied to find real and synthesized heart conditions. Using an aggregate featurization technique, initial Computer Emergency Response Team cyber data, we showed the capability of these methods to both enhance the likelihood of detecting these events in real time and also to deliver likely reasons the identified behavior was anomalous.

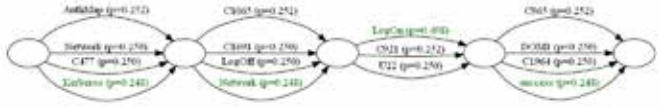
In FY 2017, we applied the extremely positive results of these methods to several additional data sets, including Los Alamos National Laboratory (LANL) cyber security data and VOLTTRON™ building efficiency sensor data. We also extended this capability by introducing a new processing method that works directly from raw system logs. This eliminates the cumbersome step of generating aggregate features that can unnecessarily restrict the expression of input data, as well as extending time to deployment. This new capability uses a two-tiered model, where the first tier learns to encode sequences of characters or tokens, and the second tier provides the temporal context across the history of raw logline data. In a standardized comparison, processing the raw loglines showed better performance than the aggregated feature methodology.

Several approaches to visualization were created to demonstrate presentation of detected anomalies to a user and provide supporting information.

In FY 2018, we will extend some of the ongoing work in interpretability and investigate semi-supervised techniques to refine model performance by integrating user assessment of discovered anomalous events.



Example of anomaly detection using SAFE Technique for LANL cyber security data.



Example of a detected anomaly and anticipated alternatives for analysis of raw data logs.

PN16044/2821

Streaming Data Characterization

Mark T. Greaves

We are creating an integrated resource of high-speed, streaming data characterization and analysis algorithms upon which future projects can build.

Initial processing of streaming data sources is an area in which there is substantial previous work. We planned a set of related projects that will survey the scientific literature on initial stream processing algorithms for co-design domains, implementing the promising projects and tasks, as well as evaluating their performance. We will use these studies to qualify potential university partners for their suitability within longer-term cooperative research and development relationships. There is currently no repository that gathers together the state of the art in powerful online analytic algorithms that can process data on the fly. With this in mind, our repository will be not only critical for success but also a lasting resource for PNNL.

In FY 2014, we directed four exploratory projects within different organizations, as described below.

Foundational streaming data preparation methods. This work was performed by Rensselaer Polytechnic Institute (RPI) and carried many initial research tasks, including a categorization, survey, and component listing for popular open-source and commercial off-the-shelf (COTS) streaming frameworks; an analysis of the specific uses of the selected frameworks that identified the overall strengths and any breakthroughs in these frameworks; and a new formal ontological classification of the chosen frameworks based on the technical approach.

Visualization of hypotheses about streaming data. This work was performed by the University of Utah's Scientific Computing and Imaging (SCI) Institute and addressed statistical hypothesis testing as a visual tool for quantifying and understanding complex phenomena. An approach was developed that relied on simultaneously leveraging the analytic capabilities of both humans and computers using visualization as a knowledge-sharing interface between the two. For example, computers are good at extracting linear models, dealing in high-dimensional spaces, and processing large quantities of data. On the other hand, humans tend to be much better at more abstract tasks, such as identifying nonlinear models and classifying data. In this work, the human helped to build the computer model, while the computer simultaneously helped the user to build the mental model of the data that resulted in a prototype.

Data summarization for streaming scientific video. At the University of Washington, this project supported initial architectural work to address the significantly growing scale of microscopy video data. The scale of imagery coming from new systems, such as PNNL's dynamic transmission electron microscope, requires developing new streaming algorithms even for common tasks such as retrieval, interactive browsing, or compressing video content. This project developed architectures and basic summarization algorithms to identify pertinent scientific content from sequential imagery on the fly and produce a condensed version of video as an output. The techniques also meet single-pass and high throughput requirements.

Event summarization in microscopy imagery. At the SCI Institute, we collaborated with biologists that had streaming image analysis needs. For two selected bioimagery problems, we determined the best set of nonstreaming algorithms and proposed modifications to methods to adapt them for streaming image analysis in experimental settings. Expected impacts include techniques for automated analysis of image data from microscopes to drive experimental techniques, as well as the development of new techniques for quickly post processing acquired image data.

In FY 2015, we expanded and formalized our collaboration with university partner RPI. Together, we focused on studying specific symbolic techniques to consume and reason over a volatile data stream. The project leveraged selected COTS and open-source components of the Semantic Web software stack and explored the interaction of these different software frameworks on choices for streaming data models. Along with RPI interns on site at PNNL, we commenced work on developing a library of algorithms to optimally manage a cache for symbolic data. This new family of algorithms allows us to account for the interaction between the particular symbolic knowledge representation formalism (e.g., description logics of various expressive powers), the different reasoning engines provided by the Semantic Web software stack, and the parameters of the streaming analysis task.

PN14076/2660

In FY 2016, we supported an RPI doctoral student to further develop the cache management algorithms created in FY 2015 and published three papers weaving their operating principles into a new account of *semantic importance*. We provided a formal ontological account of importance ranking and cache replacement based on priority vectors. We used this account to build streaming cache management algorithms for offsides detection and for real-time insider threat detection in cyber networks. In each case, we were able to experimentally show that algorithms based on semantic importance exhibited higher performance under resource limitations than traditional approaches to cache management.

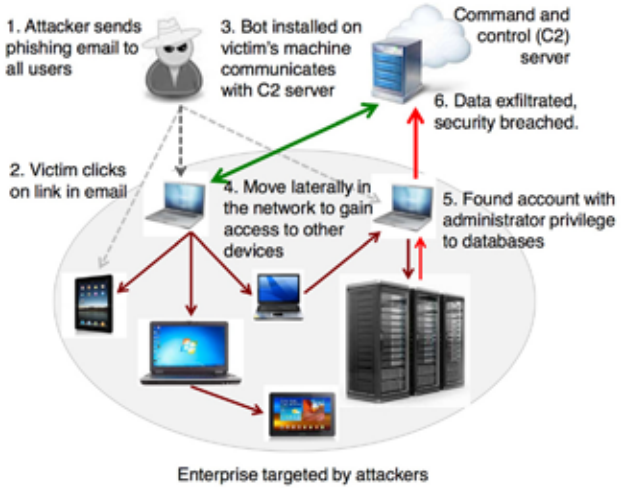
In FY 2017, we completed the Streaming Data Characterization project. We continued to work with our RPI doctoral student on characterizing and refining the algorithms that we developed in FY 2016. We brought this work to completion by creating a stable implementation of a streaming cache management system using semantic importance and used the software we wrote to benchmark our work against 28 competing algorithms. Using our offsides detection data, we were able to show that using semantic importance to manage the stream cache led to dramatic improvements in precision (1.8x precision increase), speed (20x faster), and memory consumption (728x less memory footprint). The implementation software has been placed in PNNL's GitHub repository. In FY 2017, this work also generated three presentations, one publication, one conference presentation, one poster, and one submitted PhD dissertation.

Tell Me Why – Enriching Resilient Action Recommendation with Explanations

Sutanay Choudhury

This project seeks to provide cyber defenders with tools that give context for a threat indicator and potential recommendations on the actions necessary, which is a critical need not being met by current cyber defense systems.

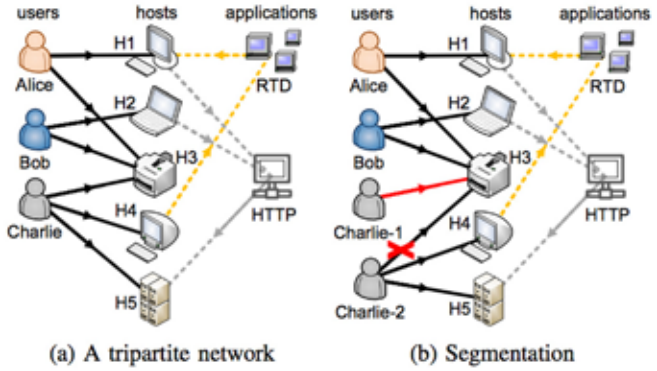
The lifecycle of a cyberattack often begins with the compromise of a single user account in the targeted organization, typically via phishing emails. Once a user clicks on the email, the attacker can obtain access to his/her workstation. From this initial foothold, the attacker begins to explore the network, possibly compromising other user accounts until he gains access to a user account with administrative privileges to the coveted resource: files containing intellectual property, employee or customer databases or credentials to manage the network itself. Skilled attackers frequently camouflage their movements to blend into the normal network traffic, making these attacks particularly difficult to detect and insidious. Since the authorized user plays the role of an unwitting accomplice in these attacks, there is an increasing consensus that designing large enterprises to be resilient against such attacks is the preferred defensive approach.



Lifecycle of a cyberattack.

Most state-of-the-art cyber defense systems typically stop at providing an indicator of anomalous activity without any explanations. "Black-box" style tools that simply produce a yes/no decision or a score have a major usability problem, in that an analyst has to spend an enormous amount of time to sort through false positives and identify the highest consequence attacks to the network's security posture.

We use a graph theoretic approach to model the cyber data sources as a collection of heterogeneous interactions between users, machines, and applications. We propose a set of algorithms that, given a snapshot of a cyber system, provides recommendations to improve its resilience by modifying user access privileges (such as disabling a user account or splitting into multiple roles), selectively disabling applications associated to a user or a workstation, or lowering the vulnerability of a machine through software upgrades.



Example of recommendation algorithms increasing network resilience.

While theoretically robust, the adoption of these algorithms can be limited by the lack of interpretability. As stated earlier, our goal is to provide explanations such as "split user d3x456 into two accounts, because he is the administrator of a highly accessed web server; therefore, a compromise of his account poses large risk." We accomplish this via the following steps: 1) soliciting input from a human expert and building a machine-interpretable version of the mental model, 2) building a knowledge representation of the same underlying graph targeted by recommendation algorithms, and 3) developing a set of deep-learning-based algorithms that find relevant subgraphs in the knowledge representation to produce as explanation.

In summary, we introduce three key techniques: 1) to detect anomalies using a combination of graph-based algorithms and neural network; 2) determine if an anomalous event presents a risk to an enterprise and automatically explain the risk to a user via historical example; 3) model complex interactions between users, machines, and applications, and recommend actions to exploit this structure to minimize risk around impacted resources or users.

PN17062/2952

Temporal Modeling in Streaming Analytics

Lisa M. Bramer

As data recording capabilities continue to grow, research on classification and predictive modeling on data in near-real-time plays an important role in modern data analysis. This project aims to improve current streaming analysis methods, making model output more useful and timely to decision-makers, by leveraging the temporal nature structure of streaming data.

Machine learning algorithms can easily be implemented on streaming data. In traditional machine learning settings, both supervised and unsupervised learning algorithms operate under the assumption that the data is stationary (i.e., the mean and covariance structure are the same for all parts of the dataset), and observations are independent of one another. Analytic methods for streaming data must differ from standard machine learning practices, because in the streaming context, there is a continuous flow of data rather than a static set of data, the data distribution is expected to evolve over time, and observations are seen over time in sequential order and likely are not independent.

This project aims to broaden the understanding of the role that temporal dependence has on the performance of state-of-the-art algorithms for streaming data. This is to be accomplished by leveraging the temporal nature of streaming data sources to improve prediction and classification analytics capabilities by incorporating temporal information into algorithms and models for streaming data. These new methods will be computationally scalable to operate within a streaming environment.

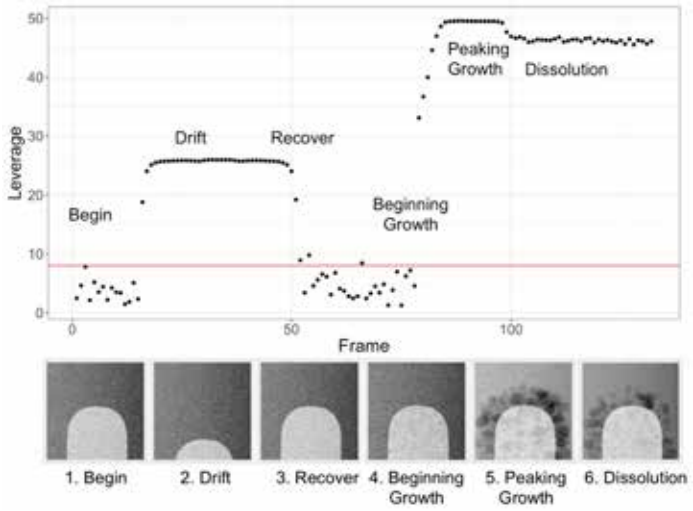
Additionally, we will strive to provide insight into quantifying the benefit of incorporating temporal information into predictive models over a variety of conditions.

Finally, we aim to develop methods for integrating multiple data sources that may be observed at different and/or irregular time scales, which generate reliable predictions at user-specified time steps.

In FY 2017, we conceptualized several models and algorithms that incorporated temporal dependence and applied these algorithms to three unique domains—mass spectrometry, electron microscopy, and social media—with data collected in a streaming context.

Non-parametric testing methods were refined and applied to mass spectrometry data for the purpose of quality control (QC) and were shown to significantly improve the detection of low-quality sample runs in real-time. The final product for mass spectrometry QC has been implemented and connected to the mass spectrometry instrument and is being utilized in real-time, with instrument operators acting as humans-in-the-loop.

For the electron microscopy (EM) use case, we refined and extended the methods developed in previous fiscal years for sparsely sampled images from a microscope. The streaming event detection algorithm we developed in FY 2016 for gray-scale images (REDSI) can now be applied to high-dimensional data such as RGB and hyperspectral images. We showed that the specificity and sensitivity of REDSI remains reasonable as the dimension of the application space grows. We also showed that REDSI can help with real-time *in situ* data collection. This was accomplished by deploying the REDSI algorithm in a graphical user interface (GUI) installed on the microscope's computer, which typically only collects the data. This GUI was used during several live, *in situ* experiments conducted in FY 2017. The microscopists that used the GUI confirmed that it was accurate and useful.



Event detection identification and corresponding anomaly scores for EM video. Results are based on data for 5% of the image shown being available to the model.

The method previously developed to quantify features in EM experiments was refined and implemented in the GUI described above. In order to make this possible, the method was made faster and more automated in order to deal with arbitrary experiment conditions and to give near-real-time

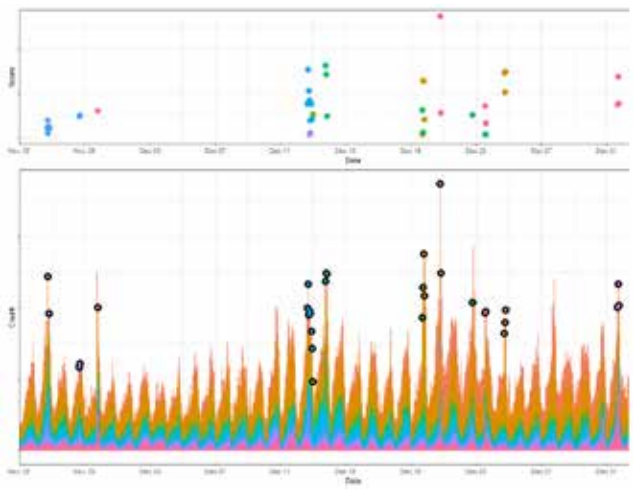
PN16039/2816

results. We successfully applied the event quantification algorithm to several new types of *in situ* experiments, including sophisticated nucleation experiments where events are differentiated based on both grayscale value and event shapes, which vary based on the nucleation site relative to the field of view.

In the social media use case, we worked on Twitter data to detect anomalies in tweet counts containing relevant text strings. Signatures existing in multivariate time series may not be visible through individual univariate analysis. In order to detect unusual anomalies in the data, careful consideration of temporal and multivariate dependence is necessary.

We approached the problem using multiple modeling techniques; both non-parametric and parametric approaches were developed. The value in using multiple techniques is that we gain confidence in our unsupervised algorithms and allow for multiple flavors of anomaly to be detected.

In FY 2018, we will develop an ensemble approach to combine our current methodology and deploy our tool to the appropriate sponsors.



Lower: Normalized tweet volumes for multiple streams (denoted by color), with events detected by multivariate algorithm denoted by points. Upper: Streams contributing to anomalous event score with magnitude of anomaly shown.

Test Bed Federation Tools for Control of Complex Systems Research

David O. Manz

We are enabling high-fidelity experiments on complex systems by designing and implementing tools for federating geographical disperse test beds to support the sharing of cyber and physical resources for collaborative experiments on control of complex systems, such as the electric power grid.

The highest form of research evidence is gained through experimentation with real systems. However, when attempting to research systems like buildings and their interaction with the power grid, it is impossible for one research organization to fund and support enough equipment to model at realistic scale and complexity. The most reasonable solution is to collaborate among the research community to achieve the desired scale and complexity to fully explore new control architectures, paradigms, and algorithms for operational use. This project focuses on developing the tools necessary to enable federation among individual test beds that are geographically dispersed.

A goal of this work is to facilitate formation of a coalition of facilities owned and operated by multiple participating organizations. The tools needed for enabling this collaboration include secure tunnels across the internet, access controls, experiment model sharing, and information and control transmission. Specific needs for tools will be defined in a parallel effort of all participating organizations.

This project supports identification of the needs for tools, design of a system necessary to fulfill requirements, and implementation and testing of the tools. To foster collaboration and speed progress, the tools developed will be packaged so they are easy to distribute to partners for operation within their test bed environments. This packaging can include the creation of an operating system image, a virtual machine (VM), or even a prebuilt and preconfigured piece of hardware. What we have identified is three options: the first is a VM, the second is a bare-metal installation, and the third is a portable, deploying hardware solution to bring the resources to the collaborators. As every organization has their own policies and acceptable risk levels, the tools are being developed with as much flexibility as possible. This flexibility can extend through every layer of federation, including multiple types of tunnels with varied security protocols, multiple field protocols, and multiple methods of integrating equipment.

As the goal of the test bed is to study new control methods, the federation tools also must enable the distribution, management, and operation of control logic. In support of this goal, the federation tools team is investigating the use of the VOLTTRON™ platform developed at PNNL as one of the tools to enable the distribution of control agents into collaborator test bed environments.

In FY 2016, the initial design of the federation tools was created. This design includes the packaging of the tools within a VM that can be shared with collaborators. The VOLTTRON platform provides a translation interface that allows communication with devices on collaborator sites, with operational protocols like Modbus to a routable protocol like the VOLTTRON Interconnect Protocol. Initial implementations of this design were created to start discussions with collaboration partners.

In FY 2017, the design was improved based on collaborator feedback. The VM approach was augmented with a bare-metal installation for collaborators who are unable or unwilling to run virtual machine images that PNNL produces. In addition to this approach, initial designs have been made on a hardware-in-a-box solution that can bring limited collaboration resources to the partner sites, if needed. Additionally, trial federation has been conducted with one collaborator, and multiple collaboration sessions walked them through the federation connection process. The pilot ended in a full end-to-end connection being established.

In the final year, this initial pilot will evolve in a more deliberate and integrated collaboration among multiple sites. Pairwise and multiple-party federations will be explored, and collaboration among parties will enable live distributed experimentation on and study of control methods applied to complex systems. Two or three more collaborators are expected to join the pilot for FY 2018 efforts. Refinements based on feedback will be made, and the resulting tools will be integrated into a cohesive package.

PN16004/2781

Topological Analysis of Graphs in Cyber Security (TAGs-CS)

Emilie A.H. Purvine

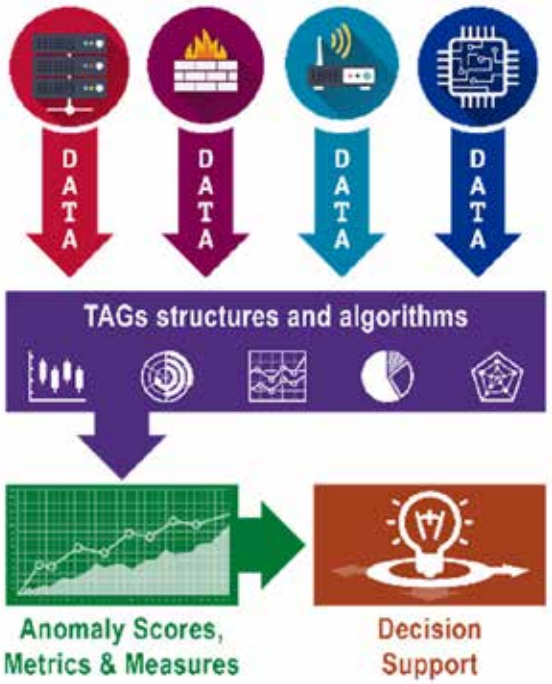
We are discovering shapes, structures, and overall cyber system behavior using topological data analysis of spaces derived from data, providing analysts with information to determine the state of their cyber system.

Topological methods have been deployed in real-world problem domains with great success (e.g., signal processing, medical imaging, and drug discovery). These methods were developed to identify the largest, most robust shapes and structures in data sets and to discover average system behaviors. In cyber systems, the network is dynamic, as elements (e.g., computers, printers, and mobile devices) sign on and off, as communications begin and end, and when interactions with other networks initialize and complete. In this project, we analyze these dynamic systems using topological methods and create new tools for system administrators to judge resiliency and ensure that missions can still be completed.

Our initial hypothesis was that this type of analysis will provide a way to differentiate between resilient and non-resilient system states, possibly even identifying different causes of non-resilient states. Within the broad theory of topological data analysis, we are using two major techniques, both coupled with novel statistical graph analysis.

The first technique, persistent homology (PH), is used to discover the most prevalent higher order structures within a topological space. We use a time interval of NetFlow data to generate a point in a high-dimensional space by calculating relevant features such as number of IP addresses seen, number of flows, and total number of bytes transmitted. Over time, these points accumulate into a point cloud that we consider as a topological space. The PH of a space can serve as a fingerprint of the structures contained within. We compare the fingerprint of a query space—corresponding to the current time period—to that of a known baseline space using established distance metrics on these topological fingerprints.

In FY 2016, we showed by example that this methodology can be used to identify instabilities or unusual events in simulated NetFlow data sets (e.g., a port scan or a data exfiltration). Then, during FY 2017, we refined our methodology for calculating relevant features, and tested on real, full-scale PNNL firewall log data. Without the luxury of ground truth in this real data set, we could not differentiate between resilient and non-resilient states. However, we were able to identify outlier behavior within the data and validated with system analysts that this behavior is interesting and would be worth investigating further.



The TAGs tools are able to ingest a variety of data sources to understand the cyber system as a whole rather than through separate analysis of individual sources. Anomaly scores and other metrics are calculated, which can be passed on to decision support engines.

Our second technique tracks the behavior of each individual IP in a network according to its level of data source or data sink. We consider cyber traffic as a dynamic "flow" on a graph where information (bytes) moves between nodes over time. This flow is dynamic, since communications only occur for short time periods, and some network components can be turned on and off regularly. By applying discrete Hodge theory, an area of differential topology, to the dynamic flow, we derive an ordered list of all IP addresses (i.e., a ranking), starting with those that are most like data sources (where information mostly flows from) down to those that are most like data sinks (where information mostly flows to). The ranking is computed for each time period, so that when a particular IP

PN14086/2670

address has a major behavioral change, we flag it for further investigation. When our PH metrics indicated some level of instability in the system, we were then able to discover the perpetrators of port scans and data exfiltration in test data sets during FY 2016.

During FY 2017 we scaled this method up to handle the larger firewall data by breaking the graph into smaller connected components. Additionally, during this year, we extended the method to incorporate IP-type information. Intuitively two IPs with the same function or type should behave similarly, whereas those with different types should not have the same expectation. Incorporating this information into the method allows us to consider anomalies as either instances when 1) an IP changes behavior drastically, or 2) an IP's behavior deviates from that of other IPs of the same type.

Finally, in FY 2017, we continued to explore the use of information measures on neighbor distributions in cyber graphs. An IP address with all communications going to the same destination IP has a different behavioral profile than one who spreads out communication across multiple destination IPs. We used this observation to create behavior profiles for the network, which evolve as IPs change their communication patterns.

We had previously observed that deviations in this behavior profile indicated the occurrence of an anomalous event in synthetic data sets. This year, we also scaled up to enterprise size authentication data available from Los Alamos National Laboratory, which contains ground truth red team events. In some cases, we observed that our information measures for red team machines were abnormal during the attack time period, whereas during other time periods, there was not an obvious deviation. However, based on our prior experience with these information measures, we still believe that they have value in identifying some classes of abnormal behavior.

Overall, in FY 2017, our final year of the project, the theme was scaling up of all of our methods, moving away from synthetic test cases to analyzing real data.

Toward Enabling Complex Sensemaking from Streaming Data

Lyndsey Franklin

The goal of this work is to increase the cognitive complexity of analytical tasks that can be performed by analysts in environments that use streaming data.

A typical streaming data environment is a human-in-the-loop system that supports basic situation awareness of pre-defined indicators or generates alerts to which humans respond. There is little, if any, support for sensemaking tasks. Sensemaking involves the simultaneous development of a frame to think about data in and the fitting of data to that frame. A great deal of cognitive effort is required in order to explore available data and come to an understanding of it in the sensemaking process. This level of effort is not possible in human-in-the-loop streaming data environments, where a human is tasked with constantly attending to system-processed data, and so such sensemaking tasks must be taken offline and away from the stream, which adversely impacts the utility of their analytical outcomes.

The cognitive complexity of the analytic tasks that humans can perform on streaming data can be increased by rebalancing the effort between humans and machines. This rebalancing entails better partitioning of tasks between humans and machines, reducing the amount of effort required by humans, and understanding how the machine's actions support or disrupt the human's analytic process. The goal of our research is to develop theoretical measures of human task complexity to address task partitioning, develop technology that supports implicit interaction with machines and computational models to reduce required human effort, and study the effects of machine-/system-generated recommendations on human reasoning during analytical tasks.

In order to develop theoretical measures of human task complexity, we applied the Human Oracle Model and extended it, so that it could describe the impact of a resource-constrained human. This meant developing ways to account for limited human attention, limited working memory, decision times, etc. This approach allows us to quantify the operations performed by humans and machines, where streaming data enters and impacts a workflow, and the level of support provided by the machine. We applied this resource-constrained model to typical streaming tasks, such as monitoring with and without system-generated alerts, and tasks that required decisions to be made in response to system-generated alerts. We evaluated the resulting theoretical predictions about time and effort with a Mechanical-Turk study involving human participants performing these tasks online.

To reduce the amount of effort required for a human to interact with supporting computational models, we have developed an analytic-interaction library. This JavaScript library is compatible with web-based visualizations (such as those developed using D3) and can be used to collect information about a person's tasks and interests passively, while the person completes their work. Logged information includes frequently used data items and attributes, which can be used as refining input to steer underlying computational models.

Additionally, we have developed metrics for assessing the potential for bias in a person's analysis, which can be used to

This figure shows two screen captures from the rate-of-stream study. Each row represents a "user" of the simulated network. Color encodes either each user's role (top) or department (bottom). The individual dots represent log on/log off times (top) or internet activity (bottom). All participants in all conditions had access to these color and data options. The various conditions differed on how quickly information updated and whether or not the participant had direct control.

PN16037/2814

redirect their attention to other parts of the larger data stream and improve the quality of their analysis. By collecting this information passively, we can avoid interrupting the analytical process but still keep the human-in-the-loop system closed with feedback to the machine.

In order to develop an understanding of how system-generated recommendations or alerts affect a human's reasoning process, we have conducted a user study that baselines the performance of participants completing a ranking task under varying rates of streaming data. In this study, we have asked participants to rank the users of a synthetic computer network based on whether or not their network use appears to be malicious.

Our preliminary results indicate that participants experiencing the fastest rates of streaming data do not have the time to adequately explore all of the data they are presented with; therefore, their performance on the ranking task is limited to prioritizing two or three of the total number of network users. This finding will be an important baseline in our upcoming work, when we begin to augment the study environment with increased machine support. It allows us to determine the level of support required from the system to reduce the impact of the stream speed on this prioritization task.

Towards an Understanding of the Role of Hydration and Hydrodynamic Forces in Modeling Synthesis

Huan Lei

We are developing a computational framework for accurately modeling nanoscale multiphase fluid in the presence of non-local interfacial energy and thermal fluctuation. This can lead to a deep understanding of the fundamental mechanisms that control the formation of hierarchical structures composed of simple shapes of rods, bricks, and sheets interacting in aqueous solution.

During the past few decades, theoretical modeling of the self-assembly of hierarchical structures has been dominated by focusing solely on the role of repulsive forces under the conditions of packing that give rise to the formation of complex structure. In this work, we focus on additional solution mediated forces for a more traditional approach to nanoscale synthesis. These forces have their origin in the theory of hydrophobicity. Moreover, the dynamical response of a fluid that gives rise to a non-trivial hydrodynamic interaction may be an important ingredient in understanding the kinetics of synthesis at the nanoscale.

Currently, there is no well-established framework for this problem. Our aim is to establish quantitative understanding of the interplay and propagation of the molecular-based, fluid-mediated interactions and their effect on assembly of hierarchical structures at the macroscopic scale. In particular, we will focus on the role of hydrodynamic and hydration forces on the process of assembly. These forces will likely depend on the molecular details that modify the solid-liquid boundary and the ultimate pathway to synthesis.

Accurate modeling of nanoscale fluid-fluid interfacial fluctuation: We have developed a numerical method to accurately impose nanoscale surface tension effect and the important granularity of water density fluctuations to quantitatively describe the hydrophobic effect across length scales. In particular, we have validated the developed computational framework by modeling the nanoscale fluid-fluid interfacial fluctuation. The obtained capillary wave spectra show consistent results with respect to different model resolutions and agree well with the simulation results obtained from molecular dynamics (MD). The computational efficiency is > 100 times faster than the MD simulation.

Since the material synthesis process is closely related to the probability of the void formation adjacent to the target boundary, we further studied the rare event of the void formation process (e.g., the non-Gaussian tail). Our method can accurately predict the size-dependent probability of the void formation process, which yields consistent results with the direct MD simulation.

Theoretical framework for modeling non-local fluid-fluid interface: At nanoscale, interfacial energy becomes scale-dependent. A computational model based on the unified Young-Laplace equation fails to capture the non-local interfacial energy across fluctuating interface. To reconcile the scale-dependent interfacial energy and model-resolution-dependent surface energy in the computational model, we further developed a rescaled computational modeling for modeling a resolution-dependent multiphase fluid system. The novel idea relies on mapping the Lagrangian particle model into Euler lattice model, while the non-local interfacial energy and density distribution are converted into the free energy functional minimization of the Van der Waals theory. The developed computational framework enables us to accurately impose interfacial energy across fluctuating interface with adaptive modeling resolution.

Free energy (potential of mean force) between nanoparticles: Since the synthesis process is essentially controlled by the free energy between the nanoparticles, accurately quantifying free energy is essential to understand the mechanism of the nanoscale synthesis process. We have computed the free energy between the two nanoparticles using our developed method. The obtained free energy function exhibits a small energy barrier and agrees well with the MD simulation results. Furthermore, we computed the probability of the void formation near-adjacent-to-boundary of various hydrophobicity properties. For all the boundary conditions considered, the present computational framework yields consistent results with the MD simulation results.

PN17034/2924

Shape-dependent assembly process: With the developed computational framework and validation presented above, we were ready to study the nanoscale material synthesis process. In particular, we considered the assembly process of nanoscale particles of two different shapes. First, we considered the assembly process of two spherical particles. Due to the symmetry of the shape, the two particles assemble in isotropic way. Second, we considered the assembly of two cubic particles. Different from the spherical particles, the synthesis process exhibits strong orientation assembly (OA) manner. This result is consistent with the observations reported by experimentalists. While such a process is beyond the computational limit of MD simulation, our developed computational framework provides an appropriate approach to understand and quantify the mechanism of the synthesis process arising from different particle shapes.

In FY 2018, we will conduct a systematic study to understand the mechanisms of the OA process for different particle shapes. In particular, we suspect that the OA originates from anisotropic free energy landscape. We will further formulate and understand the role of the hydrodynamic force experienced between moving particle in solution, as well as the kinetics of assembly through comparison study with experimental measurement.

Transpire: Transparent Model-Driven Discovery of Streaming Patterns

Aritra Dasgupta

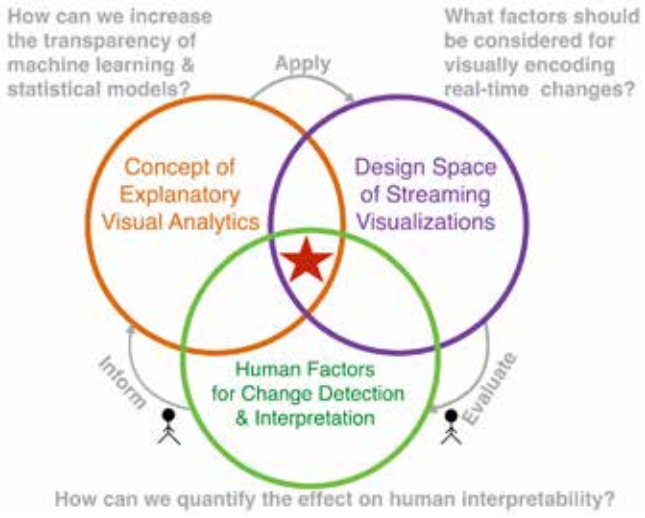
In streaming data environments, the only thing that is constant is change. To the human observer or analyst, changes are often too fast to notice, too many to remember, and too complex to understand or make predictions. To help analysts discover and reason about changes, we will be developing a transparency-based, explanatory visual analytics framework for integrating streaming data visualization techniques with automated methods that increase confidence and utility of human judgment for streaming data analysis.

Most existing visual analytics techniques are based on assumptions about how human-stream feedback should work and are not supported by concrete empirical evidence. In our work, we fill these gaps through three main contributions.

First, we tested how varying levels of model transparency can best leverage the high bandwidth of human perception systems for communicating key insights. By conducting qualitative studies through surveys and interviews of analysts from diverse domains, we will develop a characterization of model transparency and its impact on analysts' perception of streaming patterns.

Second, we used this transparency characterization for developing visual analytics methods and techniques for context-aware human reasoning about streaming data. We study how differences in model transparency affect efficiency, effectiveness and the trustworthiness of human insights.

Finally, we aim to demonstrate the efficacy of our methods through case studies and quantitative user studies that demonstrate how these visual analytics methods can significantly improve human interpretability of machine-detected streaming patterns.



Illustrating the three-fold objectives of Transpire.

In previous work, we established a streaming data visualization design framework for analyzing trade-offs related to change perception. We also conducted a user study demonstrating the explicit relationship between transparency of analytical methods and human trust. We developed a visual analytic tool for instance-level explanations of binary classifiers. Finally, we developed a visualization technique for comparing baseline patterns in a stream of many data objects (e.g., users in a network, items in a transaction database, etc.)

During FY 2018, we will continue to investigate the issue of domain experts' trust in automated methods and visual analytics by conducting design studies and controlled experiments with domain experts who are not trained in machine learning.

PN16038/2815

Visual Analytics Platform for Large-Scale Hierarchical Control System Data

George Chin, Jr.

To enable control system designers to develop and evaluate modern-day control systems more effectively, advanced visual analytics capabilities may be developed and deployed to allow designers to better explore, interpret, and analyze the behavior of complex systems and the data and results they generate.

Infrastructure systems today are generating ever-increasing amounts of data that are highly complex, relate to large numbers of sensing and/or control end points, and involve multiple scales of operation. Innovative and advanced visualization tools and platforms are needed to enable control system designers to visually test, validate, and refine control systems; effectively explore big data to identify critical events, relationships, and patterns; access, evaluate, and compare experimental results at multiple levels of detail and aggregation; understand how controls manage, impact, and interact with devices, agents, and systems; provide control system-specific and general modeling visual analytics capabilities currently unavailable to control system designers.

This project focused on researching and developing novel visualization techniques and tools for navigating and exploring large-scale, massive-entities of control system data. Tools feature intuitive visual representations, zoomable interfaces, multi-scale layout and navigation, and scalable performance. In addition, the project developed a visual analytics platform (to manage and integrate visualization tools with complex models and data) that combines and examines results across multiple experiments. This platform provides connections to data sources, coordinated views of data and models, and capabilities to search, filter, and integrate data from control applied to complex systems.

In FY 2017, the project abstracted the geospatial modeling capabilities of the visual analytics platform to enable the embedding of different types of topologies, from buildings and campuses to electric distribution and transmission systems. Individual topology elements are now mapped to encoded visual primitives, which may visualize associated data in a variety of ways. The visual analytics platform collects and manages a data hierarchy that enables the exploration of experimental results at different levels of aggregation or resolution. Thus, a system designer may visually zoom from a building to a campus to a city to a distribution system to a transmission system and view experimental data relevant to the level of the current view.

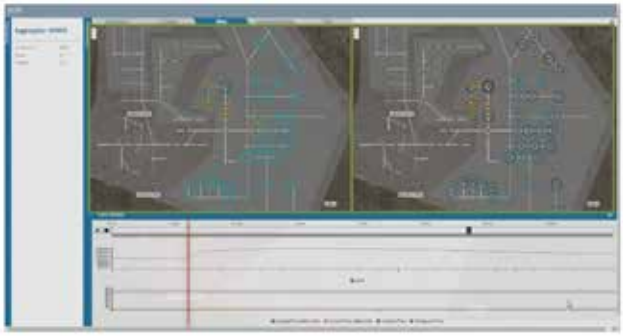
Beyond the geospatial view, topological elements were also mapped to alternative logical structures such as sets and hierarchies corresponding to those inherent in the alternative packed circles and sunburst visualizations. Thus, a new topology may be easily integrated into the visual analytics.

New animation capabilities have also been added to the visual analytics platform to allow the system designers to view the temporal behavior of topological elements, systems, and devices. Additional capabilities have also been developed to visually alert the system designer when critical data thresholds in the experimental data have been breached. The combination of the topological, animation, and alerting features provides essential capabilities for conducting geospatial and temporal analyses of control system data.

Animation and alerting features convey emerging temporal and spatial patterns and conditions in time-series data.

To support comparative data analysis, side-by-side visualizations may be displayed in a shared window. With this new feature, a control system designer may view and contrast experimental data across two different topologies, local sections of the same topology, or the same topology at different points of time. This critical feature enables system designers to better understand and assess how changes to a control system affect its underlying behavior.

PN15003/2678



New side-by-side comparison view allows control system designers to compare results from different experiments animated across time.

Additional optimizations to the visual analytics platform were implemented to enable more efficient rendering. Loading of data into memory is now limited only to data in the current resolution and view. Further optimizations are under development to only load new data that has changed from the previous view.

Nuclear Science and Engineering

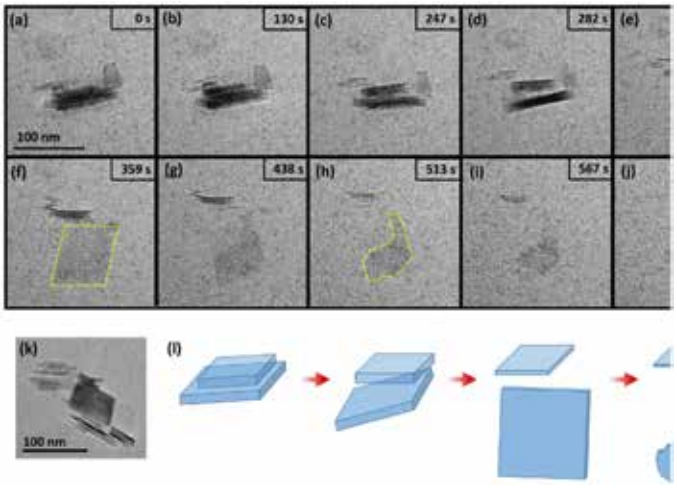
An *In Situ* Investigation of Boehmite (gamma-AlOOH) Dissolution under High pH (Potential of Hydrogen) Conditions

Edgar C. Buck

We are using a combination of advanced atomic resolution and* in situ *liquid cell transmission electron microscopy techniques to determine why Hanford-derived boehmite is more resistant to dissolution than other forms of boehmite.

Boehmite (γ-AlOOH), together with gibbsite (γ-$Al(OH)_3$), are the most common aluminum (oxy)hydroxide (AOH) mineral phases in nature and are both classified as layered minerals. Over the past year, we have been examining the potential intrinsic effects that may affect boehmite reactivity. To investigate dissolution, we have utilized liquid cell electron microscopy (LCEM) because of its high temporal and spatial resolution. Oddly enough, we found that when an aqueous suspension of nano-sized boehmite was exposed to electron irradiation, the particles were observed to dissolve, even at circumneutral pH and room temperature.

This dissolution step followed a pathway that first included rotational and translational movement of the particles, their dissociation from each other, and internal delamination before and dissolution. However, gibbsite nanoparticles exposed to identical conditions did not dissolve. The fact that dissolution and delamination are closely coupled for boehmite, but that gibbsite neither delaminates nor dissolves, strongly suggested that destabilization of the hydrogen bond network holding the structural units together is an integral part of the dissolution process, as the hydrogen bond network for boehmite and gibbsite are very different.



LCEM imaging of boehmite dissolution. (a)-(j) Sequential transmission electron microscopy (TEM) images showing boehmite particles in deionized water under uniform TEM illumination (200 keV) resulting in the particles initially separating/delaminating and rotating, followed by dissolution; (k) stacking behavior of the boehmite was also seen in cryo-TEM imaging of boehmite in water; (l) schematic representation of the delamination, rotating, and dissolution steps.

In earlier caustic dissolution studies, we examined the effect of transition metal dopants on boehmite—in particular, the effects of Cr, Mn, and Fe additions. It was found that Fe-doped boehmite did not dissolve. In our later LCEM work, Fe-doped boehmite also did not dissolve, but likely for different reasons: that Fe can serve as an electron trap and electron/hole recombination center, points to another mechanism that could explain differences in mineral stability under the beam.

A detailed, theoretical effort is warranted to help untangle the possibilities, some of which have not yet been realized, but the present study has provided context and constraints that can serve as guideposts for future work. More generally, this study provides a better understanding of and appreciation for the importance of interlayer hydrogen bonding in layered mineral stability, as well as specific, baseline information concerning the stability of Al (oxy)hydroxides in water during exposure to the electron beam.

PN15090/2765

Characterization of Radiation Induced Defects Across Scales

Luke Sweet

Relating the microstructure properties, in particular those that can be affected by radiation damage, observed at a microscopic level to bulk material observations is challenging. The goal of this effort is the development of a bulk material powder X-ray diffraction (XRD) analysis technique that incorporates observations made at the microscopic level to improve microstructure properties analysis. Such an approach to improving the understanding of microstructure properties of a material across scales will, if successful, allow more timely sample throughput while substantially increasing the quantity of representative material that can be analyzed.

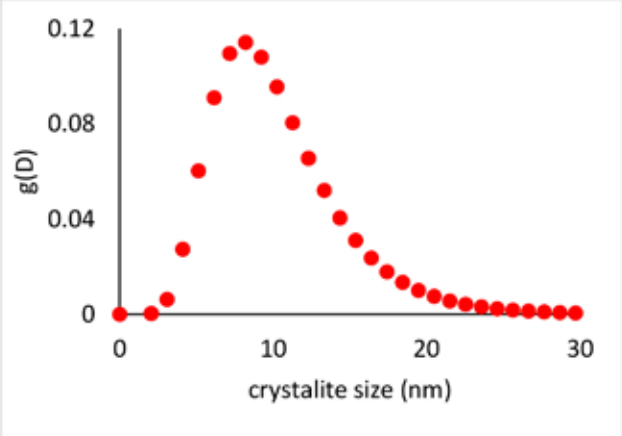
Microstructural characterization, the quantification and description of grains and defects in a solid, is done by a variety of methods. One of the least ambiguous methods to characterizing microstructure in a material is by high-resolution microscopy. Transmission electron microscopy (TEM) can give a clear picture of domains of ordered atoms (grains or crystallites) and various types of defects in the ordering of atoms in a solid. The shortcoming of TEM is that the quantity of material that can be probed is very small (approximately 100 μm^2).

Microstructure characterization by powder XRD analysis can be performed on grams of sample in a manner of minutes to hours; however, the interpretation of the data can be ambiguous and, in cases where a prior knowledge of some microstructure properties are not known, the results can be inaccurate.

An approach to improve microstructure characterization of bulk materials by using insights obtained from TEM analysis in order to come up with physically meaningful evaluations of XRD peak profiles was developed by Paolo Scardi and Matteo Leoni, called whole powder pattern modeling (WPPM), in 2000. The WPPM approach has not yet been applied to actinide containing materials nor used to understand radiation induced defects. This project is working on applying the WPPM approach to improve microstructure characterization of actinides and radiation induced damage.

Started only in the final weeks of FY 2017, the project team has initially focused on gaining familiarity with some common models and software used for WPPM. We have done some modeling on some previously existing data on CeO_2 (a good surrogate for PuO_2). We also explored methods for doing pair distribution function analysis, a complimentary total X-ray scattering technique, using an existing lab-based diffractometer.

In FY 2018, we plan to fully kick off the project by systematically irradiating materials under different conditions and then analyzing the samples by the WPPM approach for XRD analysis using complimentary microscopy techniques to develop models.



A statistical distribution of cubic crystallite sizes (g(D)) for a 1 gram sample of CeO_2. The distribution and shape of the crystallites in this sample were obtained through a best fit using the WPPM method to explain the observed XRD pattern. In the future, we will be using complimentary techniques to inform the selection of models instead of trying models until a best fit is found. Understanding the crystallite size distribution is just one step in microstructure characterization.

PN17109/2999

Correlation of Colloidal Interactions and Macroscopic Rheology in Concentrated Electrolyte Solutions

Jaehun Chun

This project will develop a "science-based" fundamental understanding of macroscopic rheology (flow-properties) of slurry under conditions relevant to storage under harsh conditions and nuclear waste treatment process. In particular, we will connect phenomena at different length scales (i.e., particle interactions, aggregate structures, and bulk rheology). The project will build a scientific framework to accomplish safe and efficient nuclear waste treatment processing, in addition to proper design and operation of waste disposition processes (e.g., pretreatment facility in the Hanford Waste Treatment and Immobilization Plant).

Particle forces in simple colloidal systems have been studied and are well understood; colloidal forces have been modeled using the Deryaguin-Landau-Verwey-Overbeek (DLVO) theory. Resultant rheology of such suspensions has been reasonably understood via the DLVO theory. However, this cannot be simply applied to Hanford waste slurry due to its unique nature such as high salt concentrations and particle size distributions. Furthermore, owing to surface roughness of various minerals in the dispersed phase, rheology can be influenced by friction forces, and solvent structuring can fundamentally influence particle interactions.

This project will consider these unique particle interactions to understand how they arise, in terms of the chemical physics at solid-liquid interface and, especially, to determine how these interactions are correlated to long-range ordering of colloidal waste particles and bulk rheology in tank-waste-relevant-systems. We have begun work to develop an atomic force microscopy (AFM) method to gain insights for particle interactions. Furthermore, we have initiated 1) a simulation framework to predict bulk suspension rheology and 2) scanning electron microscope (SEM) and nuclear magnetic resonance (NMR) spectroscopy implementing System for analysis at the liquid vacuum interface (SALVI) to understand such complicated correlations.

AFM for particle force measurements: We have developed a method for measuring boehmite/boehmite particle forces in aqueous solutions using AFM. The AFM tips are functionalized with boehmite crystals. We have acquired and characterized natural boehmite crystals from the Smithsonian museum and Arizona State University to act as the substrates for the measurements; X-ray diffraction (XRD) and Raman analyses show that the crystals are pure, while SEM and AFM imaging show that the surface is sufficiently large and smooth. Our preliminary data suggest that the particle force increases with increasing ionic strength, which is responsible for a yield stress, according to the particle-based simulations we developed. These results agree with dynamic light scattering measurements that show rapid aggregation with increasing sodium nitrate concentrations. We will conduct more measurements by varying the solution pH, ionic strength, and loading rate in FY 2018.

Growth of large boehmite single crystals: To obtain large boehmite crystals, we initiated the use of a high-temperature vapor-phase approach, namely pulsed laser deposition (PLD), to fabricate large single crystals (> 5–10 microns) of boehmite. Using PLD, we have successfully deposited films composed of mixtures of γ-Al_2O_3, as well as boehmite. Our results showed that amorphous films with stoichiometry similar to the target were produced at deposition temperatures below 200°C. At higher deposition temperatures, film crystallinity increased, resulting in oriented boehmite crystals aligned with the lattice-matched TiO_2 anatase substrate, as well as a polycrystalline secondary phase γ-Al_2O_3. We are planning to optimize deposition conditions in order to synthesize boehmite films consisting of large single-crystal regions in FY 2018.

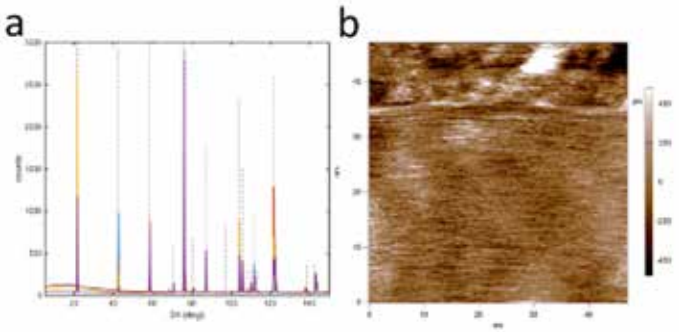
In situ *liquid SEM imaging using SALVI for particles in caustic conditions*: We worked on imaging of boehmite particles (< 100–200 nm), under more alkaline conditions using *in situ* liquid SEM equipped with SALVI. Important optimization includes 1) the use of SEM/focused ion beam to make more observation apertures along the microfluidic channel and 2) both secondary electron (SE) and backscattered electron (BSE) imaging at the same location in a given experiment. Our results show that it is possible to observe particle morphology and aggregates using *in situ* liquid SEM under strongly

PN15091/2766

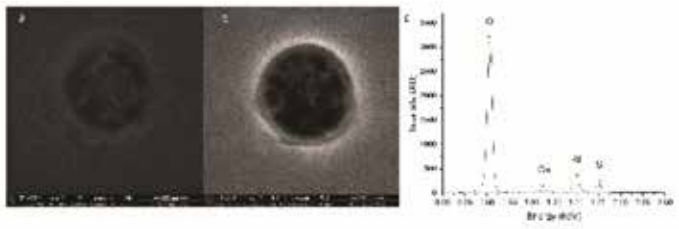
alkaline conditions. In addition to more work with SEM, NMR spectroscopy and magnetic resonance imaging will be used to study the diffusion of colloids *in situ* coupled with particle morphological observations in FY 2018.

Particle-based simulation for bulk rheology: Based on Lubricated Flow-Discrete Element Method (LF-DEM), we completed the simulations for bi-disperse Brownian colloidal dispersions to understand effects of various physicochemical parameters such as particle concentrations and colloidal/friction forces on emergent phenomena (i.e., correlations between scales). Importantly, we found that 1) shear-thickening occurs as a function of a shear rate and strength of cohesive force, and 2) the cohesive force is closely correlated to rheological characteristics through friction force, and 3) the cohesive force is directly related to contact network.

We have been also working on the effects of particle size distributions on rheology, along with a new scheme to implement and interpret particle size distributions with LF-DEM. We will complete this project in FY 2018 and will also work to develop a plan to study the effect of surface roughness of particles on rheology in FY 2018.



Characterization of natural boehmite crystals: a) XRD analysis and b) SEM/AFM imaging.



In situ liquid SEM analysis of pH 12 boehmite particles in deionized water using SALVI: a) SE image, b) BSE image, and c) energy-dispersive X-ray spectroscopy spectrum of the sample.

Determining Radiolytic Transient Intermediates and Interfacial Species and Their Roles in Aluminum Oxyhydroxide Reactivity)

Zheming Wang

In this project, we are developing and adapting molecular spectroscopic techniques to determine transient reaction intermediates and the structure and energetics of the interfacial species, both* in situ *and* ex situ*, which are crucial to understanding radiation-initiated chemical reactions relevant to the nuclear industries, legacy nuclear waste re-processing, and fundamental radiation chemistry. Modern computational techniques are applied to help interpret our data, as well as to guide principled experimental designs.

Ionizing radiation from radioactive decay processes deposits a large amount of energy in a small volume and in a very short period of time, leading to the formation of highly energetic radicals, electrons, and other secondary molecular fragments or simply excited molecules, which further react among them or with other ground state molecules to form stable products. Due to the high energy density, both spatially and temporally, all these reactions occur at very short time scale. For example, for water radiolysis, it is well-accepted that both the initial event of ionization/excitation and the pre-chemical reactions complete within femtoseconds, while the subsequent chemical reactions range on the time scale of femtoseconds to microseconds.

Detailed mechanisms of radiation-induced chemical reactions are rarely known because of the transient nature and overall low concentrations of the intermediate chemical species, as well as difficulties in their identification and quantification. In solid-liquid mixtures, radiolytic reactions are further complicated by the presence of ionization in the solid state, causing formation of various radiation-induced defects and energy transfer processes at the solid-liquid interfaces.

In this project, we will develop a suite of unique spectroscopy techniques, including electron paramagnetic resonance, laser-induced fluorescence, and ultrafast pump-probe laser spectroscopy, to investigate radiation chemistry both in the bulk and at interfaces of solid and liquid in real time. We will use these techniques, the key radiolytic intermediates, their contribution to interfacial reactivity, and their pathways to the final radiolytic products, including their dependence on the type of irradiation and their temporal profiles in solution, in the solid and at the solid/liquid interface for the chemical systems of aluminum oxyhydroxides and zirconium dioxide. The structure and energetics of the interfacial hydroxyl groups and hydration waters under radiation conditions, as well as their dependence on the surface hydration content, will be investigated by sum-frequency generation-vibrational spectroscopy and pump-probe transient vibrational spectroscopy. The ultimate goal is to understand how reaction conditions can be tuned to allow controllable surface properties under irradiation, tailoring to desired solid dissolution, precipitation, or aggregation behavior.

Two types of electron paramagnetic resonance (EPR) have been performed to compare radical production in common forms of alumina: room temperature spin trapping experiments to compare solution state radicals, such as hydroxyl/atomic hydrogen/e-, and low temperature studies of defects in the solid phase. Both experiments used a mercury vapor lamp to provide *in situ* irradiation of the samples with ultraviolet (UV) light.

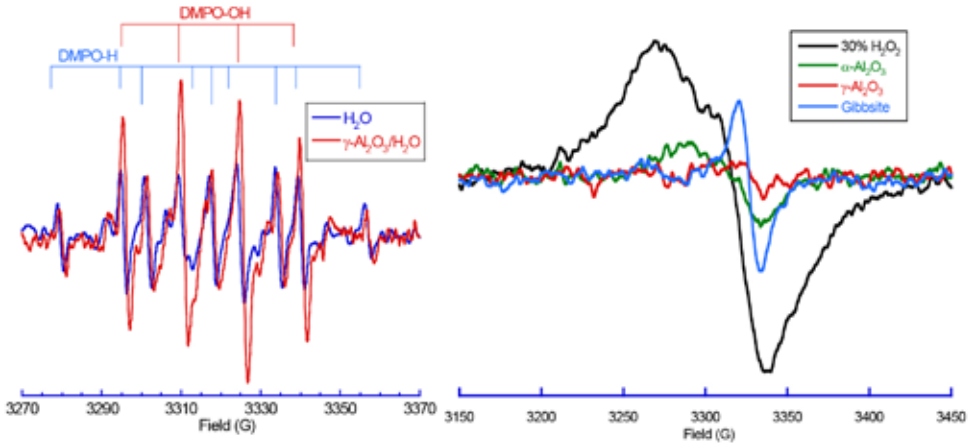
From the spin trapping experiments, it was found that gamma-alumina produced many more hydroxyl radicals than water alone and showed no change in $\dot{H}^{\bullet}/e^{-}$. However, alpha-alumina and gibbsite produced only slightly higher hydroxyl radicals than water, and actually lower amounts of $H^{\bullet}/e$. This would indicate that these minerals have surface species that can scavenge or quench these radicals. Likewise, all three minerals showed different defect spectra when irradiated at

PN17101/2991

low temperature. None of the samples showed detectable hydroxyl radicals under these circumstances, but did show spectra consistent with persistent defects in the minerals themselves. Since we know from the spin trapping experiments that hydroxyl radicals are formed, they must have been able to recombine in spite of being frozen.

We have started to use the Chemsimul software to model water radiolysis, which involves 78 interdependent chemical reactions. Preliminary results indicated that all but three of them match with the reference time-evolution profiles. The discrepancies in the simulation of the time-dependent concentration profile of $\dot{H}^{\bullet}$, e^-, and $HO_2^{\bullet}$ is due to the fact that Chemsimul does not allow the input of three-body chemical reactions. We will investigate some options to overcome this deficiency.

In FY 2018, four sub-tasks will be performed: 1) develop experimental procedures to detect and quantify radiation-generated hydroxyl radical in the ns-µs time scale using laser-induced, time-resolved fluorescence technique; 2) develop/adapt EPR techniques to detect/quantify $\dot{H}^{\bullet}$, e^-, and $HO_2^{\bullet}$ at ms and longer time scales; 3) Chemsimul-based computational modeling will be carried out to simulate the outcomes of the experimental results across the time scales; and 4) develop ultrafast stimulated Raman spectroscopy instrumentation targeting the monitoring of the vibrational spectral signatures of aluminum oxyhydroxide transformation in both aqueous solution and solid phases. All these research efforts are critical to elucidate the reaction mechanisms of radiolytic transformation of aluminum oxyhydroxide.



Room temperature EPR with 5,5-dimethyl-pyrroline N-oxide (DMPO) spin trap (left); positions of peaks for hydroxyl and hydrogen radical adducts are shown as stick spectra. Low temperature (125 K) frozen solution (water) spectra of alumina after *in situ* irradiation with UV (right side). Hydrogen peroxide is shown as a reference for hydroxyl radical in frozen solution.

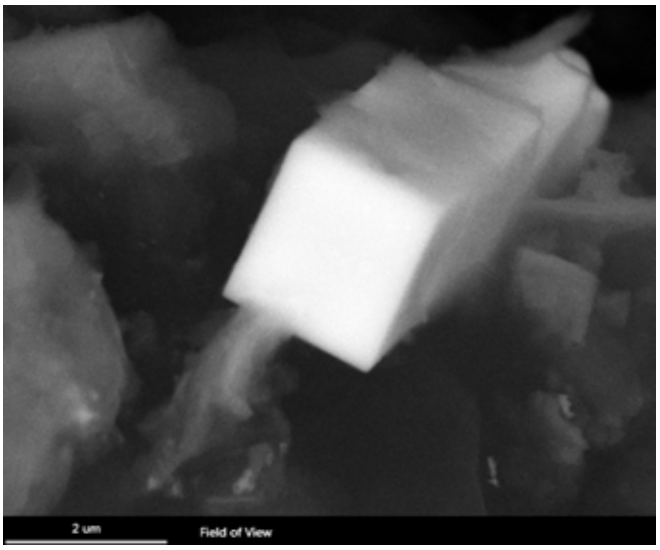
Hot Particle Analysis Aided by a State of the Art Focused Ion Beam

Dallas D. Reilly

This project seeks to use focused ion beam (FIB) technology to subsample highly radioactive materials for analysis, advancing the capabilities for radioactive materials study around PNNL and the DOE complex.

The main intent of this project is to use a combination FIB/scanning electron microscope (SEM) to analyze and subsample highly radioactive materials for movement into laboratories and instruments at PNNL that do not currently or regularly accept highly radioactive samples. The Radiochemical Processing Laboratory (RPL) is a facility at PNNL that has the capability to perform chemistry and analysis on many radioactive materials up to very large quantities. However, it lacks in analytical capabilities in which other parts of the laboratory excel. A state-of-the-art FIB/SEM has been acquired at RPL for subsampling and analysis of radioactive materials. These instruments have been used previously in the semiconductor industry for microscopy sample preparation and analysis, but as radioactive material subsampling tools, they show much promise.

Thus far, this project has focused on various subsets of samples, including low enriched uranium metal fuels, used nuclear fuel remnants, and plutonium samples, with applications ranging from environmental management to national security. Under this project, a "free-release" workflow has been created to reduce sample sizes to essentially non-radioactive levels. The paperwork for these releases is in place, and those samples have been moved to both ultra-trace and completely non-rad facilities for advanced analysis, including nano-secondary ion mass spectrometry (nanoSIMS), large geometry SIMS, atom probe tomography, single crystal X-ray diffraction, and transmission electron microscopy. These analyses have resulted in new science, publication potential, and a high likelihood of follow on collaborations. This is the final year of the project.



Crystalline plutonium particle extracted from Hanford Crib soil. This material is extracted and studied to determine the chemical speciation of radioactive contaminants in the soil.

PN15049/2724

Interfacial Diffusion and Crud Formation at the Liquid:Liquid Interface of Solvent Extraction Processes

Amanda J. Casella

The radiation present during reprocessing of used nuclear fuel breaks down the extractant that isolates uranium and plutonium from fission products. Instead of separation into the organic phase, uranium and plutonium can complex with the degradation products and form crud (solids) at the liquid-liquid interface. This project seeks to gain a fundamental understanding of these complexes in order to enhance processing operations and increase safeguard accountability resulting from the associated materials unaccounted for (MUF).

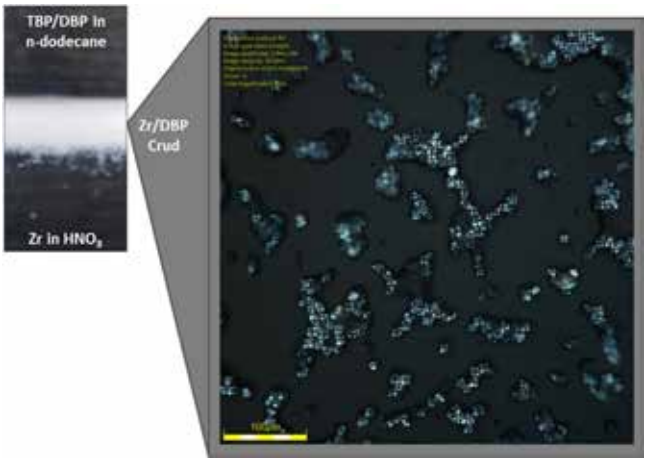
This work targets a fundamental understanding of how crud (solids) forms at the liquid-liquid interface during solvent extraction processes. By using *in situ* measurement techniques coupled with molecular dynamic modeling, the phenomena and drivers leading to crud formation can be captured, providing a more holistic understanding than was previously achievable. The impacts from this study will lead to a stronger characterization linking to processing history, improving process efficiency, and enhanced safeguards accountability due to the complexation with uranium and plutonium.

This 3-year project is currently finishing up year two. The first year focused on an in-depth review of historical Hanford records, along with other open-literature published work, to provide the framework and prioritization for experimental investigations in years two and three.

The focus in year two has been on establishing and validating the macro-scale testing along with implementation of the complementary microfluidic studies. Both systems use *in situ* monitoring by Raman spectroscopy, which enables capture of rapid transients along with detailed kinetics, prior to, during, and after extraction. The macro-scale tests are performed in a Lewis cell, a long-standing/validated method for kinetic analysis. Computational fluid dynamic modeling, also developed on the project, provides the necessary additional fluid transport insight to enable similar kinetic measurements within the microfluidic chip. The microfluidic studies greatly enhance the number of conditions and parameters that can be studied, as they offer more efficient change in parameters along with a six-order-of-magnitude reduction in solution volume.

In addition to experimental work, computational approaches are being used to gain further molecular insights regarding liquid-liquid interfacial processes. Examination of the relationship between adopted configurations of the solute, orientation of the solvent, and the ability of the solute to enhance microsolvation provide clues regarding molecular organization of precursors to solid phase nucleation and formation of interfacial solids. These computational results are being integrated with experimental results in order to gain molecular insight.

Year three will expand conditions/parameters beyond those used for benchmark validation developed in FY 2017. Refined extrapolations between the macro- and micro-scale will further enable the efficient testing that can be performed on the microfluidic platform. At the end of FY 2018, the desired product is a fundamental understanding of the drivers, complexations, extractions, nucleation, and transients that contribute to crud formation at the liquid-liquid interface.



Photograph and optical microscope image of Zr-dibutylphosphate (Zr-DBP) crud, which forms at the liquid-liquid interface during solvent extraction of used nuclear fuel.

PN16090/2867

Modeling of Used Nuclear Fuel Canister Mitigation and Repair Techniques

Nicholas A. Klymyshyn

This research explores the best options to mitigate or repair cracks in spent nuclear fuel (SNF) canisters to ensure safe storage and transportation of SNF. The goal of this work is to apply PNNL's modeling and material science expertise to guide industry to the most effective solutions in dealing with chloride induced stress corrosion cracking (CISCC).

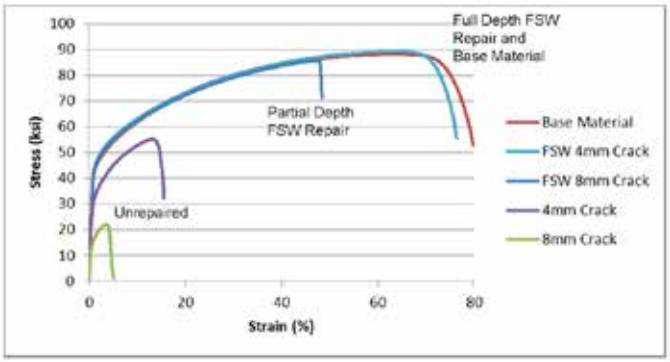
The conditions necessary to cause CISCC of SNF canisters (e.g., tensile residual stress state, environmental conditions) are expected to exist at some dry storage sites in the United States. PNNL has expertise in structural modeling and analysis of dry storage systems under seismic, normal handling, and hypothetical accident conditions. PNNL also has expertise in solid state joining processes that are candidate technologies for repairing CISCC cracks, as well as expertise in materials testing related to CISCC.

This project was initiated to use numerical modeling, informed by joining technology expertise and material science, to evaluate different options for addressing CISCC and to demonstrate PNNL's technical expertise and ability to lead the research needed to solve this critical problem. This work demonstrates the first analytical and experimental steps needed to classify, evaluate, and rate the suitability of a few mitigation and repair techniques, including friction stir welding (FSW) and cold spray. The results of this work will guide future testing and analysis needed to fully resolve this complex problem.

Significant tensile residual stresses are expected to exist in dry storage canisters from the fabrication process, and this is one of the conditions that make CISCC possible. Canister weld regions, specifically, are where through-wall tensile residual stress states are expected, so CISCC is only expected to occur near welds. Repair and mitigation technologies will alter the pre-existing residual stress state when they are applied. One of the ways a repair or mitigation technology can achieve success is by altering the residual stress state such that existing cracks will not propagate and new cracks will not form. Methods to accurately model residual stress were explored, and it was determined that residual stress modeling and analysis is a significant topic for future research.

A number of SNF dry storage systems were modeled and analyzed to predict the response to historical earthquakes. Ground motion data from the U.S. Geological Survey was used to load LS-DYNA finite element models of SNF dry storage systems that approximate, but do not precisely match, existing dry storage systems. The goal was to estimate the range of seismic loading conditions that can be expected on canister walls. The loads were found to be relatively low, but because they are tensile, they have the potential to cause cracks to propagate, especially in regions of high residual tensile stress. It was determined that the canister seismic loads are very sensitive to the design of the storage system, which suggests that mitigation and repair strategies may need to be tailored to specific dry storage system designs.

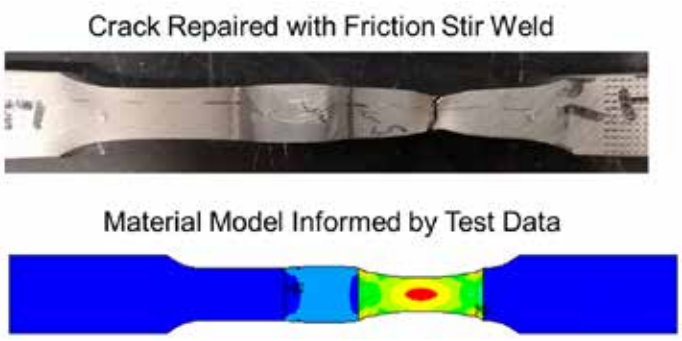
FSW techniques were tested on steel plates that contained simulated cracks. A number of tensile tests were performed to establish the strength of the repairs and to provide significant numerical model validation data using digital image correlation (DIC) methods. The limited data set



Demonstration of the potential for FSW to repair simulated cracks in stainless steel plates. The plate thickness is 9.5 mm. Simulated cracks were 4 mm and 8 mm deep. Unrepaired, the simulated cracks reduced the strength of the uniaxial tension test sample. The Partial Depth FSW Repair intentionally did not match the full depth of the simulated crack, leaving a flaw in the middle of the test sample, but it still achieved a repaired strength that was comparable to the base metal. The tool depth of the FSW 4-mm Crack case was deeper than 4 mm, so the crack was fully repaired by the FSW operation and achieved a strength that was approximately equal to the base material. This test series demonstrates that FSW can achieve promising results, even in cases when the depth of the crack might not be precisely known.

PN17027/2917

created under this task demonstrates FSW as a viable technology for addressing CISCC.



Repair technologies were tested on steel plates with simulated cracks. DIC data of the uniaxial tension test were collected at low and high strain rates for dynamic material model validation. The DIC data records the material behavior through time, providing a basis to match material behavior of repaired cracks under dynamic loading conditions. This figure is an illustration of tuning the material model to achieve results that match the test. In this case, the weld nugget proved to be stronger than the base material, so the yield strength in the base material was reduced, and the simulated tension test results better matched the real tension test results.

Monitoring Diffusion of Actinide Daughters and Granddaughters in Metals for Chronometer Applications

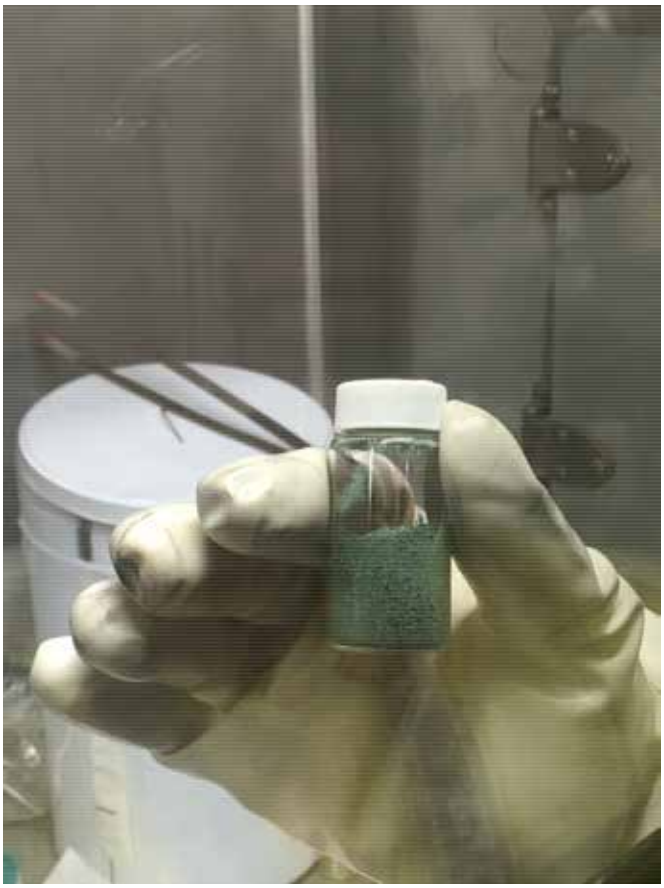
Dallas D. Reilly

This project seeks to understand trace element fractionation at the solid-liquid interface during the production, casting, or melting of uranium (U) metal. These processes are important to the nuclear forensics community for the application of radiochronometers to investigations involving nuclear material.

Past studies have observed trace elements like thorium (Th), the daughter product of U decay, fractionate quantitatively during metal casting. However, the mechanisms controlling this phenomenon are not understood. The goal of this project is to elucidate these mechanisms by combining high-temperature controlled experiments, molecular-scale observations, and theoretical modeling. If successful, determining the fractionation mechanisms of Th during U metal casting will not only confirm the use of U/Th parent/daughter radiochronometers as an endogenic nuclear forensic signature of U casting process age, but also lay the groundwork for interpreting the next generation of chemical process signatures within actinide metals.

Initial experiments involved the doping of Th into uranium tetrafluoride (UF_4, a common precursor to the metal form) prior to metal conversion to better understand this process at the earliest stages of the metal production cycle. These initial experiments were successful in showing quantitative separation of Th from U at trace concentrations, indicating some mechanistic information and a high likelihood that the U/Th chronometer can be applied to U metals. A prototype induction furnace system capable of melting U metal samples has been constructed and will be used to melt U under other conditions to provide similar information for other potential processing methods.

Modeling efforts have been undertaken to both interpret experimental observations and test our understanding of the mechanisms controlling Th fractionation during U metal casting. These efforts have resulted in a publication discussing the molecular dynamics of carbon diffusion, which could play a role in Th fractionation, within U metal.



Mixture of UF_4, Th, and the reactants used to convert U to metal form.

PN15095/2770

Particle-Filter Surface Interactions and Dynamics in the Presence of Cross-Flow

Richard C. Daniel

By direct observation of the accumulation of mineral oxides on sintered steel filtration media, this project will provide unique characterization of the fouling processes that hinder efficient and sustainable filtration of industrial slurries. These measurements, which hereto have been limited to biological systems on morphologically simple filter media, will guide development of filter operational strategies and optimizations that reduce cost and lifecycle for filtration-based solid-liquid separations.

Treatment and disposal of the approximately 55 million gallons of highly radioactive tank waste stored at the Hanford Site in Washington State will cost approximately $110 billion and take 40 years. Hanford waste treatment costs are driven by process bottlenecks that limit waste throughput. One such potential bottleneck is separation of radioactive waste solids from liquids and non-radioactive components by cross-flow filtration. Indeed, bench- and engineering-scale evaluation of cross-flow filters used in Hanford waste treatment have identified waste separations are subject to an unexpected loss in filter performance that results from a slow but persistent fouling of the filter by waste solids. This fouling will necessitate frequent, time-intensive chemical cleaning operations to restore filter performance.

Modification and optimization of the filter process to avoid or minimize such fouling that results in improved performance, even by a modest 5%, could result in billions in saved operational costs and years in reduced mission lifecycle. The goal of the current project is to develop filtration optimizations through improved understanding of waste solid/filter interactions that lead to lost filter performance.

Recent studies of Hanford waste filtration have focused primarily on operational feasibility of waste filtration to prove cross-flow filtration as a viable separation technology for Hanford wastes. These studies have allowed the maturation of cross-flow filtration operations for Hanford waste treatment but have been unable to aide in the development of physical-realistic fouling mechanisms that could form the basis of optimization strategies. Existing studies of Hanford waste filtration infer fouling using indirect methods, such as the rate of decline in the filtration production rate. More direct methods for visualization solid filter interactions require destructive filter testing. *In situ* methods for evaluating solids accumulation, such as real-time direct visual observation using light microscopes, have only been used for biological systems on filter media and at operational conditions not relevant for Hanford operations.

This project aims to apply recently developed visualization techniques to image how solid waste simulant particles accumulate on sintered stainless steel-filter elements in the presence of cross-flow. This imaging, coupled with simulation of particle-filter interactions under flow conditions, will help improve understanding of the unique filtration behavior observed for mineral oxide slurries undergoing cross-flow filtration, lead to the development of improved predictive models, and help engineers improve optimization

FY 2017 research efforts built on FY 2016 work, namely the development of a cross-flow visualization cell (CFVC).

The gear pump originally installed in the CFVC during FY 2016 was upgraded to a 2 HP motor and a rotary lobe pump to enable evaluation of performance at velocities up to 5 ft s^{-1}. The new rotary lobe pump and motor allow PNNL researchers to study the relevance of turbulence and high axial velocities in scouring solids from the surface of the filter and contributed greatly to improved filter evaluations.

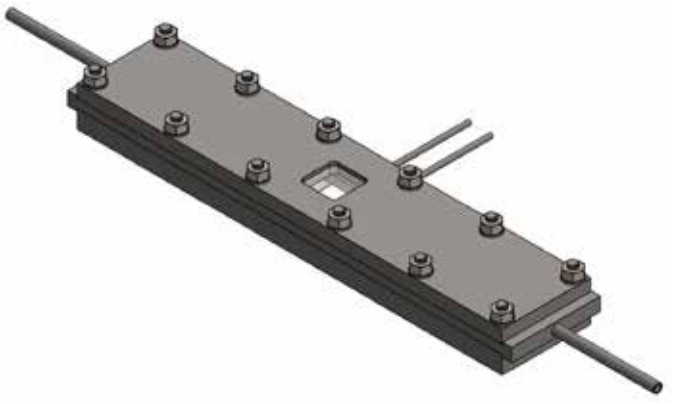
FY20 17 CFVC testing experiments provided basic insight into the cross-flow filtration behavior of solids on the sintered stainless-steel filters used in actual waste separations. Using direct visual observation, CFVC testing attempted to correlate the time rate loss in filter performance with the loading of particles on the filter surface as a function of cross-flow velocity. Two mono-disperse polyethylene (a 5-µm blue fluorescent and a 40-µm red fluorescent) spheres dispersed in water were used as solid particle simulants; while unsuitable as chemical simulants, these two spheres are similar in size to waste particle aggregates and can be visually tracked with ease under an ultraviolet light source. FY 2017 CFVC studies have found that, as expected, filter performance roughly correlates to bead coverage, with high coverages corresponding to lower fluxes

PN16058/2835

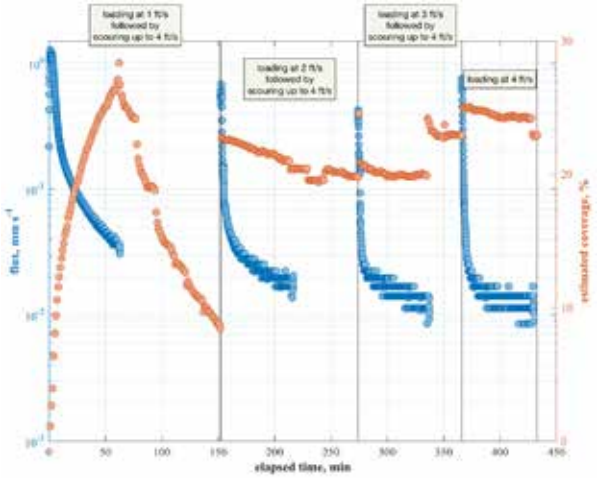
(on average), and the time rate decline in filter performance with time only correlates well with increases in bead coverage during the first stage of bead loading (suggesting further changes result from compaction or re-arrangement of existing beads). We also found that bead coverage does not exceed approximately 20–25% (which is similar to the porosity of the filter medium). Finally, the cross-flow is effective at removing large 40-µm beads, but is unable to effectively remove 5-µm beads (suggesting particles small enough to fit into the surface roughness or flow boundary layer are of greatest concern for fouling).

These results provide new insight into the mechanisms that govern particle-filter interactions and fouling on cross-flow filters, which will enable improved operational strategies for mitigating fouling and accelerating waste clean-up at the Hanford Site.

Continued work in FY 2018 will involve the following: 1) perform well-controlled single component dispersion filtration tests with 1-ppmv dispersions of 5-µm and 40-µm fluorescing glass beads, 2) perform well-controlled filtration tests of bi-disperse mixtures of 5-µm and 40-µm beads, and 3) evaluate the impact of bead concentration and filtration pressure on surface fouling morphology.



Schematic of filtration visualization flow cell.



Permeate product per unit area (flux: left axis, blue circles) and filter bead coverage (right axis, orange circles) as a function of time for a 1-ppmv suspension of 5-µm polyethylene beads.

Phase Field Modeling of Microstructure Development in Pu (IV) Oxalate Precipitation

David Abrecht

In much the same way that no two snowflakes are alike, crystals of plutonium compounds form a variety of shapes and structures based on the conditions under which they were grown. Our research looks to develop techniques to correlate the various crystal shapes of a common plutonium compound, plutonium (IV) oxalate, with conditions (i.e., time, temperature, concentration) under which they grew to aid forensic analysis of interdicted nuclear materials for law enforcement.

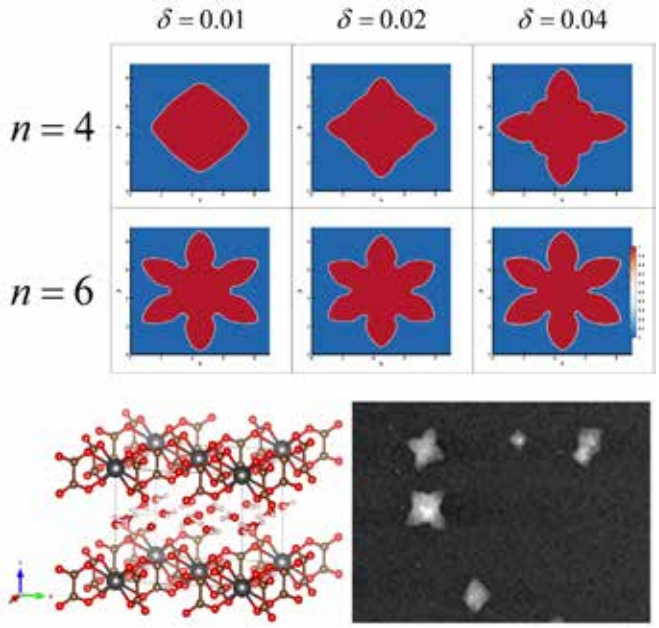
The microstructure of materials, including their grain size and orientation and defect concentration, is responsible for many of the physical properties of materials. As such, crystal growth and grain structure evolution modeling has long been a goal of material scientists attempting to control material properties. However, simulating the moving interface of a material undergoing phase changes has been computationally difficult until recently. Advances in computing power and modeling techniques, such as phase field modeling where the interface is treated as a continuous mathematical field, have allowed these dynamic simulations to become more accessible.

This effort aims to apply phase field modeling techniques to a highly different problem—the chemical precipitation of actinide materials—in order to establish a relationship between crystal shapes observed by microscopy and the conditions of the chemical system, including temperature, concentration of reagents, and reaction time.

This 3-year project has just finished its first year. Year-one efforts have focused on establishing the physical constants for the plutonium oxalate system (i.e., plutonium diffusion rates, thermal conductivity, surface tension, solubility, etc.) needed to describe the crystallization behavior in phase field models.

While many properties are available from the literature, some of the less common parameters have required additional modeling efforts to elucidate. Detailed atomic scale calculations of plutonium oxalate bulk and surfaces are being performed using density functional theory to establish the energy requirement of creating new plutonium oxalate surfaces, a key parameter for establishing the crystal growth rate. These parameters will then be used in the higher-scale phase field modeling calculations to establish the length and time scales over which the model needs to run.

At the phase field modeling (continuum) level, efforts have been focused on establishing the thermodynamic properties of the highly acidic, highly saline solutions from which plutonium oxalate is normally precipitated in order to establish the mathematical relationship between the precipitation rate and the process conditions. Efforts have also aimed at establishing the strength of the surface energy anisotropy required to produce the octahedral shapes observed for plutonium (IV) oxalate by microscopy. Efforts continuing into next year will combine all of these aspects to produce full phase field models, followed by validation of these models for precipitation conditions against microscopy experiments.



Aspects of the modeling effort. (Top) Modifications to the model parameters that generate different crystal shapes. (Bottom right) Efforts aim at reproducing the shapes of plutonium oxalate crystallites seen by scanning electron microscopy. (Bottom left) Atomic scale modeling of plutonium oxalate crystals helps provide insight into model parameters, such as surface energy.

PN17043/2933

Provenance and Pathways Investigation of Uranium Oxide Particles Using Oxygen Isotope

Dallas D. Reilly

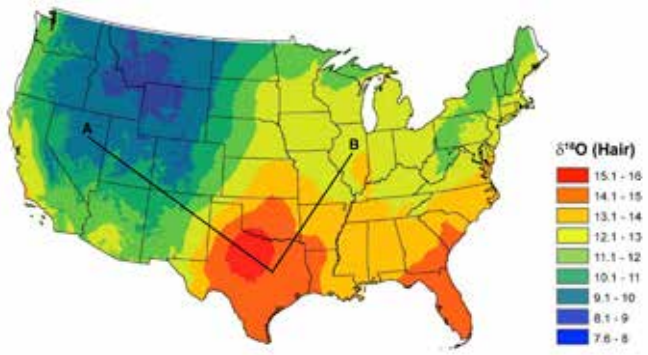
This project is developing new ways in which nuclear forensics scientists can look at material to gather the most information possible for a given investigation. When successful, this research will allow end users to understand questions about nuclear material such as: Where was this material produced? Where was it stored? When was it chemically altered?

This project was initiated because there has been a significant amount of groundwork laid by the geochemistry and climatology communities on oxygen isotope (oxygen-18 compared to oxygen-16) fractionation to study the history of our planet, and there is a potential for using similar studies on materials of interest to the national security community. For example, there is a very good understanding of the oxygen isotope ratio in water vapor across the planet; if a material like a uranium oxide is diverted from a known location and is interdicted at another known location, the water vapor absorbed by that material may indicate which route it took between those points. That information may help nuclear forensics scientists answer important questions relating to their investigation. By the end of this project, we hope to be able to apply oxygen isotope measurements to compounds like uranium oxides to provide qualitative information on its origin.

The main focus in the first two years of this project was to develop capability to measure oxygen isotope signatures on the bulk (approximately 5 mg via fluorination separations) and particle (via large geometry secondary ion mass spectrometry [SIMS]) scales. The bulk analysis is intended for larger samples, when available, and to standardize samples for SIMS analysis since matrix and standard calibrations are vital for quantitative information at the particle scale. Both capabilities were advanced to a preliminary stage in year one and were able to be utilized on non-nuclear materials.

In year two, method development was fully completed. The bulk method selected uses chlorine trifluoride and heat to fluorinate oxide compounds, leaving behind elemental oxygen gas, which is collected for basic gas mass spectrometry measurements. The SIMS system uses an ion beam to remove individual atoms, which get ionized and accelerated to a highly sensitive mass spectrometer. In FY 2017, the bulk method was tested on both non-nuclear (silica) and nuclear (uranium oxides) materials with success. SIMS measurements were performed on standardized silica spheres to better understand fractionation effects that are a common problem for that method. Further, SIMS was used on uranium oxide particle standards.

In FY 2018, both methods will be used to study oxygen isotope ratios on uranium dioxide fuel from two different production locations, uranium metals stored under different conditions, and plutonium oxides produced at two different locations. Further, other production methods will be evaluated to test the utility of these measurements in the real world.



Map of oxygen isotope ratios of human hair in the United States; uranium oxides likely absorb local water vapor in a similar way that the human body does. As an example, if a material was lost at point A, traveled through a region with a high concentration of oxygen-18, and was recovered at point B, investigators may be able to determine a likely travel route.

PN16034/2811

Radiological Atomic Force Microscopy: Coupled Radiation Source/Liquid-Cell Atomic Force Microscopy to Study Radiation-Induced Interfacial Processes

Greg A. Kimmel

This project is developing an atomic force microscope (AFM), the radAFM, to investigate the role of β and γ radiation in driving interfacial chemistry far from equilibrium.

The radAFM will enable proof-of-principle measurements targeting the following science questions: 1) how does the radiation environment impact the structure of nanoscale aqueous films at liquid/solid interfaces, and how does the structure influence the overall interfacial reactivity in the presence of radiation? 2) What is the molecular-level impact of energy deposition from radiation on the interfacial dynamics in highly alkaline environments over long time domains? 3) How do radiolytic conditions influence interfacial dynamics, including dissolution and agglomeration/precipitation? What are the fundamental mechanisms and pathways for these processes?

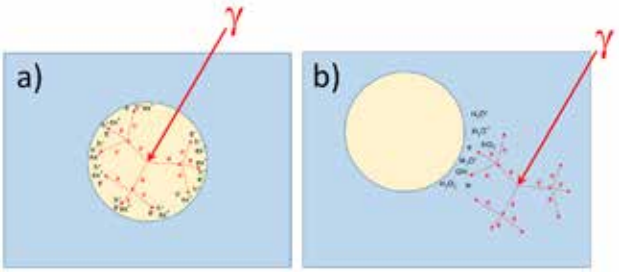
Determining the effects of ionizing radiation on interfacial dynamics is essential for developing a comprehensive understanding and predictive capabilities of the unique interfacial chemistry relevant to the materials currently stored in high-level waste (HLW) tanks.

Previous studies for relevant oxide/water systems clearly indicate significant impact of ionizing radiation on chemical and physical characteristics of both components (solid and liquid) of the dispersion. Two major mechanisms are typically mentioned to describe radiation effects in aqueous heterogeneous systems

Solid-based: Electronic excitations (electrons, holes, excitons) and radiation defects (structural and electronic) generated in the solid phase can react with molecules of the liquid phase at the interface. For example, the radiolytic yield of H_2 at oxide/water interfaces can be enhanced by more than an order of magnitude due to reaction of excitons and/or electrons with the adsorbed water molecules.

Liquid-based: Reactive radiolytic species from aqueous phase can attack the surface of the dispersed solid, affecting solubility, nucleation, growth particle-particle interaction, and many other important interfacial processes.

In spite of the significant potential for radiation-induced reactions for water/solid dispersions, we cannot model or predict the effects of radiolysis for specific systems, such as the Hanford HLW, based on the existing literature. The overall objective of the proposed research is develop the radAFM capability to enable fundamental insight into the role of ionizing radiation in interfacial chemistry relevant to HLW and to help ultimately predict the impact of ionizing radiation on long-term aging of the waste. We have designed and built a novel radAFM facility based on an AFM microscope coupled to a commercial X-ray source.



Two major mechanisms of radiation effects in aqueous heterogeneous systems: absorbed γ-photon (or β-particle) creates cascades of secondary electrons scattering through medium. Primary scattering event may happen in a) solid orb) liquid phase, initiating different mechanisms of radiation impact. Once electron energy decreases < 100 eV ionization and excitation of medium becomes most efficient, the produced highly reactive species (electrons, holes, excitons in solid and molecular ions, radicals, reactive molecules, hydrated electrons in liquid) become the radiolysis driving force.

In the initial configuration, a liquid-cell AFM is coupled to a compact X-ray source via a thin window on the bottom of the liquid cell.

The primary accomplishments for FY 2017 included designing, constructing, and testing the radiation shielding needed for safe operation of the radAFM. Prior to construction,

PN17039/2929

Monte Carlo simulations of the X-ray dose distribution were performed to verify the design parameters. Tests of the radiation levels outside the enclosure during operation of the X-ray source were conducted and verified to be well below acceptable limits.

In addition, the microscope base for the AFM has been modified to allow mounting the X-ray source below the x-y translation stage, and the radiation shield has been installed on the vibration isolation table for the AFM. We developed the operational protocols, training procedures, and safeguards needed to operate the X-ray source, and dose calibrations for the X-ray source using radiochromic films are currently being conducted.

We have begun trial irradiations and are working on solving technical issues needed to achieve full operational capabilities. These technical issues include, for example, thermal drifts associated with the heat load from the X-ray source, the ability to relocate individual features before and after irradiation, and particle mounting techniques.

To develop protocols for imaging irradiated nano-particles, batches of boehmite and gibbsite particles were irradiated and then compared with unirradiated samples of the same material. Initial analysis suggests that gross morphological changes were not observed for samples irradiated under vacuum, but the fracture characteristics may have changed (i.e., different yield strengths were observed).

Signatures of Plutonium TetraFluoride and Plutonium Metal Processing

Karl Pitts

The goal of this project is to reconstitute subsets of Hanford's Plutonium Finishing Plant (PFP) processes used to produce plutonium (Pu) metal and recover Pu from various waste forms.

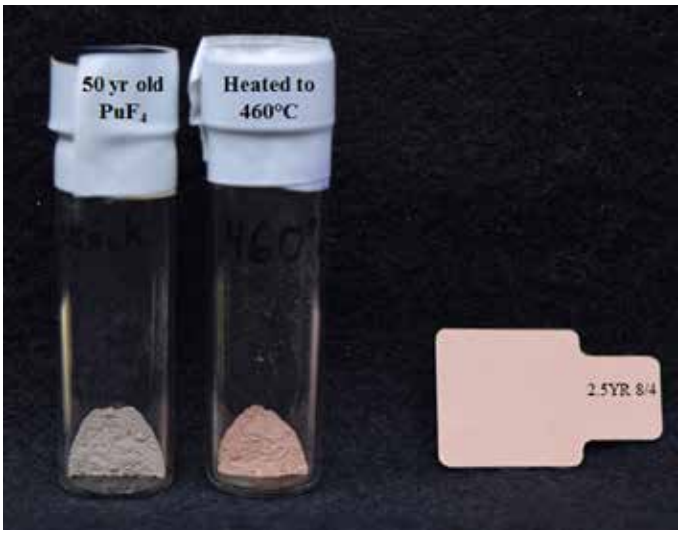
The production of plutonium metal from plutonium tetrafluoride (PuF_4), its shaping, and the recycling of scrap materials are all of interest to the nonproliferation community. While these production operations can be indicative of weaponization activities, they might also have legitimate application to produce metallic nuclear fuel for fast reactors.

This project was the study of the legacy PuF_4 reduction process developed at Los Alamos and applied to the bulk of plutonium metal production at Hanford and other sites in the American production complex. Its scope was later extended to selected features of uranium production with UF_4. The two most important outcomes of this project are the successful production of small scale (about 1–10 grams) samples of both Pu and U metal and new insights into the radiation-induced structural damage of the legacy PuF_4 during the approximately 50 years after its production at Hanford.

This legacy metal reduction process uses the reaction of PuF_4 feedstock with calcium metal. Calcium metal, iodine booster, and PuF_4 are placed in a ceramic crucible and heated in an induction furnace using process parameters derived from the Hanford production process. At several hundred degrees, the iodine reacts with the calcium, igniting the PuF_4 and calcium in a reaction similar to thermite. The dense metal separates from the calcium fluoride slag and cools to a metal button.

The structural damage study was required to understand the material for reaction studies and as reference archive material. This particular sample was welded in stainless steel pipe and used as a sealed neutron source after 1972. When opened in 2010, the stored material was a gray powder with an amorphous crystalline structure rather a salmon pink powder with a monoclinic crystalline structure. Characterization of this material was required for this project, as well as other applications such as material libraries. After demonstrating that thermal annealing restored the crystalline structure, a new task was justified and added in FY 2017 to prepare fresh PuF_4 and surrogate UF_4 for further damage studies.

During FY 2015, the major accomplishments were the analysis of the legacy PuF_4 feedstock and the design of the reduction system. The chemical analyses confirm that the material was chemically separated from other radioactive materials in late 1966, with subsequent conversion to PuF_4. Numerous characterizations showed that, while the PuF_4 has undergone metamictization during its 50 years of storage as a reference neutron source, it retained its particulate morphology and structure known from historic studies performed at the PFP. The reduction system was designed for installation in a plutonium-certified glovebox and tested with a CeF_4 surrogate feedstock. The resulting button had a high metal yield and clean separation from the slag.



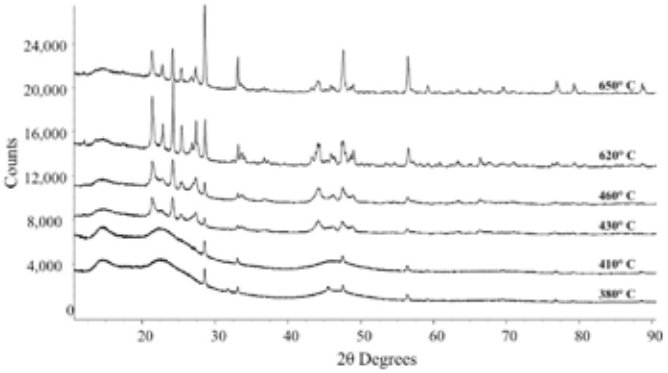
Color changes of the Hanford PuF_4 material after heating and restoration of crystalline structure indicative of Hanford "pink powder."

During FY 2016, the reduction apparatus was installed in a glovebox for plutonium operations. Both a 1- and 10-gram plutonium button were made with high metal yield. In addition, the project scope was expanded to include production of uranium metal as a plutonium surrogate with nearly identical process parameters. This enabled conservation of the legacy PuF_4 for additional feedstock analyses, other new projects, and, in particular, an extensive set of measurements to further investigate the measurements of structural damage.

PN15037/2712

During FY 2017, the major tasks were an in-depth study of structural damage in legacy PuF_4 and initial progress toward producing new tetrafluoride feedstock. The key characterization issues were the material's crystalline structure, oxide content, and moisture content using measurements of differential thermal analysis, X-ray diffraction (XRD), and other chemical measurements. The legacy material was predominately amorphous with an oxide content near 10 percent. During heating, this material had a reproducible exothermic event, which occurred near 414°C. After heating, the color of the material was the bright pink expected of the "pink powder" produced at Hanford. The underlying structural change with higher temperature annealing was measured using XRD. There is a clear restoration of crystalline structure, as shown in the measured diffraction plots as a function of final temperature.

Further work in FY 2017 was the production of UCl_4 from other chemical forms of uranium. Approximately 50 grams of clean UCl_4 was prepared using anion exchange. Both precipitation of UF_4 in hydrofluoric acid and pyrolysis of $U(CF_3CO_2)_4$ are being implemented, with XRD analysis in progress.



These plots show the increase of crystallinity, as shown by the sharp X-ray diffraction peaks, as the aged Hanford-produced PuF_4 material is heated. There was an associated color change from gray, indicative of amorphous material, to salmon pink, indicative of the high-quality, crystalline material produced with good process control.

The project is now complete, and these restored capabilities transitioned to new projects that address the production of and similarities between plutonium and uranium metal.

Solving the Plutonium-238 Problem

Jeffrey A. Katalenich

PNNL is developing next-generation fabrication techniques for radioisotope heat sources to simplify processing and enable future missions that rely on plutonium-238 for heat and power.

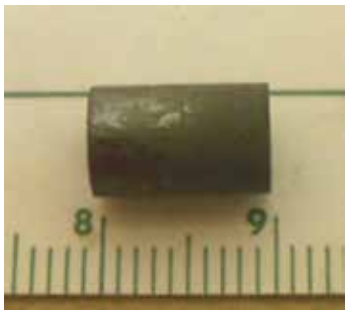
Plutonium-238 (^{238}Pu) has unique properties that make it useful as a radioisotopic heat source. Coupled to electricity conversion devices, such as thermoelectrics, ^{238}Pu can provide highly reliable power on scales from mW to 100 s of watts, with lifetimes as long as several decades without maintenance.

However, use of ^{238}Pu is constrained due to its cost of production and challenging processing requirements. To obtain plutonium with high ^{238}Pu assay, neptunium-237 targets must be irradiated in a nuclear reactor and then chemically processed to recover and purify the plutonium. Past experiences at Mound, Savannah River, and Los Alamos to process separated ^{238}Pu into oxide heat sources have shown that ^{238}Pu poses unique hazards, particularly as a fine powder.

Current techniques for producing ^{238}Pu oxide heat source pellets at Los Alamos are based on historic powder processing methods that generate highly dispersible plutonium powders. Processing ^{238}Pu oxide powder in gloveboxes and hot cells results in plutonium dust dispersal and increased potential for a breach in the engineered containment and worker exposures.

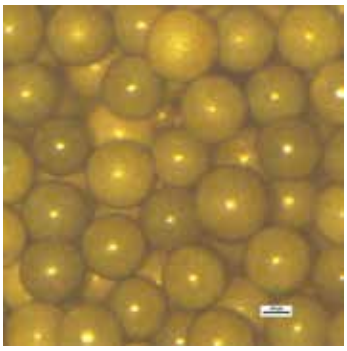
To solve this problem, PNNL developed sol-gel techniques that prevent the generation of dispersible plutonium powders. Using sol-gel, plutonium is kept in solution until macroscopic oxide microspheres are formed, eliminating the powder generation and handling steps. Plutonium oxide microspheres may be used directly as heat source granules or could be pressed into pellets.

Research and development thus far has led to the production of cerium oxide, neptunium oxide, and plutonium oxide microspheres. Additionally, cerium oxide microspheres, which serve as a non-radioactive surrogate for plutonium, have been hot-pressed into a pellet.



Cerium oxide microspheres were hot-pressed into a 1-cm long pellet by collaborators at the University of Dayton Research Institute.

Internal gelation sol-gel technology, which was developed to produce uranium oxide microspheres for high temperature gas reactor fuels in the 1970s, was modified to enable the fabrication of uniform, crack-free 238Pu oxide microspheres. ^{238}Pu sol-gel microspheres, imaged eight weeks after production and calcination, did not exhibit any observable degradation. Based on work to date, sol-gel processing would dramatically reduce the generation of ^{238}Pu dust, simplifying operations and mitigating contamination and worker exposure risks. It is anticipated that cleaner ^{238}Pu processing methods would reduce material holdup, which would decrease worker dose, reduce degradation of engineered containment, and lessen decontamination and decommissioning liabilities.



Plutonium-238 oxide microspheres made using sol-gel methods with an average diameter of 250 ± 16 µm.

With basic process development and demonstration complete, PNNL is ready to transition sol-gel activities to process refinement and scale up. Current equipment capabilities are limited to several grams production per day. Additionally, material analyses are required to characterize material properties relevant for microsphere pelletization.

PN15027/2702

Tunable Irradiation Testbed

Andy Casella

This project explores how a tunable irradiation testbed, based on an off-the-shelf cyclotron with customized targets, could be used to study many current topics of interest in the field of radiation materials science. These topics include irradiation signatures and impacts on material performance; prediction and control of interfacial phenomena; design of radiation resilient materials; the coupling of chemical dynamics with energy deposition from ionizing radiation over relevant spatial and temporal scales; and the evolution of materials degradation, transient species, and material behavior in radiation environments.

Much is known about the high intensity operation of cyclotrons; many associated issues were first addressed in the 1950s, and the fundamental feasibility of accelerated beams of the order of several mA has been established for cyclotrons. Off-the-shelf cyclotrons of approximately 30 MeV and > 1 mA tailored to medical uses, such as medical isotope production and patient treatment, are commercially available.

Previously, PNNL investigated a unique accelerator target system coupled to such an accelerator that could generate an intense neutron spectrum with an energy distribution that could simulate the environments of fast, fusion, and thermal reactors. The present study is focused on how the proton beam could be directly applied to address materials problems important to high energy physics. This project will provide performance information to enable the fabrication and testing of materials of interest in nuclear reactor, particle accelerator, and outer-space environments under irradiation conditions of interest for subsequent post-irradiation examinations.

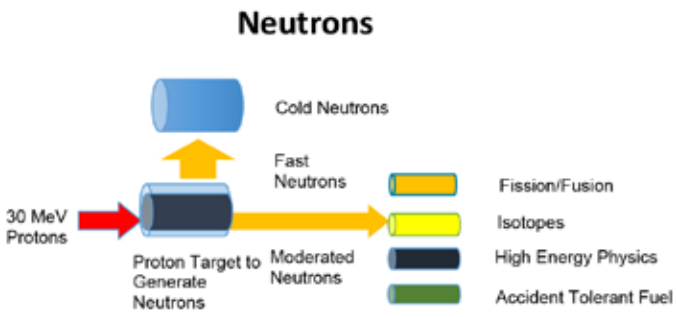
During FY 2017, customized accelerator target systems were developed for a variety of both neutron and proton irradiation applications. Particular target configurations were shown to allow characterization of radiation damage in materials applicable to signature science and legacy waste, as well as the development of resilient materials for fast, thermal, or fusion reactor applications without the complexity and expense of a reactor irradiation. Neutron spectra could be tailored to approximate that of a high temperature gas reactor, a molten salt reactor, a light water reactor, and a liquid metal reactor. Additionally, targets were developed to assist in tritium production research and development (R&D) to better understand phenomenology, ^{252}Cf neutron source replacement, characterization of radiation damage to materials of interest to facilities characterizing neutrinos, and isotope production R&D by protons and neutrons.

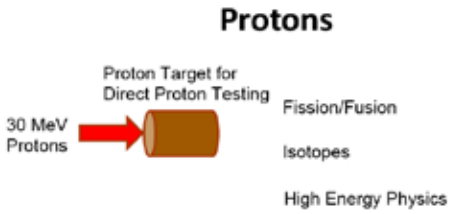
The characteristics of the direct proton beam target were developed to allow the capability of rapid target material change-out, removal of 30 kW of heat from a small area, maximization of the target volume, and full utilization of the proton beam. Material damage in terms of displacements per atom and gas (hydrogen and helium) production were compared between the 30 MeV cyclotron beam and the GeV accelerator beams used in high energy physics. Our results show that, depending on the materials, comparable damage rates can be obtained. Hydrogen and helium gas production rates are generally lower, but can still be useful.

The study developed a proposed cyclotron-based irradiation facility that would be a single shielded building with multiple beam lines from a small, flexible, high flux irradiation source. The proposed system has several key advantages compared to a reactor or isotope source, including on/off capability, independence from source decay, and re-configurability.

In our vision, the ultimate capability would contain a conventional, approximately 55-ton cyclotron with the ability to generate a 30-MeV proton beam with a 1-mA current, and it would accommodate up to 4 end stations for proton or neutron irradiations. Space should be made available for sample preparation and examination, a chilled water system for cooling, a control room, a maintenance and service room, and a shipping and receiving area.

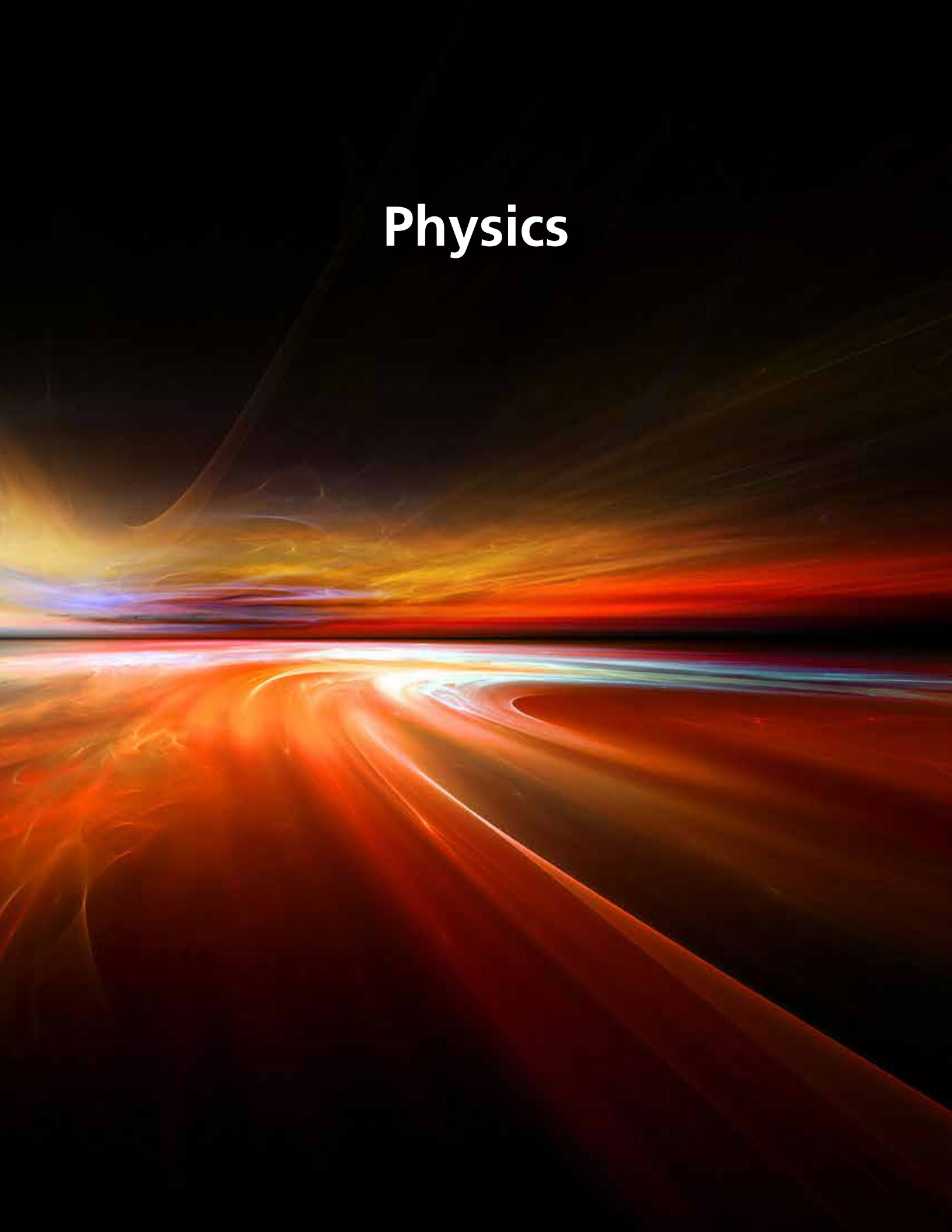
PN17075/2965





These plots show the increase of crystallinity, as shown by the sharp X-ray diffraction peaks, as the aged Hanford-produced PuF_4 material is heated. There was an associated color change from gray, indicative of amorphous material, to salmon pink, indicative of the high-quality, crystalline material produced with good process control.

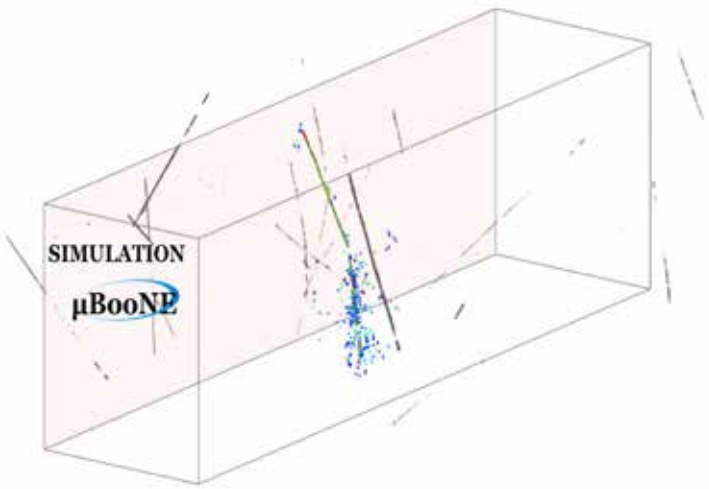
Physics



Accelerator Neutrino Physics in Liquid Argon Time Projection Chambers

Eric D. Church

This project seeks to understand the nuances of the detectors that allow one to make very accurate cross-section measurements and oscillation measurements, and eventually, be able to demonstrate development of data acquisition systems that will allow for paradigm-changing proton decay measurements at the Deep Underground Neutrino Experiment (DUNE) detector in South Dakota.



A simulated proton decay event in MicroBooNE. Overlaid in gray are the cosmic rays that plague surface detectors and that must be subtracted to find the signal event. MicroBooNE will not discover proton decay; however, it will develop the techniques for discovery in future underground LArTPCs.

The U.S. High Energy Physics (HEP) plan is committed to a long-term investment in Liquid Argon Time Projection Chambers (LArTPCs) and related facilities. These unique detectors offer unprecedented event resolution and full calorimetric reconstruction. The investment is for large international collaborations to measure properties of neutrinos arriving from accelerators at Fermilab to detectors such as MicroBooNE (micro Booster Neutrino Experiment). Other goals of the long-term LArTPC program include measuring supernova neutrinos, atmospheric neutrinos, and, importantly, looking for evidence of baryon number violation. Such an observation would be paradigm-changing, insofar as understanding the physical laws of nature.

Expertise at PNNL in the relevant physics analyses and in data acquisition systems suggests PNNL is uniquely positioned to make valuable contributions to this program. Nucleon decay searches, particularly, require sophisticated data acquisitions (DAQs) and triggers, as well as advanced real-time and offline reconstruction and simulation techniques.

PNNL researchers performed key MicroBooNE detector service work, as well as engaging fully in nucleon decay signal searches for DUNE and MicroBooNE. Additionally, they worked on implementation of algorithms for modern electronics, with an eye to future LArTPC DAQs.

A crucial issue to establish for the future LArTPC program is that nucleon decay backgrounds are, indeed, low at DUNE. Given how pernicious cosmic ray events are in the surface MicroBooNE detector, one must demonstrate that they do not remain so at DUNE. Improvements on particular branching fraction sensitivities that surpass the global limits set by SuperKamiokande on the 2035 timescale are expected to be on the order of factors of six. Theoretical considerations and novel reconstruction techniques suggest that a discovery of paradigm-changing nucleon decay is possible in DUNE.

Commercial off-the-shelf PCIe cards in a commodity server likely represent the future of LArTPC DAQs. This is a natural conclusion of using FELIX [3]-like DAQs, which will likely hold the entire data of one anode plane assembly in one commodity server. PNNL is in possession of a high performance DAQ server in a teststand holding a FELIX card and a Brookhaven Warm Interface Board (WIB). WIBs are the means of readout in both Short-Baseline Near Detector and proto-DUNE and, quite likely, DUNE itself. As a step toward producing a realistic deep-learning-trained "Level 2" trigger for off-beam nucleon decay and supernova events that will be a candidate for large underground LArTPC DAQs, PNNL researchers worked on FELIX DAQ systems and field-programmable gate array boards in a teststand on this project. That teststand will continue to be the basis of future triggering systems for advanced DAQs for HEP and national security applications.

PN16081/2858

Advanced Detection Techniques for a Linear Collider Detector

Jan F. Strube

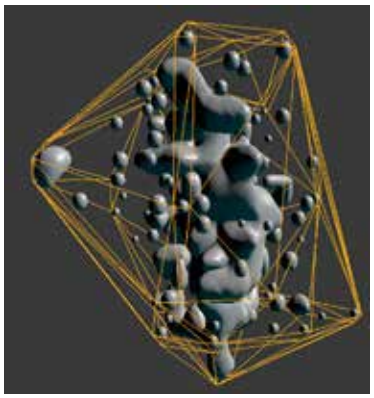
This project proposes to develop technology to improve the reconstruction performance of next-generation detectors at high-energy particle colliders. The physics aim of this project is to increase the precision of Higgs boson measurements and the detection capabilities for signatures of dark matter.

Detector concepts for the next generation of linear accelerators have been designed for precision measurements of Higgs boson properties and to be sensitive to a wide range of new phenomena. In particular, they measure the energy of quark jets with a resolution of 3.5–5% over relevant energies, compared with 10–30% at the Large Hadron Collider (LHC) Compact Muon Solenoid (CMS) experiment. To this end, the CAlorimeter for LInear Collider Experiment (CALICE) collaboration has developed sampling detectors with extremely fine granularity—a technology that was also chosen for the upgrade of CMS. The information that these detectors provide per particle is at least an order of magnitude higher than what the current generation of detectors provide. This additional level of detail calls for new algorithms to fully exploit the potential of these detectors.

PNNL is developing algorithms that bring techniques from image reconstruction and recognition into the domain of event reconstruction for particle physics experiments. These algorithms will be developed with current and future generations of computing hardware in mind to facilitate efficient running at the time when the experiment becomes operational. The developments of these algorithms will be guided by physics simulation studies and evaluated on beam test data from the CALICE collaboration.

Since the start of the project in April 2016, the preparation phase of the project has successfully completed. PNNL is member of the CALICE collaboration, which affords the group access to the data samples of the test campaigns and allows the validation of reconstruction algorithms on measured data.

We have started to design an analysis of invisible Higgs decays that will eventually be used to benchmark improvements to the reconstruction. This analysis is extremely challenging at the LHC and its upgrades, and it benefits from the precision available at a next-generation lepton collider.



Visualization of the three-dimensional measurements of a 50 GeV pion in a CALICE detector.

We have created a first prototype implementation of algorithms that take techniques from computer graphics and apply them to the reconstruction of events in a collider detector. Preliminary attempts at using these techniques to classify the particles look promising and encourage us to develop the methods further.

PN16083/2860

Atomic Tritium for Project 8

Brent A. VanDevender

Determination of the absolute neutrino mass is one of the priorities of modern nuclear and particle physics. Although extremely light, neutrinos are also extremely abundant, possibly outweighing all of the normal matter in the universe and affecting the evolution of large-scale structures such as galaxy clusters. This project supports the development of an atomic tritium source for Project 8, the next-generation experiment to measure the mass of the neutrino.

The 2015 Nobel Prize in Physics was awarded for the discovery of neutrino oscillations, which show that neutrinos have non-zero mass, in contradiction to the Standard Model of Particle Physics. However, that discovery does not determine the absolute mass, apart from setting a lower limit of about 0.01 eV/c^2. Past experiments to directly determine the absolute mass have not yet had sufficient sensitivity, resulting in an upper limit of 2 eV/c^2, with a new experiment coming online that will reduce that to 0.2 eV/c^2 by about 2022 if neutrinos are still lighter than that. Even between those bounds, the implications of neutrino mass for the evolution of the universe and the development of a quantum field theory to replace the Standard Model of Particle Physics are profound. Project 8 anticipates sensitivity to neutrinos as light as 0.04 eV/c^2.

The atomic tritium source development supported by this research will improve the sensitivity of neutrino mass measurements by removing an irreducible systematic experimental uncertainty that appears in any experiment done with a molecular tritium source. Like the experiments that set the current upper limits, Project 8 will use the tritium endpoint method, which determines neutrino mass by its effect on the energy spectrum of electrons emitted in the beta decay of tritium.

In its natural state, tritium gas exists as a two-atom molecule. When one of those atoms beta decays, the daughter molecule rotates and vibrates, resulting in a smearing of the electron energy spectrum. That smearing limits sensitivity to the mass of the neutrino also emitted in the decay. That limit can only be removed by using a source consisting of individual tritium atoms. Confining those atoms in the sensitive region of an apparatus and preventing them from recombining to molecules is a major challenge to be addressed by this research.

The first step is to develop a conceptual design of the source and an understanding of how its parameters affect the neutrino mass sensitivity of the Project 8 experiment. The pressure of tritium in the experiment will be very low, so that recombination of atoms to form molecules will happen only on the walls of the containment vessel. We plan, therefore, not to have a physical vessel at all, but rather a magnetic field configuration that will confine magnetic tritium atoms while nonmagnetic tritium molecules quickly escape. Such a source by itself would self-heat because of its radioactivity and evaporate away, so we will add a small amount of nonmagnetic helium, which will pass freely through the magnetic trap to exchange heat with the cryogenic physical vessel.

Last year, we developed a plan for a simulation-driven design effort. That plan will address the trapping and cooling issues above and allow us to specify quantitative design parameters, such as the required purity of tritium atoms relative to molecules and the strength of the magnetic trap. We chose an existing simulation framework that will naturally handle the magnetic trapping of atoms and the paths of decay electrons in the magnetic field. That framework must still be extended to include the thermodynamics of the source to account for the helium heat exchange and thermal evaporation of tritium atoms from the trap. Quantum mechanical interactions that could result in additional trap losses must also be added.

The design parameters will be driven by their effect on the neutrino mass sensitivity of the experiment. We must, therefore, also develop the uncertainty budget of the experiment, so that we can complete a quantitative conceptual design with the ability to map the design parameters to their contributions to the total uncertainty. Major progress was made this year coordinating efforts to analyze data from a prototype instrument. That data analysis is a critical ingredient to the understanding required to develop the rigorous uncertainty budget. Significant technical progress was also made in the development of a data acquisition framework for the future generations of the Project 8 experiment.

PN16080/2857

Axion Dark Matter Experiment (ADMX)

Noah S. Oblath

Axions are one of two theoretically well-motivated candidates for the dark matter that is known to make up about one quarter of the total mass of the universe. The ADMX experiment is searching for cold dark matter in the local galactic dark-matter halo. The existence of the axion would not just solve the mystery of dark matter, it would also explain an apparent "fine tuning" question in particle physics, known as the Strong CP Problem. At the heart of ADMX is a tunable microwave resonant cavity that increases the ability to detect an axion of a particular mass by a factor of approximately 10,000.

A variety of techniques have been used to try to detect axions, but only ADMX would be able to detect an axion with the appropriate properties to make up a significant portion of the cold dark matter in the universe. Those axion properties include its mass and the strength with which it interacts with normal matter. ADMX searches for different axion masses by changing the resonant frequency of the cavity.

The ADMX experiment started in 1995 and has been relentlessly working to improve its ability to detect axions. The current experimental plan will cover the theoretically best-motivated range of axion properties, and PNNL has capabilities that will enable ADMX to be successful.

In 2015, the ADMX collaboration went through a process of expanding the collaboration membership to bring in new expertise that would be needed to enable the experiment to be sensitive to higher axion masses. They sought the involvement of PNNL, in particular to take advantage of our expertise in microwave engineering, as well as to grow the existing relationships between PNNL and the University of Washington (where the ADMX experiment is based). We will use our expertise in microwave engineering to provide the receiver electronics that process the signals coming out of the experiment and to do the modeling of the microwave resonant cavity that helps us understand the cavity resonant modes that are critical to the experiment's success.

This fiscal year, the main contributions to ADMX from the PNNL group included work on cavity modeling, developing a prototype cavity-locking system, and analyzing the data to search for axion signals. The ADMX collaboration collected data from January through June 2017. Though the full data analysis is still in progress, preliminary estimates show that ADMX has made the most sensitive axion search ever. The PNNL group is participating in the analysis of the data collected this year with both the main cavity and the higher-frequency prototype cavity that was run simultaneously.

The next phase of the ADMX experiment presents a variety of challenges in the coming years. The PNNL group worked on several fronts to address these challenges

A new design for the microwave cavity is required to be sensitive to higher axion masses. We participated in the design process for the new cavities using advanced radio-frequency simulation tools.

Multiple cavities will be used simultaneously, which requires adding a system to lock the resonant modes of the cavities, such that they are sensitive to the same axion mass range. PNNL has expertise in cavity locking, and we built one of the prototype systems for cavity locking. Based on previous experience from a different project, we proposed and built a simple and reliable digital locking system. That prototype will be evaluated by the collaboration, and PNNL will be responsible for building the full cavity-locking system in the coming year.

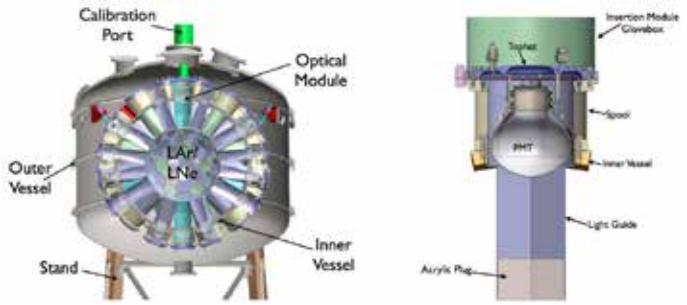
To support our increased participation in ADMX, in March 2017, we hired a postdoc who was already a central figure in the operation of the ADMX experiment. In August 2017, we hired a second postdoc, who will be stationed in Seattle at the site of the experiment. He will be involved with work on the cryogenic electronics systems. We continued to work on improving the data-acquisition and experiment controls systems, as well as developing new analyses to increase the scientific output of the ADMX collaboration.

PN16087/2864

CLEAN Detection of Dark Matter and Low Energy Neutrinos

Christopher M. Jackson

Monolithic detectors using noble liquid targets hold promise to significantly extend the sensitivity and potential discovery of particle dark matter and new physics in the neutrino sector well beyond the state-of-the-art anticipated for the Generation-1 (G1) and G2 experiments. The efficacy of the so-called "single-phase" approach for the detection of dark matter and neutrinos will be examined with the MiniCLEAN and DEAP-3600 experiments operating deep underground at SNOLAB, and the "dual-phase" approach will be examined with DarkSide-20k at Gran Sasso.



(Left) Model of the MiniCLEAN central detector with its 4π target viewed by 92 optical cassettes. (Right) The optical cassettes are 30-cm long and consist of a 10-cm thick acrylic plug, the front surface of which is coated with a wavelength-shifting fluor (tetraphenyl butadiene [TPB]) and 30-cm light guide leading to the photomultiplier tubes. The inner target, defined by the TPB surface, contains 500 kg of LAr within a nominal radius of 44 cm.

The astronomical and cosmological observations of dark matter in the universe are well established and are one of the very few places where there is direct experimental evidence for physics beyond the standard model. Primordial nucleosynthesis argues that the dark matter is non-baryonic; observations of structure formation indicate only a small contribution from neutrinos. Therefore, dark matter is exotic, and Weakly Interacting Massive Particles (WIMPs) are a compelling candidate that provide a cold, dark relic of the Big Bang and could be directly detected as they scatter from massive, ultra-pure detector targets operating deep beneath the earth's surface.

The challenge to realizing sensitive dark matter detectors is in separating backgrounds from the nuclear-recoil events characteristic of the WIMP signature. Future progress hinges on improving sensitivity by several orders of magnitude beyond existing or planned experiments before the background floor imposed by the coherent scattering of neutrinos from the atmosphere and from the sun will limit sensitivity to WIMPs. Detectors based on the noble liquids are particularly well suited to achieving the target masses (approximately 200 tonnes) necessary to pursue this challenge.

CLEAN (Cryogenic Low-Energy Astrophysics with Noble liquids) exploits a *single-phase* approach where only scintillation light is measured. Simplicity of design affords the construction of a very large detector at low cost, while pushing the major sources of background outside of the region of interest for a sensitive WIMP search. MiniCLEAN will serve as technical demonstration of the principles of a single-phase detector using 500 kg of liquid argon (LAr) as the target material and including three-dimensional event reconstruction. The MiniCLEAN detector is deployed in the Cube Hall, 6,800 feet underground, at SNOLAB.

This year, MiniCLEAN reached a major milestone and finished the cooldown to the required LAr temperatures. An extensive SNOLAB Operational Readiness Review, examining the state of the detector, documentation, safety, and the readiness of the collaboration for science running was successfully completed in February 2017. Operations permission was granted and the ongoing LAr fill started.

PNNL contributions to MiniCLEAN operations included a new role leading the production simulation and data processing. A test bench to assay the argon-39 content of a sample for use as a calibration source was designed and deployed. Staff members also led the analysis group, performing cold gas studies that demonstrated the world-leading purity of the CLEAN technology.

PNNL also established a role in the next-generation LAr dark matter detector, DarkSide-20k, focusing on developing general tools necessary to determine radioactivity limits required for detector construction materials. This included

PN16097/2874

a framework that consists of an assay organization and recording tool, drawing together capabilities across the entire collaboration and allowing efficient sharing of expertise. Additionally, in order to determine the required assay level, a tool to estimate the level of uranium and thorium permitted in materials due to neutron production in (α,n) reactions was created.

FY 2018 will see the execution of the full MiniCLEAN science and calibration program using an atmospheric LAr target, allowing technical demonstration and a dark matter search, followed by injection of argon-39 to test the limits of background rejection with this technology.

Deep Learning Applied to Accelerator Neutrino Physics in Liquid Argon Time Proportional Chambers

Eric D. Church

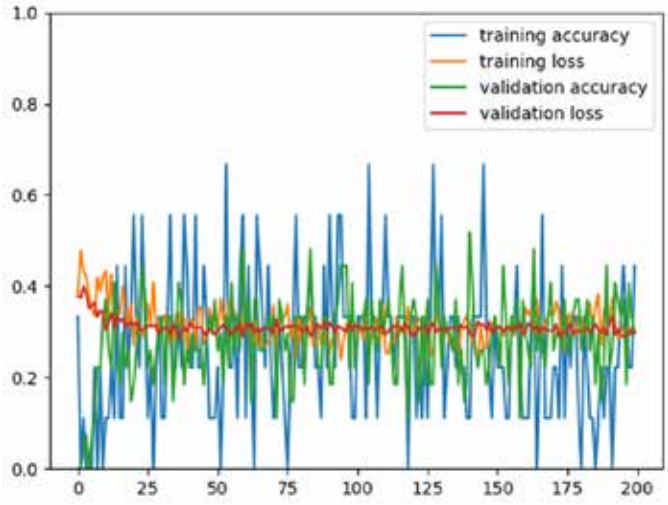
For this project, we will measure properties of neutrinos arriving from accelerators at Fermilab to detectors such as MicroBooNE (micro Booster Neutrino Experiment). PNNL will contribute computationally sophisticated deep learning (DL) techniques to lead these measurements.

Since beginning in February 2017, we have produced simulated and real data in the hdf5 file format to store MicroBooNE experimental datasets to which all scientists have been applying DL techniques. With these data, we have produced DL networks that are competitive with previous work from the MicroBooNE experiment. Further, we have established roles for PNNL researchers within the MicroBooNE and Deep Underground Neutrino Experiment (DUNE) collaborations. This ensures continued access to new data, engages us in understanding the data quality, gives us a role in detector development, and gives us leadership in the scientific output of the experiments.

Within this project, we are conducting research on three fronts: 1) developing networks that optimize use of the full, correlated information from the three planes of the MicroBooNE detector; 2) working with the keras/tensor-flow infrastructure and hdf5 files to allow for quicker scientific exploration and effective collaboration; and 3) bringing these techniques to nucleon decay search strategy development and cross-section measurements within the MicroBooNE and DUNE experiments.

We have begun work recently on a semantic segmentation network for MicroBooNE, in which each pixel in the images is labeled by the particle species that produced its charge deposition. This is novel for high energy physics and nucleon decay search, and it will allow even more precise identification of a neutrino or nucleon decay signature. Previous MicroBooNE work showed a 1:2 signal to background result in neutrino event identification using DL, which we have confidence, using our new analysis techniques and incorporating the cosmic ray detector data, will improve by 50%. Further, our innovative work maintains the full detector image fidelity, which will bring large background reduction. The impact here to suppress MicroBooNE's most pernicious background—that from irreducible cosmic rays—could not be overstated, if achieved.

The nucleon decay search with DL, which shows an early indication of increasing DUNE's sensitivity for this paradigm-changing discovery by an estimated 40%, would immediately make PNNL the leader in this exciting physics on DUNE. Analyses will be enabled and informed by the MicroBooNE DL work we have done already and will demonstrate an increase in signal efficiency and an increase in background rejection.



A typical loss and accuracy plot versus epoch. For us, an epoch is one batch, which is 10 events. Each 10 events is a new set, not seen in another epoch. This plot is typical of the state of the research at this time, in that after early training, the loss and accuracy stall, despite varying learning rate and other parameters.

Currently, we are seeing an approximately 2x-random classification accuracy for 5 particles of about 35%, but with a large, not-understood volatility from batch-to-batch. We expect to achieve a mid-80% accuracy. We believe a full examination of the underlying data integrity must occur next and will embark on this in FY 2018.

PN17064/2954

Image Fusion – Secondary Ion Mass Spectrometry (SIMS) and Microscopy

Benjamin E. Naes

Image fusion of SIMS and scanning electron microscopy (SEM; equipped with an energy-dispersive X-ray Spectroscopy [EDS]) is a powerful correlative imaging capability that enables the analysis of isotopic and elemental microstructures at spatial resolutions near those obtained via electron microscopy alone. The research in this project improves the ability to identify Uranium (U)-bearing particles under non-ideal sampling conditions for nuclear safeguards, which has helped with the analysis of Al-rich domains within lithium aluminate ($LiAlO_2$) ceramics used as tritium-producing burnable absorber rods (TPBARs), as well as Molybdenum (Mo)-rich domains within U-Mo alloys, developed for light water power reactors.

Image fusion is the process of combining input from two or more images into a single image. The resulting image is, by definition, more informative than any of the stand-alone images. The intent of image fusion, as applied to the micro-analysis of nuclear materials, is to capitalize on the power of combining two or more powerful analytical techniques to enable the observation of spatially unique domains with high precision isotopic specificity. The successful development of an image fusion capability involving SIMS and EDS has led to a novel approach for the analysis of isotopic and elemental spatial distributions, aiding analytical measurements ranging from the identification of actinide-containing particles to the irradiation behavior of micro-structured nuclear fuels.

As a side step to image fusion, and with some correlation to the fusion framework, a convolutional neural network was developed to post-process SIMS images and, ultimately, determine isotope ratio distributions of particles. This initial phase of utilizing artificial intelligence spawned other exploratory applications to assist with SIMS imaging, post and pre-acquisition.

In FY 2016, image fusion of SIMS and EDS X-ray maps for the isolation and reanalysis of isotopically unique U-bearing particles was studied and published. The work represented the first use of image fusion for nuclear safeguards applications, resulting in improved accuracy and precision of isotope ratio measurements for U-bearing particles that fall below the spatial resolution of SIMS images.

From this work, a machine learning approach was hypothesized to improve particle identification following the automated pre-screening SIMS imaging step; as such, a fully convolutional neural network (FCN) was developed to supersede automatic or manual thresholding approaches typically used for these large data sets. In this case, the FCN performed a binary classification of pixels in each image, belonging to a particle or not, which effectively removed background signal without manually or automatically determining an intensity threshold. Eight thousand images from 28 different particle screening analyses were used to train the FCN, with near 99% accuracy.

Once the FCN was implemented, background eliminated images were then segmented using a watershed technique in order to determine isotopic ratios of identified particles. A comparison of the isotopic distributions of an independent data set, segmented using the neural network with a commercially available automated particle measurement (APM) program developed by CAMECA, showed that an order of magnitude gain in absolute signal per particle was achieved by the FCN approach, combined with significantly less user bias and improved robustness.

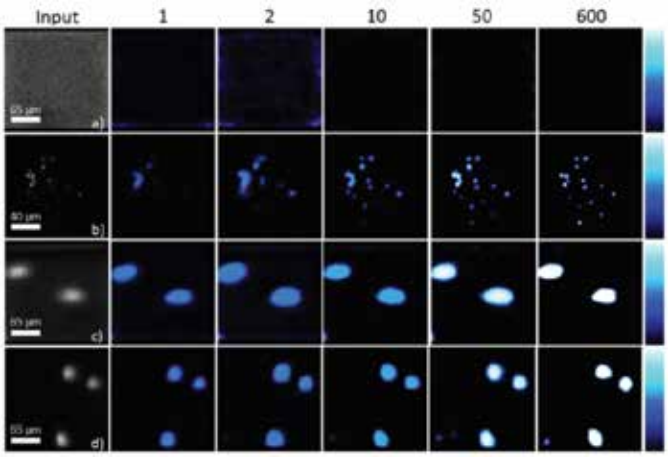
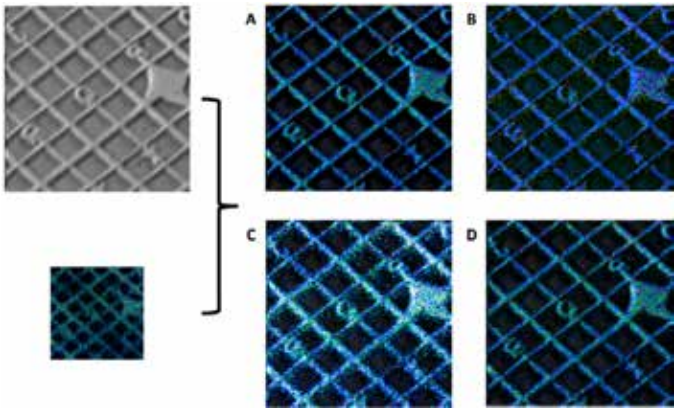


Image series for four field of views (FOVs) after various epochs during FCN training, highlighting the evolution of the network's ability to decipher foreground from background and identify particles in differing FOVs.

PN16025/2802

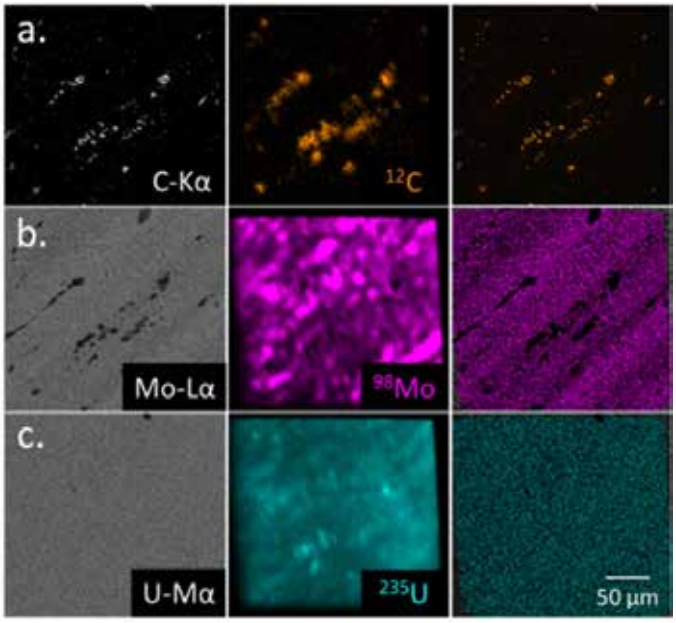
In FY 2017, a study was undertaken to compare different image fusion algorithms versus the hue saturation and lightness (HSL) method employed across the applications to date (including particle analysis, U-Mo monolithic fuel foils, and $LiAlO_2$ TPBAR components). Each of the methods involves scaling up and registering the low-resolution image to the high-resolution image, decomposing the images into intensity and detail images, histogram matching the high-resolution image to the panchromatic image intensity, and producing the final image by substituting or combining the intensity and detail images of the high- and low-resolution images. Several of these methods have been reproduced from the literature.

The method that works best is somewhat subjective (differing applications may require separate fusion methods, for instance), yet it was discovered that the HSL method, used to date, performed comparably or better than the others tested (principal component analysis [PCA], additive wavelet, and substitutive wavelet transforms) using a test case image pair.



Left: a high-resolution SEM image (top) vs. low-resolution SIMS image (bottom). Right: fused images produced using A) HSL color space transform, B) PCA fusion, C) additive wavelet transform fusion, and D) substitutive wavelet transform fusion.

SIMS, SEM-EDS, and image fusion were performed on a sample form the Advanced Test Reactor Full-Sized Plate In-Center Flux Trap Position 7 experiment. The image fusion results were characteristic of SIMS isotopic maps, yet on the spatial resolution of the EDS images. The boost in spatial resolution is expected to help understand the heterogeneity that persists due to processing selections. Future microanalytical characterization efforts on irradiated fuels may reveal that SIMS image and fusion with EDS analysis could provide a vital role in determining if elemental variations in the unirradiated material are associated with isotopic variations in the irradiated material and, ultimately, if those variations affect performance.



EDS (left), SIMS (center), and fused (right) images for a single FOV of the U-Mo monolithic fuel foil. Three elemental signals are shown: a) carbon, b) molybdenum, and c) uranium.

Low-Mass Dark Matter Backgrounds Research and Development

Raymond A. Bunker III

Dark matter comprises 85% of matter in the universe, but its nature is unknown. This project focuses on understanding backgrounds that limit searches for low-mass dark matter and on the development of reduction strategies; the ultimate goal is to design a future experiment to discover dark matter.

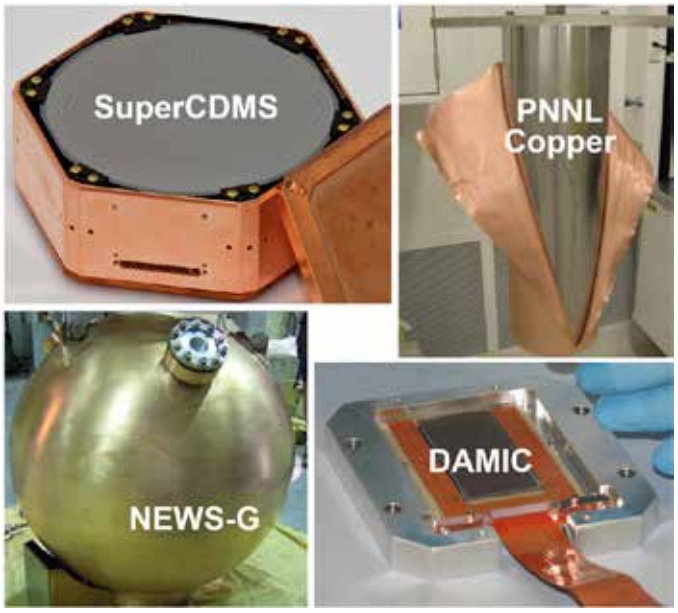
The nature of dark matter is a fundamental question in physics. There is overwhelming evidence that it exists, but its composition is unknown. Weakly Interacting Massive Particles (WIMPs) are a well-motivated class of particles that could make up the dark matter and could be detected by an Earth-based detector. Current WIMP searches have phenomenal sensitivity, but background requirements are extraordinary, and progress in the field is directly tied to the backgrounds achieved.

The hunt for WIMPs has pushed into the "low-mass" regime, in which solid-state and gaseous detectors are used to search for dark matter particles with masses less than 10x the mass of the proton. An important milestone is the detection of coherent scattering of solar neutrinos. This benchmark is referred to as the "solar neutrino floor." Although the next generation of experiments will explore a broad range of new parameter space, they will generally fall short of the neutrino floor. Trace levels of radioactivity in the detectors are expected to limit their sensitivities.

The goals of this project are two-fold: understand and characterize the backgrounds expected to limit the sensitivity of low-mass dark matter searches and develop strategies for the reduction of these backgrounds so that future experiments can reach the solar neutrino floor. Radioactive backgrounds from sources intrinsic to the detector materials are expected to dominate, particularly those created through exposure to cosmic-rays and to environmental radon. Research and development will focus on methods for measuring and mitigating these sources of background, utilizing PNNL's expertise in ultralow-background materials, low-background detectors, and ultrasensitive radiopurity assay.

In FY 2016, cosmic-ray simulations identified tritium and ^{32}Si as the dominant backgrounds for future detectors. A white paper was written outlining the potential use of pure ^{28}Si to mitigate the latter, and preliminary plans were made to explore removal of tritium from detectors by "baking" them at high temperature. Finally, progress was made on the development of advanced background modeling tools, which will be used to construct a background budget at the solar neutrino floor.

In FY 2017, a thorough exploration of the silicon industry was completed, including investigation of ^{32}Si assay methods and mitigation strategies. Results concluded that isotopically pure ^{28}Si is the most promising strategy for creating future ^{32}Si-mitigated detectors. Plans were formalized for a detector testing facility at PNNL that will be used to acquire data for validation of background models, a critical need to understand detector response and, thus, discovery potential. This facility will be constructed in FY 2018.



Dark matter detectors from (counterclockwise from top left) the Super Cryogenic Dark Matter Search (SuperCDMS), NEWS-G, and DAMIC experiments; (upper right) copper is an essential material for all three—PNNL's copper is the purest in the world and is critical to the future success of low-mass dark matter experiments.

PN16086/2863

A program of research and development was initiated to explore use of PNNL's ultra-pure copper for reduction of radon-related backgrounds. A test growth was conducted at the end of FY 2017 for the New Experiments With Spheres-Gas (NEWS-G) experiment; further tests are planned for NEWS-G and Dark Matter in charge coupled devices (DAMIC) in FY 2018. Finally, solid-state detector response was investigated from first principles, and code was developed for calculating ionization in silicon and germanium crystals.

In FY 2018, development of background modeling tools will continue, so as to provide guidance for the test facility and a generalized framework for all dark matter experiments. Plans for exploring tritium "baking" were revisited at the end of FY 2017, and measurements will begin in FY 2018. Tritium is a difficult background and must be mitigated in future detectors.

Mitigating Challenges Toward an Enduring Supply of Low-Radioactivity Argon for Ongoing Pacific Northwest National Laboratory National Security and Basic Science Programs

Henning O. Back

This project aims to understand and mitigate the challenges that were faced in the production of the first 200 kg of low-radioactivity underground argon for the DarkSide-50 dark matter search experiment and to identify the source or sources of the residual ^{39}Ar and ^{85}Kr. This will ensure the successful establishment of an enduring supply of low-radioactivity argon with the lowest ^{39}Ar levels.

While argon is easily purified to remove natural radioactivity, argon that is derived from the atmosphere contains a small fraction of the radioactive isotope ^{39}Ar. Due to the difficulty of separating argon isotopes, the ^{39}Ar is the limiting background in all very low-level radiation detectors that use argon, particularly argon-based dark matter detectors. The future successes of these programs depend on the establishment of an enduring source of low-radioactivity argon that is free, or nearly free, of ^{39}Ar.

Argon derived from the atmosphere contains the radioactive isotope ^{39}Ar, which is produced through interactions with cosmic rays. Argon from active CO_2 wells in southwestern Colorado have been found to contain very low levels of ^{39}Ar, with an upper limit of 0.07% atmospheric concentrations.

The first production of low-radioactivity underground argon (UAr) from the Colorado CO_2 wells was hampered by poisoning of sorbents and contamination freeze out in cryogenic systems. To ensure that these challenges are mitigated in the next generation of UAr production, we successfully identified the contaminations in the original adsorption and cryogenic systems, and through collaboration with the University of Houston, the contamination concentrations in the CO_2 feed were measured.

Examples of high mass compound contaminants found in UAr extraction and purification plants.

C_3H_8	C_6H_{14}	$C_6H_{12}O$	C_8H_{16}
$C_5H_{10}O$	$C_6H_{12}O$	C_7H_{16}	C_8H_{18}
C_5H_{12}	C_6H_{12}	C_6H_6	$C_6H_{10}O_2$
C_6H_{14}	C_7H_{14}	C_7H_{16}	C_8H_{18}
C_5H_{10}	$C_6H_{13}I$	$C_6H_{12}O$	C_9H_{20}
$C_5H_{10}O$	C_7H_{16}	$C_5H_8O_2$	

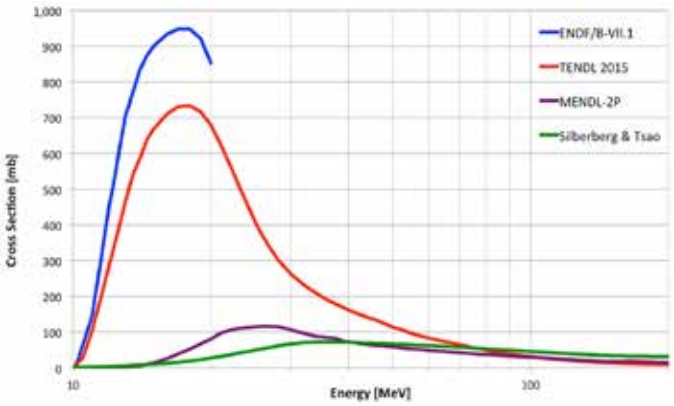
Through this project, PNNL has joined the DarkSide-20k dark matter search experiment and its effort to produce 50 tons of UAr. The project is called Urania, and it is a collaboration between Princeton University, the University of Houston, the University of Naples in Italy, PNNL, and Fort Lewis College in Durango, Colorado. The results of the contamination determination are being used by the Naples group in the design of the Urania plant.

To determine the origin of the ^{39}Ar and ^{85}Kr, we explored three possibilities for their origin: they are intrinsic, they are contaminations from an air infiltration, and/or the ^{39}Ar is produced by cosmic rays.

At PNNL we had found strong circumstantial evidence that the ^{39}Ar and ^{85}Kr in the first production of UAr for DarkSide-50 was from an air infiltration. In order to explore this hypothesis, we built a long-term gas analyzer that could monitor several gas streams at the remote location of the CO_2 wells. It was, indeed, determined that there is a minor air infiltration that can account for the amount of ^{39}Ar and ^{85}Kr in the original production of UAr.

PN16036/2813

In exploring the cosmogenic production of ^{39}Ar in the UAr, we found that the cross sections for the relevant reactions are only estimates and vary by an order of magnitude. By using the known history of the previous UAr production, we estimated that up to 20% the ^{39}Ar in DarkSide-50 could have come from cosmogenic production.



Calculated cross sections versus energy for ^{40}Ar(n,2n)^{39}Ar reactions.

Knowing the production rate of ^{39}Ar due to cosmic rays is critical in order to understand the proper handling of the UAr once it has reached Earth's surface and will determine the limits to the science reach of an argon-based dark matter experiment. In FY 2018, this project will be making this measurement by irradiating a small volume of UAr at the LANSCE accelerator at Los Alamos National Lab and counting the ^{39}Ar at PNNL.

PICO-40L (40 Liter Version of the PICO Physics Collaboration Experiment) Bubble Chamber Research and Development

Ben M. Loer

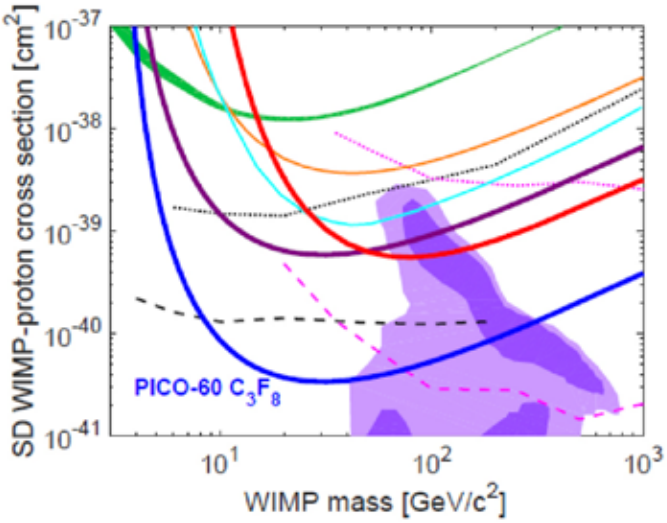
Dark matter makes up the majority of the mass in the universe, yet its fundamental nature remains a mystery. This project is focused on bubble chamber research and development, which is a compelling technology for dark matter detection due to its unique sensitivity to spin-dependent interactions with fluorine and near-complete insensitivity to the most common electromagnetic backgrounds.

The PICO collaboration has pioneered the development of bubble chambers for dark matter searches. The next-generation PICO-40L detector, currently under construction, features new technical innovations that will significantly reduce backgrounds compared to previous generations. Prior to this project, PNNL made technical contributions to PICO in the form of engineering design and material assay. Through this project, PNNL will now also contribute scientific leadership, driving the final installation of PICO-40L, developing improved detector models and analysis techniques, and preparing for the next-generation PICO-500 detector.

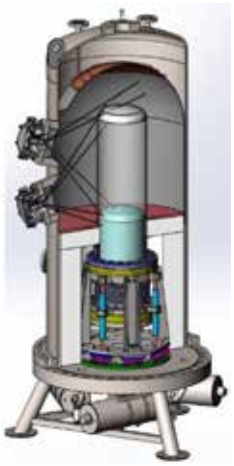
An early success of the project is a publication of data collected with the PICO-60 detector. Before running, the detector was subjected to a new cleaning procedure designed to remove particulate contamination, which is the basis for part of the PICO-40L design. This procedure successfully eliminated the particulate backgrounds, and the detector achieved a world-leading sensitivity to spin-dependent dark matter interactions, a factor of 17 increase over previous PICO results.

The technical design for PICO-40L is nearing completion and passed Conceptual Design Review at SNOLAB in April 2017. It is on track for Technical Design Review completion in October 2017.

During FY 2018, this project will switch from design to commissioning and analysis for PICO-40L to next-generation research and development. PICO-40L will be installed underground at SNOLAB in the first quarter of FY 2018, and this project will provide on-site assistance with installation and commissioning. Our team will focus on improving modeling and analysis of the detector acoustics, which are a critical component of background rejection, radioactive background modeling, and exploration of deep learning techniques for both acoustic and optical reconstruction. Finally, the collaboration will begin designing the next-generation PICO-500 detector next year, as well. We will continue to lead coordination of technical design and assay efforts, as well as take on leadership roles in other areas such as background control.



Spin-dependent dark matter interaction sensitivity from the PICO-60 experiment, compared to other experimental searches (lines) and theoretical predictions (shaded regions).



Design of the PICO-40L Detector.

PN17072/2962

Search for Lepton Number Violation

John L. Orrell

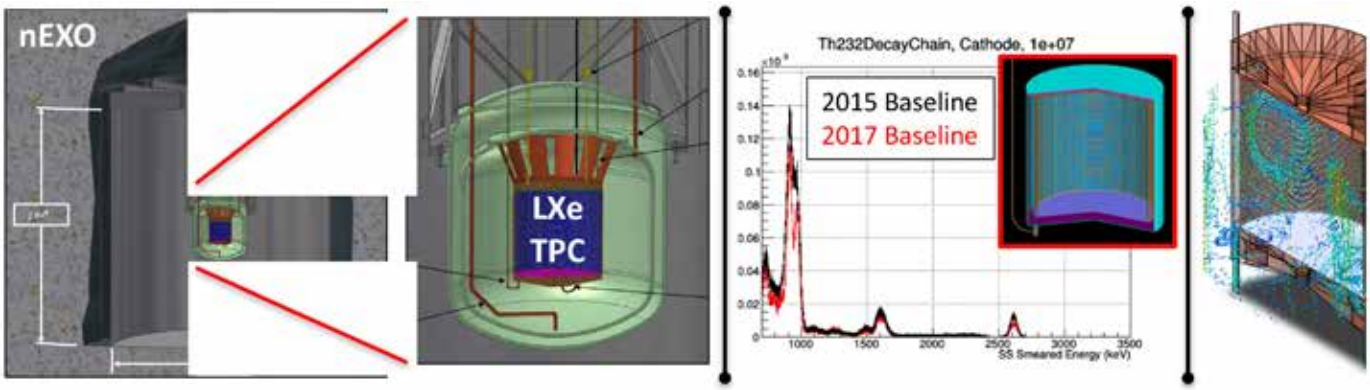
This project will assist development of a ton-scale neutrinoless double beta decay experiment employing 136Xe operated as a liquid xenon (LXe) time projection chamber (TPC) by working with the nEXO Collaboration. The focus of this project is on background estimation, material radio-impurity assay, low background materials development, and isotopic enrichment of xenon, which are areas of PNNL expertise.

PNNL is investing in the development of the next generation of neutrinoless double beta decay research, an experimental program recommended by the scientific community in the 2015 Nuclear Physics Long Range Plan to search for explicit lepton number violation and the nature of the neutrino. Lepton number—the quantum number identifying electrons, muons, taus, and their respective neutrinos—is conserved in all observed Standard Model particle interactions. However, due to their electrically neutral nature, neutrinos and their antimatter counterpart anti-neutrinos may be unique among the Standard Model particles, in that they are actually the same identical particle, proving lepton number cannot be a conserved quantity. Neutrinoless double beta decay—the double beta decay of an unstable nucleus without the emission of the otherwise requisite two anti-neutrinos—is potentially an experimental means to identify the matter-antimatter uniqueness of neutrinos.

The discovery of lepton number violation through double beta decay would open the door to a possible explanation of the matter-antimatter asymmetry of the universe: the fact that matter predominates and antimatter has only a fleeting existence in our world. The scale of future double beta decay experiments is driven by the need to observe a large number of candidate nuclei, given the anticipated half-life is on the order of 10^{28} years expected for this ultra-rare nuclear decay process.

The nEXO Collaboration is an international organization of research institutions developing a ton-scale, next-generation double beta decay experiment employing a target mass enriched in ^{136}Xe and operated as an LXe TPC. PNNL expertise in background estimation, material radio-impurity assay, low background materials development, and isotopic enrichment are significant contributions to the research and development program to demonstrate the feasibility of the nEXO experimental design.

In FY 2017, the PNNL team performed the simulation production of the "2017 Baseline" background model for the nEXO experiment's updated design. This radiation transport simulation required about 1.5 weeks of 500 computing cores to generate a complete set of background simulation outputs used by the collaboration to calculate the half-life sensitivity reach of the nEXO experiment. The projected background levels are determined based on the radio-impurity concentration levels measured in detector construction materials. PNNL analytic chemists used inductively coupled plasma mass spectroscopy to measure trace levels of radio-impurities in copper, sapphire, and silicon used for silicon photomultiplier detectors. These material radio-impurity measurements are key inputs to validating the background and sensitivity projections of the nEXO experimental concept.



The two panels on the left show the nEXO neutrinoless double beta decay experimental design concept. The two panels on the right show (from left to right) comparison of the 2017 baseline update radiation transport simulation, with the latest simulation geometry shown in the upper-right and at the right-most figure, a visualization of the LXe fluid flow model in the copper vessel of the nEXO experiment.

PN16079/2856

The nEXO experimental concept has many aspects of design. Two areas contributed to by PNNL staff members include conceptual design of the cathode and selection of materials to meet the radio-purity goals of the experiment and modeling and computational fluid dynamics modeling of the LXe flow in the vessel. Understanding the LXe fluid flow helps predict the turn-over time of the LXe in the vessel, necessary to control trace-level chemical impurities that potentially negatively impact the ability for the TPC to accurately reconstruct energy deposition event locations in the detector.

In FY 2018, PNNL will continue to work with the nEXO Collaboration to refine this ton-scale experimental concept focused on searching for lepton number violation through discovery of the double beta decay process.

Ultrasensitive Nanoscale Chemical Imaging with Controllably Tailored Electromagnetic Waves

Patrick Z. El-Khoury

In this project, we will design, fabricate, and test the performance of novel constructs that can be used to controllably guide light on the nanoscale. We will then use the nascent nanoscale electromagnetic fields to identify and image molecules on the same length scale through Raman and UV-Vis nano-spectroscopy, with joint sub-5-cm^{-1} spectral and sub-5-nm spatial resolution that has not been previously demonstrated.

Metallic nanostructures can support confined and enhanced electromagnetic fields, which may be used to chemically identify and image molecules (e.g., through surface-enhanced Raman scattering [SERS]) and tip-enhanced Raman scattering (TERS). However, the attainable spatial resolution and sensitivity in near-field optical microscopy and spectroscopy come at a price; the recorded molecular signatures are obfuscated by the complex interplay between molecules and ill-defined plasmonic metal nanostructures. Slight nanometric variations in the composition and structures of the plasmonic nanostructures are difficult to control and may completely alter the plasmonic response and structure of the local electric fields interacting with nearby molecules. This is particularly the case in emerging techniques, such as TERS, which hold great promise for nanoscale chemical and biological imaging applications. In this regard, the commonly adopted practice of correlating structural/topographic images of plasmonic nanostructures with near-field optical spectra is often insufficient; there is a need for novel approaches that can be used to controllably tailor electromagnetic fields on a few-nanometer-length scale.

Attaining high spatio-spectral resolution using conventionally macro-microscopic techniques necessitates the use of plasmonic antennae. To controllably couple, interfere, guide, and focus microscopic light sources onto the nanoscopic apex of a probe, we will employ helium ion lithography to etch nanostructures onto the metallic probes. The performance of these tips will then be tested using a state-of-the-art combined AFM-optical microscopy platform, which will be used to perform both nano-Raman and nano-UV-Vis imaging measurements.

In FY 2017, we established a protocol that can be used to simulate the optical properties of lithographically patterned metal-coated probes.

In the simulations, the "contrast" obtained by adding the lithographic pattern design to the metallic tip was evident. Under both incident laser polarizations, we observe a 5X enhancement in the electric field magnitude for the patterned design when compared to the reference construct. Although a modest increase in signal magnitude is observed in this case, this result clearly demonstrates how a simple lithographic pattern may be used to enhance the optical signals from (bio)molecules coaxed into plasmonic tip-surface nanojunctions.

Modifying the lithographic pattern design and repeating the simulations results in a much higher contrast. In this case, a 15X enhancement in the electric field magnitude is observed. Other nanostructure designs are envisioned and will be tested out in FY 2018.

Our next step was to test this premise in the laboratory. In FY 2017, we produced such probes using helium ion lithography (HIL).

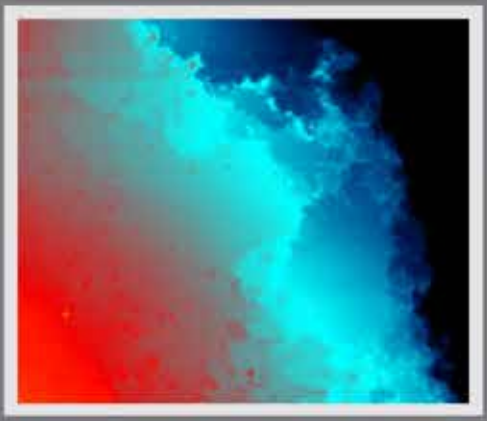
Several lithographically etched probes were produced in FY 2017. We were successful in etching trenches of various widths. On the basis of our simulations, we anticipate at least one of the tips to yield a 15X enhancement in the electric field magnitude and, hence, 50,000 times enhancement in Raman scattering cross sections in nanoscopy mode.

The optical properties of our tip of choice for this report were measured using hyperspectral optical absorption microscopy. A UV-vis absorption image (hyperspectral image slice) of the tip illustrates the intensity distribution at any particular wavelength. Similarly, each pixel contains the full spectral response of the tip in the near-UV-near-infrared spectral region. Sharp resonances at approximately 420 nm were observed from our tips, commensurate with the expected plasmonic response.

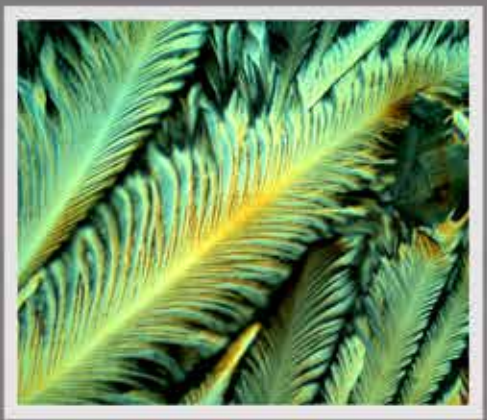
In FY 2018, we will essentially repeat the HIL experiments to tailor contaminant-free tips and substrates. This will allow us to test the performance of our novel tip designs in nanoscopy mode. To date, we have identified several different promising tip designs through finite difference time domain simulations; they will all be fabricated and tested in TERS mode.

PN17084/2974

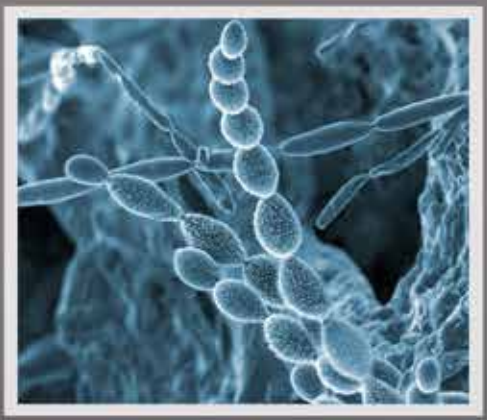
ON THE COVER



Soils are vital to our future supply of water and food, but they also play a critical role in adapting to climate change and sustaining the planet's biosphere—the land, sea and atmosphere occupied by living things. With funding from the Laboratory Directed Research and Development program, PNNL researchers studied microbes from the typical rhizosphere found in the Northwest's Central Cascades pine forests to better understand how microbes help stabilize carbon in soils. The rhizosphere is the small area surrounding plant roots where microbes interact with minerals in the soil—a hot spot for biogeochemical activity that is constantly changing.



PNNL researchers and collaborators are studying uranyl nitrate—a key ingredient in the nuclear fuel cycle—to better understand how its properties change as it loses moisture in different climates or storage scenarios. Using time-resolved reflectance spectroscopy, they have captured some of the dynamics of this transformation, watching the transformation process to examine the material's structure and behavior as it dehydrates. With support from DOE's National Nuclear Security Administration, researchers are now correlating their real-time, high-resolution observations with computational studies to help understand environmental effects on several components of the fuel cycle.



In this false-color topographic map of the surface of a piece of polished glass, the most corroded areas (dark blue) are contrasted with the areas most protected (red). PNNL researchers are exposing glass samples to different chemical solutions and studying these topographic images to better understand and predict how glass would corrode in various conditions for thousands of years—helping fine-tune new glass corrosion models to ensure vitrified nuclear waste is confined in the glass for safe, long-term storage and disposal. This research is funded by a joint DOE Office of Environmental Management and Office of Nuclear Energy project and by Washington River Protection Solutions.

2017 Annual Report

Laboratory Directed Research & Development

www.pnnl.gov

U.S. DEPARTMENT OF ENERGY

BATTELLE

PNNL-27379