Pacific Northwest
NATIONAL LABORATORY

PNNL-35165

Integrated Assessment Modeling of Grid Resilience: GCAM-to-PCM Scoping

Methodology Report

October 2023

Allison M Campbell
Malini Ghosal
Nathalie Voisin
Sarah H Davis
Dhruv Bhatnagar
Gokul C Iyer
Haewon C McJeon
Patrick R Maloney
Abhishek Somani
Michael Kintner-Meyer

U.S. DEPARTMENT OF ENERGY

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

DISCLAIMER

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
www.osti.gov
ph: (865) 576-8401
fox: (865) 576-5728
email: reports@osti.gov

Available to the public from the National Technical Information Service
5301 Shawnee Rd., Alexandria, VA 22312
ph: (800) 553-NTIS (6847)
or (703) 605-6000
email: info@ntis.gov
Online ordering: http://www.ntis.gov

PNNL-35165

Integrated Assessment Modeling of Grid Resilience: GCAM-to-PCM Scoping

Methodology Report

October 2023

Allison M Campbell
Malini Ghosal
Nathalie Voisin
Sarah H Davis
Dhruv Bhatnagar
Gokul C Iyer
Haewon C McJeon
Patrick R Maloney
Abhishek Somani
Michael Kintner-Meyer

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

Pacific Northwest National Laboratory
Richland, Washington 99354

Acknowledgments

This research was supported by the Earth & Biological Sciences Investment, under the Laboratory Directed Research and Development (LDRD) Program at Pacific Northwest National Laboratory (PNNL). PNNL is a multi-program national laboratory operated for the U.S. Department of Energy (DOE) by Battelle Memorial Institute under Contract No. DE-AC05-76RL01830.

Contents

Acknowledgments ii
1.0 Introduction 1
2.0 Background 2
3.0 Summary of Findings 3
4.0 Project Task Breakdown & Project Schedule 4
5.0 Downscaling Framework 6
6.0 GCAM 7
6.1 Tool Description 7
6.2 Example Scenario Results 8
6.3 LDRD Scenario Configurations in GCAM 10
6.4 LDRD Cost and Effort to Run and Postprocess GCAM 10
7.0 GridView 11
7.1 Tool Description 11
7.2 Example Simulation Results 12
7.3 LDRD Simulation Configurations 14
7.4 LDRD Cost & Effort to Run and Postprocess GridView 15
8.0 Plans to Leverage & Enhance IM3 Tools for Downscaling 16
8.1 Capacity Expansion Regional Feasibility (CERF) 16
8.2 Total Electricity Loads (TELL) 18
8.3 Grid Optimization model (GO) 20
8.4 MOSART-WM 21
8.5 LDRD Downscaling Gaps that Remain 23
8.5.1 Modeling Load Shape Impacts of Technology Adoption 23
8.5.2 Modeling Hourly Renewable Output that Correlates with Meteorology 27
9.0 Value by End-User 34
10.0 Long-Term Vision 35
11.0 Conclusion 36
12.0 References 37

Figures

Figure 5.1. Proposed Simulation Framework 6
Figure 6.1. GCAM-USA produces energy projections at the state level 7
Figure 6.2 Sample GCAM simulation results for US footprint - Installed Generation Capacity with respect to three different GCAM scenario configurations 9

Figure 6.3. Sample GCAM simulations results for the US footprint – Transportation load by Vehicle Type (Left) and Fuel Type (Right) with respect to three different GCAM scenario configurations 9
Figure 7.1 GridView is a production cost model that evaluates security constrained unit commitment and economic dispatch under a number of constraints (ABB 2023) 11
Figure 7.2. Sample GridView results reflecting the average daily generation dispatch in the Northwest Power Pool region............ 13
Figure 7.3. Sample GridView results reflecting nodal marginal prices during peak load hour in summer (Left) and winter (Right) 13
Figure 7.4. Sample GridView results comparing large-scale energy storage facility net revenues to installed capacity, by state, under different GridView market conditions 14
Figure 7.5. GridView input parameters for load, wind and solar will be adjusted for simulations under this project............ 15
Figure 8.1. Framework for IM3 electricity system modeling to project grid stress 16
Figure 8.2. Summary of CERF functionality 17
Figure 8.3. Example hourly load predictions for California and Rhode Island scaled to match GCAM energy output............ 19
Figure 8.4. Sample of simplified topology to be used in GO simulations in IM3............ 21
Figure 8.5. Preliminary analysis of GO runtimes without high performance computing 21
Figure 8.6. Diagram showing the inputs and outputs of MOSART-WM (Source: Voisin et al. 2013) 23
Figure 8.7. LDV load shape generation using EVI-Pro............ 25
Figure 8.8. MDV load shape generation using PNNL methodologies 26
Figure 8.9. HDV charging station assumptions used in PNNL agent-based model 26
Figure 8.10. Three different methodologies and their increasing complexity that could be developed under this project to create hourly renewable output............ 27
Figure 8.11. Renewable hourly timeseries element in the larger downscaling framework to be developed............ 28
Figure 8.12. “NREL Datasets” methodology for generating renewable hourly timeseries............ 28
Figure 8.13. “Generating production from WRF” methodology for generating renewable hourly timeseries 29
Figure 8.14. Wind toolkit power curves reprinted from https://www.nrel.gov/docs/fy14osti/61714.pdf............ 30
Figure 8.15. “GP + GBM” methodology for generating renewable hourly timeseries 31
Figure 8.16. Sliding window cross-validation for GBM 32
Figure 10.1. Subset of the larger framework in Figure 5.1 illustrating future capabilities that are enabled after accomplishing successful downscaling and production cost modeling simulations of decarbonization pathways............ 35

Tables

Table 4-1. Proposed Project Schedule 5

1.0 Introduction

Deep decarbonization towards a net-zero emissions future and the associated growth in power demand, renewables, and changes in water availability will require significant evolution in grid planning. Many obstacles threaten the viability of near-term generation capacity build out, including the need for large-scale transmission expansion. Current very-long term planning tools do not adequately model physical representations of economic dispatch and system topology down to inter- and intra-regional transmission and distribution infrastructure. This project attempts to fill that gap.

An EED directorate objective is end-use decarbonization and the joint EED-EBSD directorate objective is to develop and analyze decarbonization pathways that will meet emission reduction goals, resiliency, and minimize impact to human and natural systems. While EBSD has invested in downscaling several variables from the Global Change Analysis Model (GCAM) (e.g., load, water demand, land use, etc.) and those tools are available to leverage, there is presently no downscaling of technology implementation (i.e., load, controls, and transportation) which are key to evaluate decarbonization strategies. Therefore, the objective of this project is to enhance multi-sectoral downscaling approaches which would enable assessment of grid resilience and resource adequacy under deep decarbonization more accurately, beyond the existing 10-year planning horizon used in system planning, through integrated power sector modeling that can inform future grid design and investment decisions.

This project will link an integrated assessment model that represents dynamic interactions across energy, economy, environment, technology adoption, and water (i.e. GCAM) with an industry-facing production cost model (e.g. GridView, Hitachi 2023). This will enable more accurate assessments of grid resilience and resource adequacy under a range of clean energy transition scenarios. The downscaling of GCAM output will be done for all US states, however, a proof-of-concept production cost model simulation(s) will be demonstrated on the 14 states in the Western Interconnect.

The downscaling tool will be developed generically such that the downscaling approach can be leveraged with other production cost model tools (beyond GridView) and future evolutions of dispatch models that may move away from marginal cost dispatch and utilize new unit commitment dispatch models as renewables deployment continues to increase. Efforts under this funding will result in an initial code-set (written in R or Python) that will receive inputs from GCAM and output datasets that can be directly fed into a production cost model to perform a full year hourly simulation for a future decarbonization scenario. Future funding to enhance, tune, and use such a tool in larger resiliency assessments will be required, as this effort is just the first step in a larger long-term vision.

This methodology report captures efforts that were accomplished with kick-start funding received from EBSD that spanned over two months. The purpose of this report is to better define the scope of work required for future funding.

2.0 Background

This team was provided $40K kickstart funding from the EBSD Mission Seed at the beginning of September 2021, two months prior to fiscal year end. The intent behind this funding was to scope the work required for such endeavor. This methodology report is an outcome of this effort, and documents what couplings were explored, and how existing tools can be leveraged and enhanced to accomplish such a framework.

This methodology report was prepared independently of the Agile LDRD that was being drafted concurrent to this effort. Project task breakdown and project schedules may not be synchronized with the Agile LDRD effort and may need to be better coordinated if the Agile receives funding.

It is important to note that the distinguishing feature between this project and ongoing efforts in the Integrated Multisector Multiscale Modeling project (IM3) is that this project describes a model coupling approach to focus on the impacts of decarbonization policies, whereas IM3 is a foundational science project and does not analyze policy (IM3 2023).

Another distinguishing feature of this effort will be the lessons learned with the limitations found in the production cost modeling tool (PCM) to be used in the study, called GridView, with significant changes in load and generation mix and how this might shape PCMs of the future.

3.0 Summary of Findings

Under this project funding, various tools proposed to be used to establish GCAM to GridView capability were explored. Upon exploration of IM3 tool capabilities, several enhancements and new capabilities were identified as needed to improve load and generation downscaling required for GridView modeling input.

Load Downscaling:
Two major load downscaling components were identified as minimum viable product (MVP) capabilities that could be accomplished under additional LDRD funding. The first is incorporating better Electric Vehicle (EV) charging characteristics in future projected hourly load profile construction. The second is incorporating increased penetration of Distributed Energy Resources (DER) impact to net loads. Ideally, additional downscaling techniques for other technology adoptions (power-to-liquid technologies, etc.) for industrial, residential, and commercial electric sectors should be considered as well. However, due to limited funding and GCAM capability, the two MVP load downscaling capabilities identified are perceived to be achievable within FY22. Methodologies that have been preliminarily considered for these load downscaling components are documented in 8.5.1.

Generation Downscaling:
A major generation downscaling component identified as an MVP capability is improving wind and solar hourly timeseries construction. Traditionally, historical based wind and solar timeseries datasets from NREL have been utilized to generate fictitious site-specific hourly timeseries data. However, these historical datasets will not correlate with the weather-dependent hourly electric load generated from TELL. Additionally, as climate changes, the performance and generation output characteristics of wind and solar plants are expected to change. We propose to begin exploration of creating improved hourly timeseries datasets for wind and solar under this funding. Methodologies that have been preliminarily considered for improved renewable forecasting timeseries generation are documented in 8.5.2

GridView Simulations:
Lastly, we recognize there will be significant time and funds allocated to performing and analyzing GridView simulations. The WECC GridView cases will be pushed to limits never previously tested, with significant amounts of new generating renewable resources, placed in locations that may or may not have sufficient transmission transfer capability. In addition, for compatibility with the power plant siting model CERF, locational marginal prices (LMPs) from GridView simulations will need to be acquired in 5-year timesteps to align with capacity expansion from GCAM. Therefore, depending on the year we wish to project our simulations out to (2050 or beyond), each use-case simulated will require at least five GridView simulations (2030, 2035, 2040, 2045, 2050, etc.). Post-processing and analyzing the output of each of these scenarios can be time consuming and will need to be approached strategically.

4.0 Project Task Breakdown & Project Schedule

The key findings under this kick-start funding have influenced the FY22 project task breakdown. The following tasks have been defined to accomplish the scope of work defined in this methodology report:

Task 0: Develop methodology report for better defining scope & methods (this report)

Task 1: Generate **GCAM scenario**[1] outputs for decarbonization scenario(s) of interest

Task 2: Build MVP **load downscaling** enhancements

- Incorporate transportation charging load profile impacts
- Incorporate DER net load impact (if possible with funding)

Task 3: Build MVP **generator downscaling** enhancements

- Create renewable timeseries that correlate with weather predicted hourly load

Task 4: Convert downscaling output to appropriate GridView input formats & **run GridView simulations for use-case(s)**[2] of interest

Task 5: Analyze GridView results and **derive conclusions/insights** regarding success of framework (post-process PCM results, perform use-case comparison, evaluate transmission impact, identify simulation limitations, compare with IM3 GO simulations, etc.)

The below table depicts the proposed project schedule and approximations for the amount of time it will take to complete the various project tasks. The project schedule also incorporates sector lead presentations and stakeholder outreach activities. The intent is to complete an initial demonstration of the downscaling and production cost modeling simulations by the summer of 2022.

[1] **Scenario:** A GCAM output scenario that incorporates technology adoption, climate, policy, decarbonization, etc. input configurations

[2] **Use-Case:** Downscaling assumption conditions, such as EV charging, DER flexibility, extreme event, etc. assumptions that are used as inputs to downscaling functions. Ideally the use-cases selected under this project should represent “book-end” conditions, i.e. unmanaged charging vs. optimal managed charging

Table 4-1. Proposed Project Schedule

PROJECT SCHEDULE

		FY21			FY22												
	Task/Milestone	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct
1	**Task 0:** Better Define Scope & Methods	X	X	X													
2	**Stakeholder Outreach #1** (Present Project)					X											
3	Present Updates to Sector Leads				X												
4	**Task 1:** Prepare GCAM Decarbonization Scenarios Harmonized with WECC				X	X	X										
5	**Task 2/Task 3:** Script & Test New Downscaling Functions				X	X	X	X	X								
6	**Stakeholder Outreach #2** (Present Progress & Study Scenarios)									X							
7	**Task 4/Task 5:** Demonstrate High-Profile Scenarios with WECC PCM Simulations							X	X	X	X	X	X				
8	**Stakeholder Outreach #3** (Present Demo Results)													X			
9	Prepare DOE Briefing Material & Scientific Journal Publication													X	X	X	X
10	Present Updates to Sector Leads															X	
11	Document Additional Development Needs															X	X

Intend to complete demonstration by mid-summer

5.0 Downscaling Framework

After interviewing IM3 tool development teams to better understand the functionality and level of downscaling they are accomplishing, an updated simulation framework to downscale GCAM output to the level needed for industry-grade production cost models (ie WECC's GridView model) is illustrated below. The output of GCAM is at the annual and state (U.S. state) level, and the input requirements for GridView are hourly at the plant/node level (generation) and balancing authority (BA) level (load). The blue shaded regions encompass the downscaling efforts for load (top) and generation (bottom).

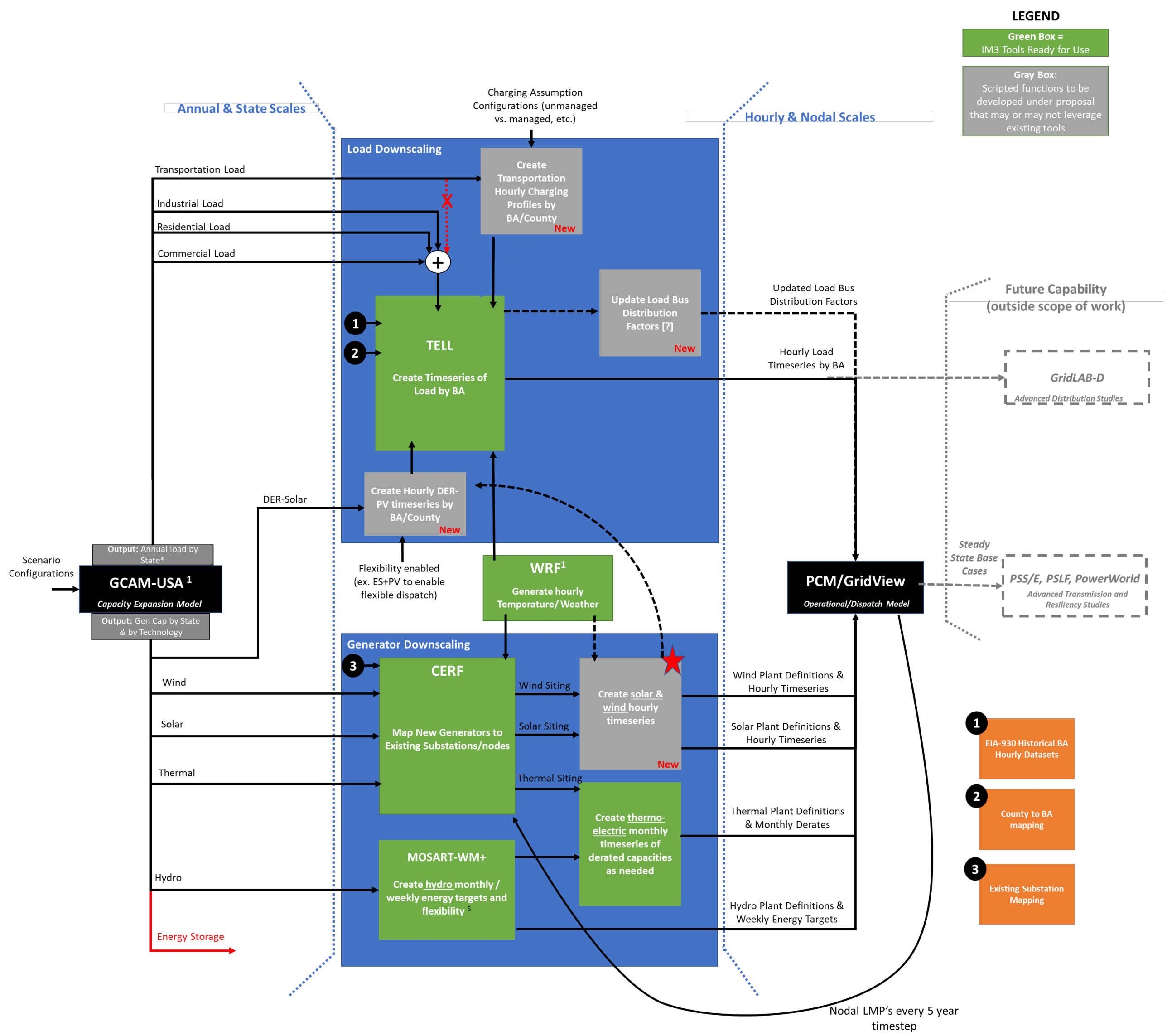

Figure 5.1. Proposed Simulation Framework

6.0 GCAM

GCAM-USA is a state-level model of the U.S. energy system embedded within a global integrated assessment model that can explore pathways for the evolution and decarbonization of electric power sector. GCAM is an open-source tool that offers unique and one-of-a-kind multi-sectoral modeling, containing representations of energy, economy, agriculture, land-use, and water systems. In its U.S. state-level representations, GCAM contains detailed modeling of nation-level economic features, socioeconomics, energy transformation, carbon storage, renewable resources, electricity markets, and consumer end-use energy demands, as shown in Figure 6.1.

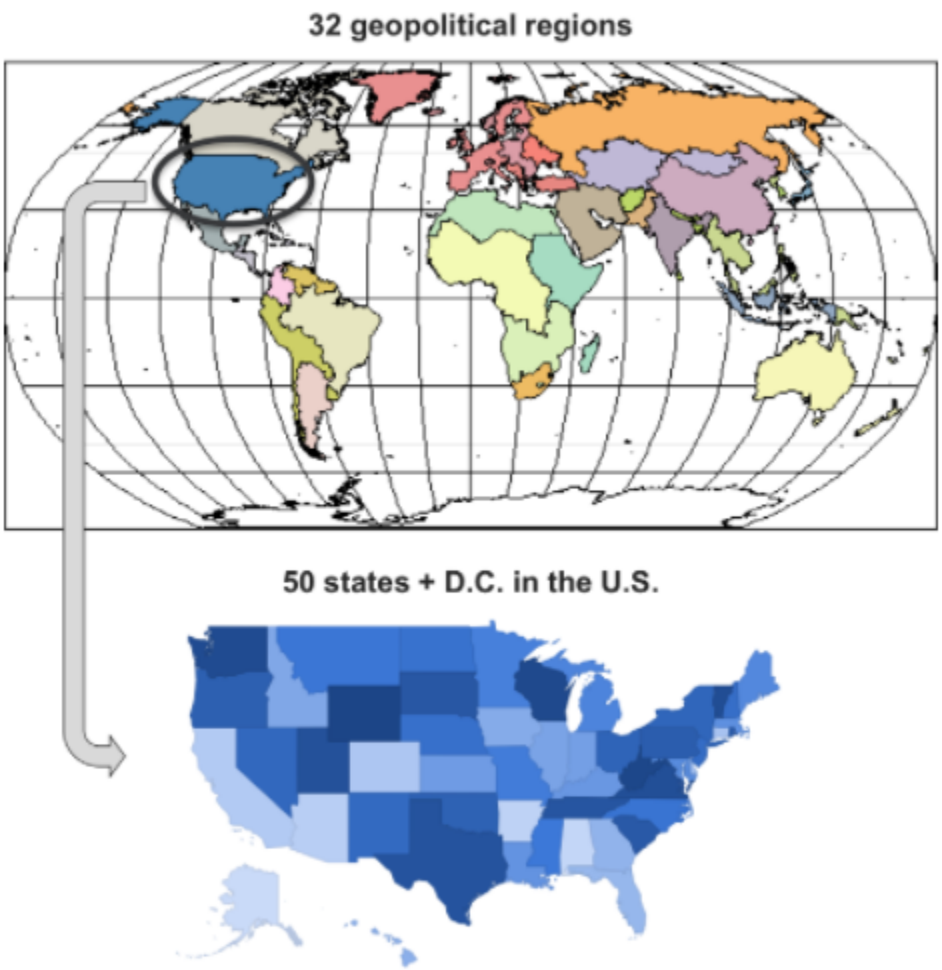

Figure 6.1. GCAM-USA produces energy projections at the state level

6.1 Tool Description

The following description of GCAM-USA was extracted from a journal article published in the Renewable and Sustainable Energy Reviews titled “Improving consistency among models of overlapping scope in multi-sector studies: The case of electricity capacity expansion scenarios”:

The GCAM-USA represents the supply and demand for electricity at the state-level for the United States. Electricity is supplied by several technologies that are tracked by vintage, with new investments that are shared out across the various technology options based on relative costs according to a non-linear logit-choice formulation. The endogenous demand for electricity in each end-use sector is deter-mined by a number of factors including population and economic growth along with the relative costs of electricity relative to other fuels. Electricity trade is represented between fifteen grid regions, with free trade between states within each grid region.
The cost of individual technologies in the power sector includes amortized capital costs, fuel costs, and operations and maintenance (O&M) costs. While capital costs and O&M costs are based on exogenous assumptions (see Iyer 2017a, Iyer 2017b for detailed assumptions), fuel costs are calculated endogenously based on supply curves. Fossil fuel resource supply curves are represented at the national level and are based on Rogner (1997). The uranium supply curve is based on Schneider (2008) and IAEA (2011). GCAM-USA assumes global trade in coal, gas, oil and uranium. GCAM-USA also includes state-level representations of wind, and geothermal resource costs based on resource curves from Zhou (2012) and Lopez (2012) respectively. Utility scale solar technologies are assumed to have constant marginal costs regardless of deployment levels.
The power sector module of GCAM-USA is embedded within a larger, multi-region, multi-sector framework. GCAM-USA includes representations of the rest of the energy system (e.g. natural gas resource pro-duction, refining), and economy (comprising of assumptions about population and GDP) for the 50 states and D.C. In addition, the model includes representations of the energy, economy, water, agriculture, and land-use systems for 31 geopolitical regions outside of the U.S. The dynamic-recursive, partial equilibrium model solves for the equilibrium prices and quantities of various energy, agricultural and greenhouse gas (GHG) markets in each 5-year time step and region through 2100 (Iyer et al. 2019).

6.2 Example Scenario Results

High level sample results from previous GCAM simulations were explored under this kickstart funding to better understand the level of detail GCAM can provide. Sample results of these already developed scenarios are shown in Figure 6.2 and Figure 6.3 below.

Generation detail includes the amount of installed capacity and annual energy by technology type by state. The generation technologies GCAM considers are hydro, wind, solar, biomass, coal, gas (natural gas dominant), nuclear, and oil.

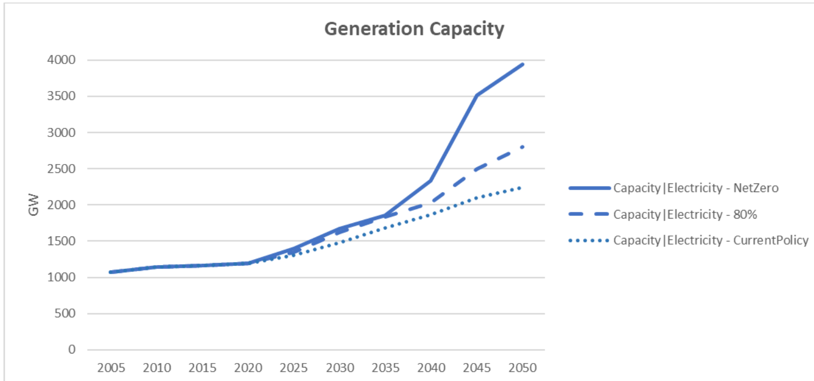

Figure 6.2 Sample GCAM simulation results for US footprint - Installed Generation Capacity with respect to three different GCAM scenario configurations

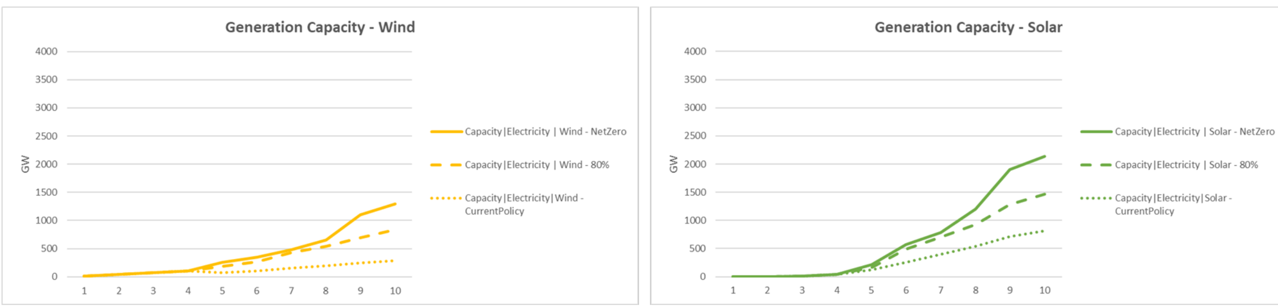

Figure 6.3. Sample GCAM simulation results for US footprint – Installed Wind (Left) and Solar (Right) capacity with respect to three different GCAM scenario configurations

With respect to GCAM electric demand output, it can be broken down to residential, transportation, industrial and commercial. The energy projections in each of these categories can be further downscaled to specific details on the type of technology/load that is contributing to the total demand.

An example of some of the level of detail GCAM transportation load can be broken down into is shown in Figure 6.3 below. Under the scenarios shown in these figures, there is still significant dependance on liquid (gasoline) fuels by the end of 2050. Under this project, we intend to explore more aggressive transportation electrification, such that more of the energy from vehicles is supplied by electricity by the end of 2050.

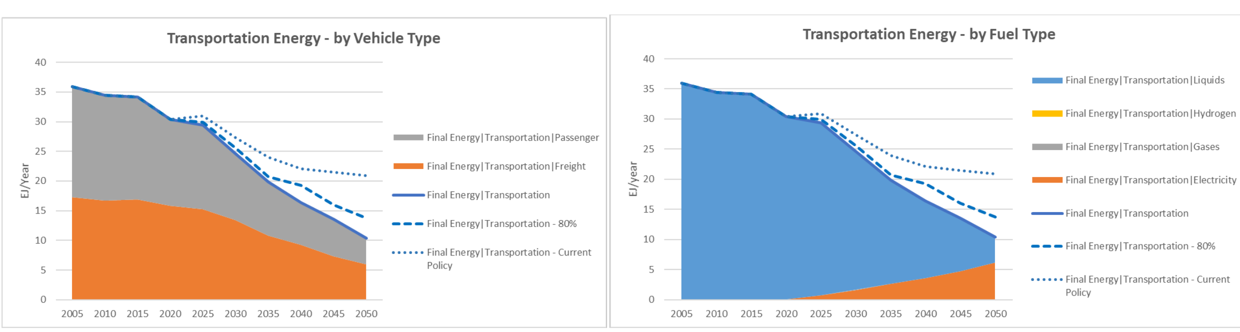

Figure 6.3. Sample GCAM simulations results for the US footprint – Transportation load by Vehicle Type (Left) and Fuel Type (Right) with respect to three different GCAM scenario configurations

6.3 LDRD Scenario Configurations in GCAM

There are various configuration knobs available to create new GCAM scenarios. Some of these configuration parameters take significantly more time to create than others.

GCAM-USA configuration options include:

- Population
- Gross Domestic Product (GDP)
- Technology costs over time
- Policy drivers (land protection, farmer subsidies, emission constraints, etc.)
- Resources (water, coal, etc. availability, modeled as supply curves)
- Parameters to equations within GCAM

We will not dramatically change most of these configuration options compared to previous GCAM simulations performed. However, under this project we will want to influence configuration options that will impact electricity supply and demand futures that influence:

- Transportation electrification
- Building electrification
- Industry electrification
- Advanced solar (after economies of scale produce “cheaper solar”)
- Advanced wind (as above, yielding “cheaper wind”)

Customizing the first GCAM scenario takes the most time, however adjusting and tweaking those scenarios becomes simpler after the first. We intend to utilize scenarios being developed under the EMF37[1] (Energy Modeling Forum) scenarios that are currently in development outside of PNNL, as a starting point for our new scenarios that intend to analyze decarbonization pathways.

6.4 LDRD Cost and Effort to Run and Postprocess GCAM

We intend to work in parallel with other GCAM-USA activities lead by Gokul Iyer, specifically the EMF37 scenario development and the Global Change Intersectoral Modeling System (GCIMS, 2023). The scenarios developed under these efforts can be used as a starting point for the decarbonization pathways we create under this LDRD. Therefore, the cost to develop new GCAM scenarios is expected to be light, as much of the heavy lifting will be completed under GCIMS.

Running GCAM-USA simulations, once all input configurations are complete, can be performed relatively quickly (within a day).

Postprocessing GCAM should be straightforward. The output of state-level GCAM-USA results can be exported in .csv and excel formats for which we can easily create scripts to post-process and extract the data needed for downscaling.

[1] Details on the Energy Modeling Forum scenarios can be found here: https://emf.stanford.edu/

7.0 GridView

GridView is a production cost model (PCM) that will be used in this GCAM-USA downscaling demonstration. PNNL has access to the most recent WECC GridView base cases of nodal network topology, BA loads, and generator inventories that will be adjusted as needed to reflect new futures in electric demand and generation mix based on GCAM-USA downscaling. We do not initially intend to alter any transmission topologies reflected in WECC's GridView case.

7.1 Tool Description

GridView is a commercial software developed by Hitachi that calculates chronological security-constrained unit commitment (SCUC) and security-constrained economic dispatch (SCED) to minimize the power system's operating costs of meeting electricity demand and reserve requirements while respecting system-level and unit-level operation limits (Hitachi 2023). Operating costs account for generating units' variable costs that largely consist of the fuel costs and start-up/down costs. In GridView, units are dispatched according to their variable cost, subject to minimum up/down times and ramp rate constraints until demand is met (usually on an hourly basis) in each BA. Transmission line constraints, emission constraints, and hurdle rates are also reflected in the dispatch.

GridView integrates engineering and economic analysis of the electric power grid to simulate SCUC and SCED in large-scale transmission networks. It is a tool that is widely used to study the utilization of generators and transmission lines, production cost of generation, locational marginal pricing (LMP), and transmission congestion.

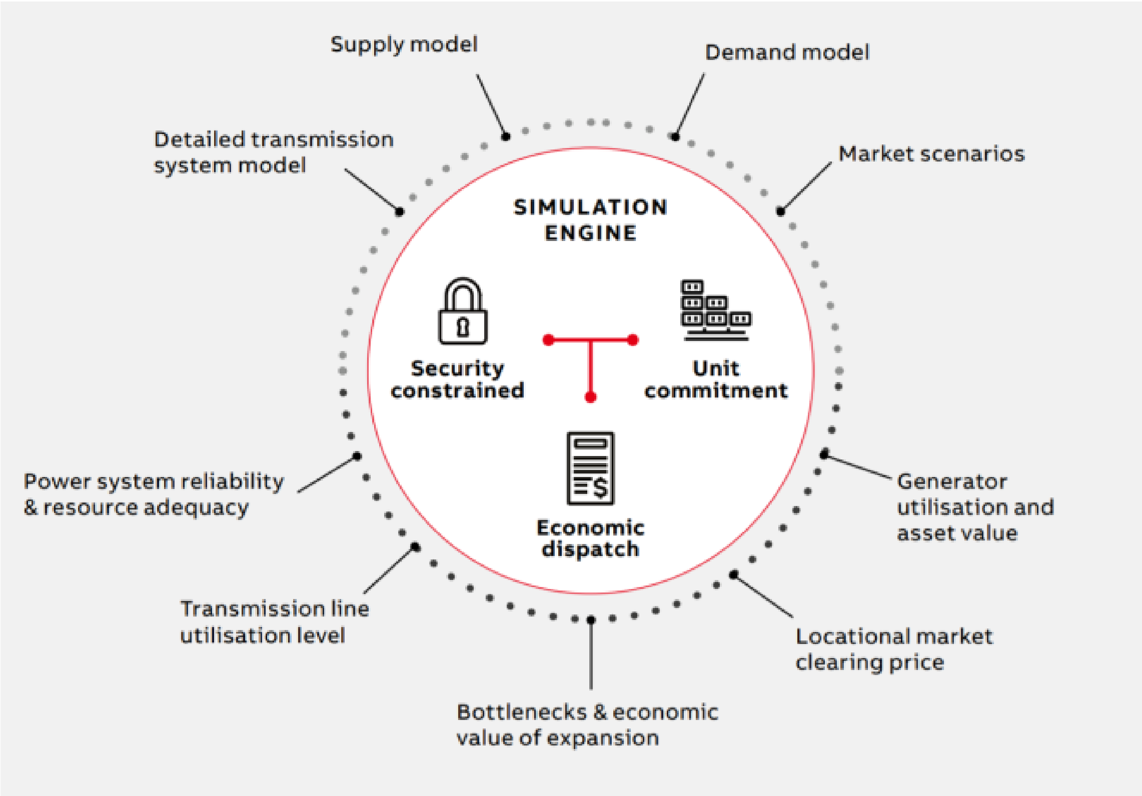

Figure 7.1 GridView is a production cost model that evaluates security constrained unit commitment and economic dispatch under a number of constraints (ABB 2023)

The WECC currently uses GridView as a tool for PCM efforts in the region. Within WECC's GridView PCM, expected loads, resources, and transmission topology ten years into the future are compiled and maintained by WECC staff. WECC's Anchor Data Set (ADS) Data Development and Validation Manual describes in more detail this data collection process and PCM practices ("System Stability Planning Anchor Data Set (ADS)" n.d.). The WECC PCM used in this study is the 2028 ADS V2.0 PCM base case made available as of July 2019. This case has the best available projection of new generation and transmission assets from the grid planning community within WECC at the time. Therefore, it was obtained and used for this project to get the most representative results. The study uses this case as-is and did not make any changes to resources, transmission, or topology contained within the case, aside from the addition of offshore wind resources.

Based on the data within the 2028 ADS V2.0 PCM, significant changes in generation resource mix within WECC are projected. However, the changes to Oregon are limited to a small number of additional solar photovoltaic (PV) plants. Otherwise, there is a significant amount of additional capacity in California, Arizona, Colorado, Nevada, and Utah, expected to come online within the next ten years, reflected in the PCM. This new capacity is predominantly forecasted to be solar PV and wind. Transmission in the WECC 2028 PCM case provides the best representation of future topology and transmission capacity available. It incorporates the addition of transmission projects in the 10-year planning horizon made publicly available to the grid planning community ("System Stability Planning Anchor Data Set (ADS)" n.d.).

In the WECC PCM, loads are modeled as hourly loads for the entire year by balancing authority. The load data within WECC's PCM is based on annual Load & Resource (L&R) data submittals from member BAs that contain monthly energy and peak load forecast for 1-10 years into the future. These data are then broken down from monthly to hourly data by applying the historical FERC Form 714 hourly load shape. The WECC 2028 PCM case currently uses a 2008 historic load shape to create the 2028 hourly load profile by applying the monthly peak load and total energy reported in the L&R. The historic 2008 load shape is an average load year with average weather conditions WECC-wide. For the purposes of this study, no changes were made to the load set by WECC in the model.

7.2 Example Simulation Results

WECC's GridView model is a powerful tool that can deliver many different types of regional insights for future market conditions. A PCM simulation in WECC using the GridView software will provide hourly load and dispatch of generation with high granularity. This level of detail can be post-processed in many ways depending on the analytical interests for a specific project. Several examples showing how GridView's output can be post-processed to deliver different insights are shown in Figure 7.2, Figure 7.3, and Figure 7.4 below.

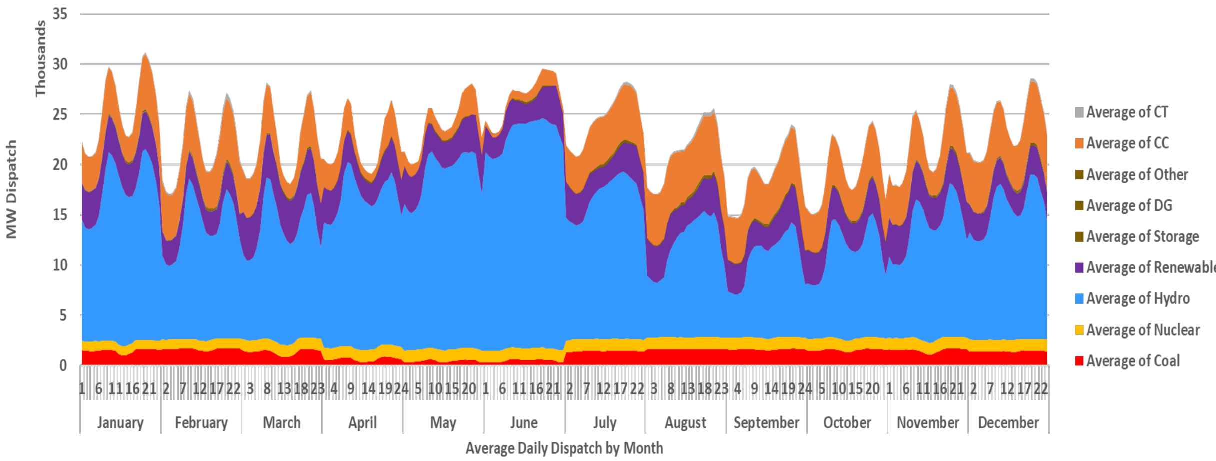

Figure 7.2. Sample GridView results reflecting the average daily generation dispatch in the Northwest Power Pool region

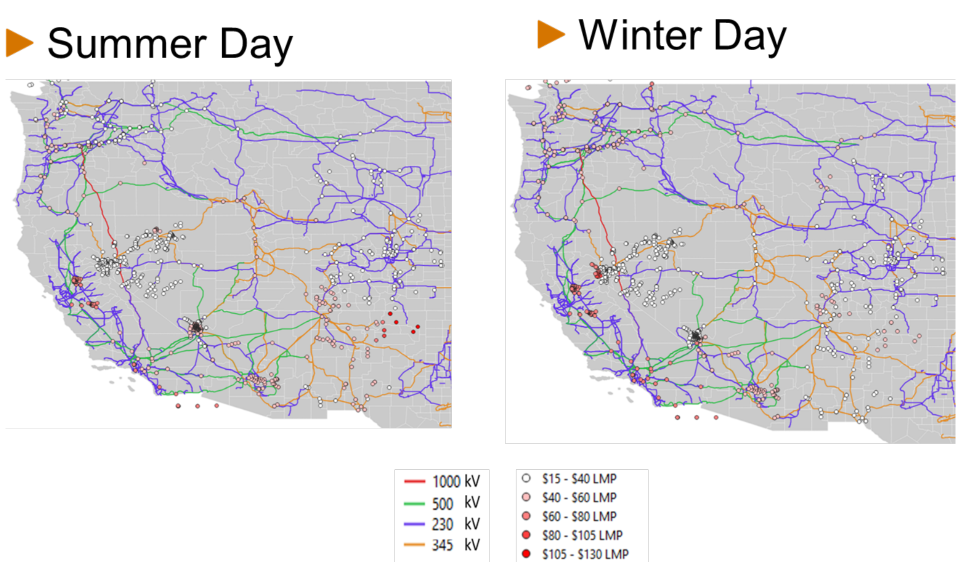

Figure 7.3. Sample GridView results reflecting nodal marginal prices during peak load hour in summer (Left) and winter (Right)

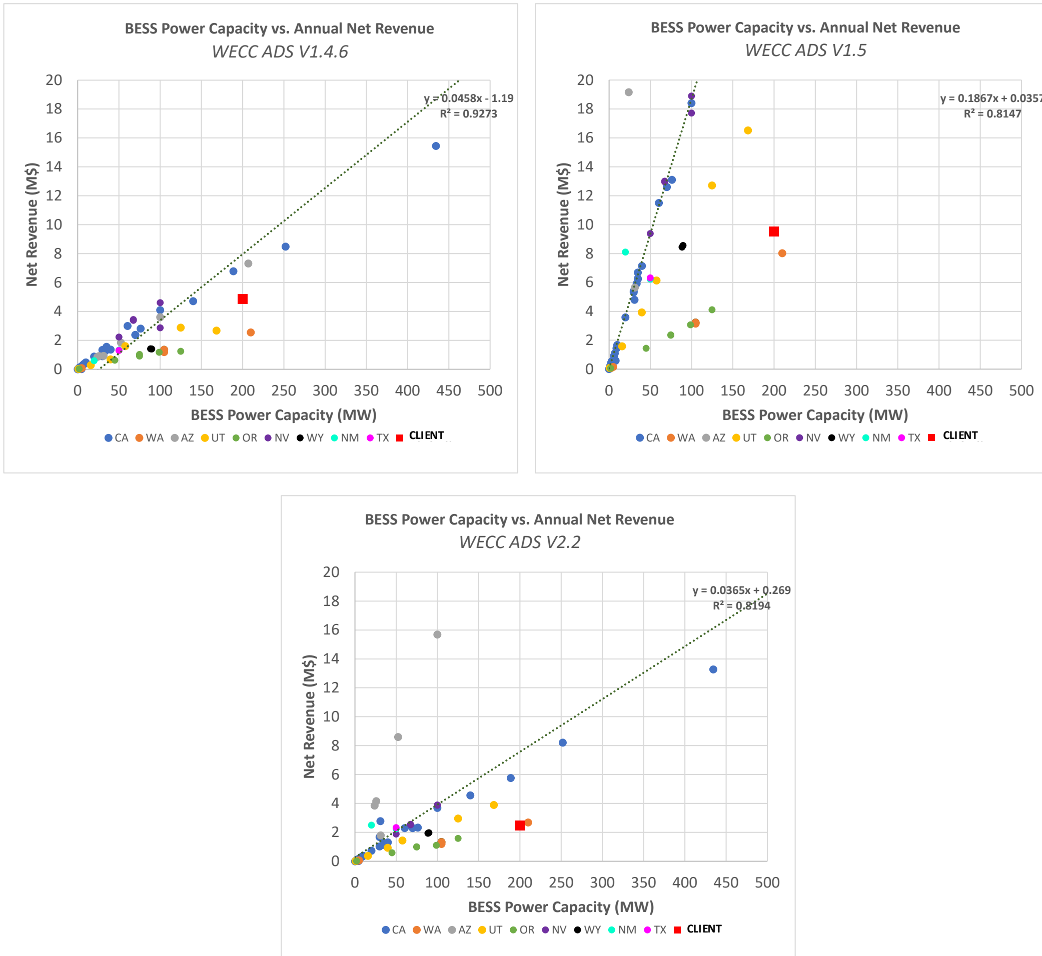

Figure 7.4. Sample GridView results comparing large-scale energy storage facility net revenues to installed capacity, by state, under different GridView market conditions

7.3 LDRD Simulation Configurations

The two main components we will alter in WECC's GridView base cases are hourly load profiles and generation mix. These inputs are reflected in the downscaling framework diagram in Figure 5.1 and are highlighted in Figure 7.5 below.

GridView simulations will be run for an entire year to acquire hourly system dispatch for all 8760 hours.

We do not intend to make significant changes to other market conditions reflected in the WECC GridView model. We will keep all other GridView simulation options equivalent to WECC's traditional practices, unless otherwise requested by stakeholders.

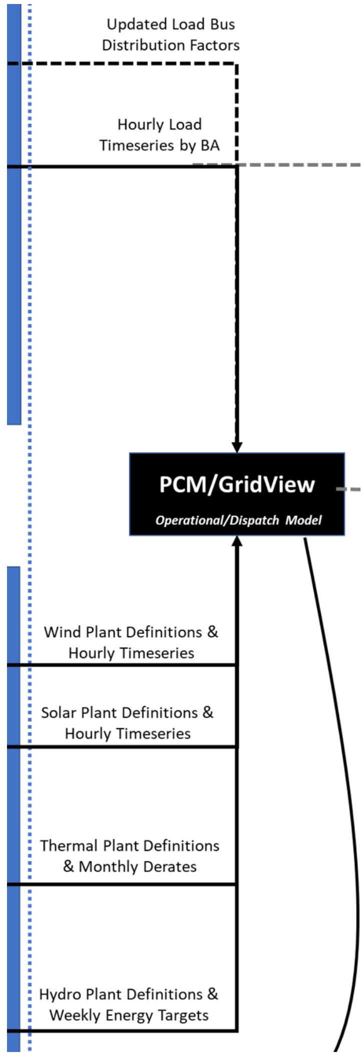

Figure 7.5. GridView input parameters for load, wind and solar will be adjusted for simulations under this project

7.4 LDRD Cost & Effort to Run and Postprocess GridView

GridView simulation setup, run-time, and post-processing will take a significant amount of time.

Configuring and importing GridView input files may take a day to a week, depending on the maturity of scripting capabilities developed to create the appropriate input file formats.

Running a year-long GridView simulation can take anywhere between 8-24+ hours, depending on available computational capabilities. In addition, under this simulation framework, GridView simulations will be required to run in 5-year increments. This simulation time can potentially be reduced by high performance computing.

Post-processing GridView results can also take significant time. Scripting the process of extracting the specific simulation result parameters of interest could be developed to make this process more efficient. An aspect of this post-processing effort will also be continuous validation of simulation results, especially as we push the system into higher renewable generation mix.

8.0 Plans to Leverage & Enhance IM3 Tools for Downscaling

IM3 includes experiments to understand how electricity system grid stress will evolve due to a variety of compounding short- and long-term stressors. A variety of tools are being developed to create hourly timeseries of load and power plant placement for future conditions looking out to 2100. This information will be used to run a publicly available production cost model (GO) on a reduced grid network (~300 nodes). IM3 tools are built on entirely publicly available data and will be made available to others as an open-source model. Note that this report represents IM3 modeling capabilities as of 2021-2022 and many enhancements have been made since then.

The IM3 modeling framework for modeling climate related grid stress is shown in Figure 8.1.

The methodology in this report follows a very similar structure to the simulation framework of IM3. The IM3 tools intended to be leveraged under this LDRD project include the siting model CERF, the load model TELL, and the hydropower and thermoelectric derating model based on MOSART-WM with its hydropower model and post processing. However, there are some gaps that we will need to address, and some preliminary methodologies for addressing those gaps are explored in this report.

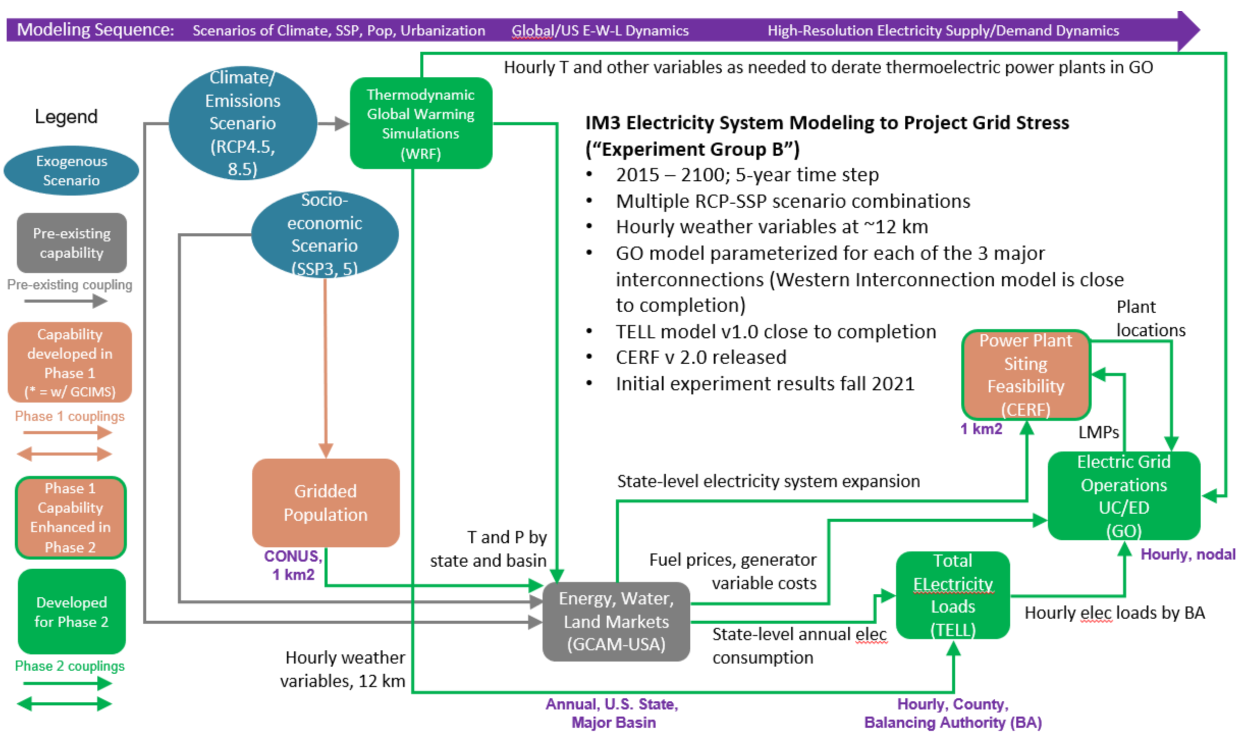

Figure 8.1. Framework for IM3 electricity system modeling to project grid stress

8.1 Capacity Expansion Regional Feasibility (CERF)

The Capacity Expansion Regional Feasibility (CERF) model is an open-source geospatial model, written in python and C++, that is designed to determine the on-the-ground feasibility of achieving a projected energy technology expansion plan. CERF is specifically developed to examine where power plant locations can feasibly be sited when considering high spatial resolution siting suitability data as well as the net locational costs (Vernon et al. 2018).

CERF converts state-level GCAM-USA electricity system expansion to individual power plants on a 1km grid for each time step. It sites renewables and non-renewables and retires power plants. A diagram illustrating CERF functionality is shown in Figure 8.2.

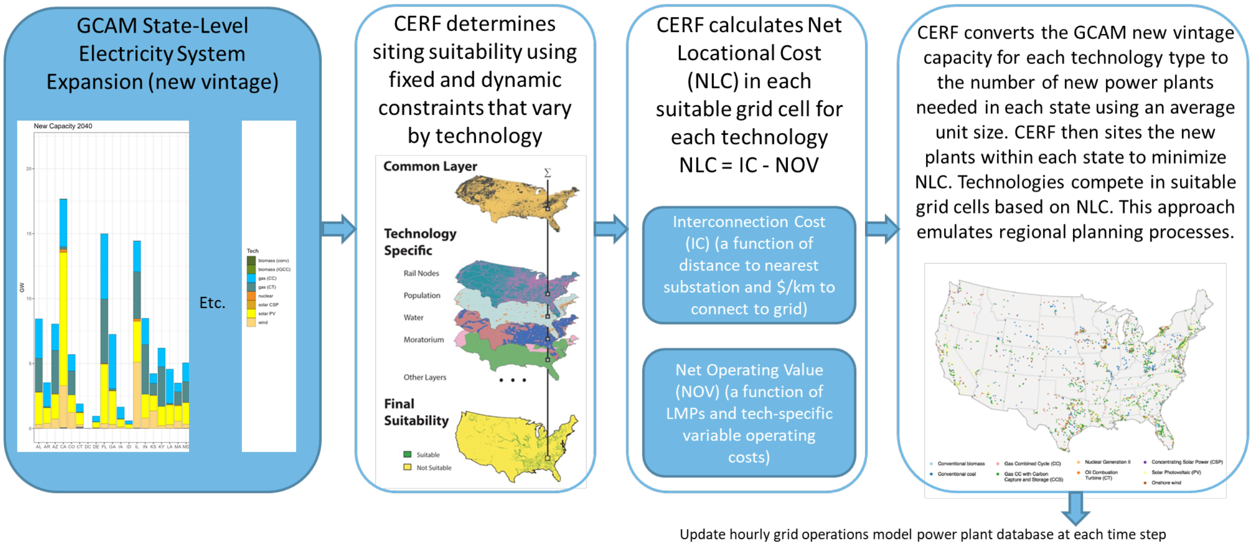

Figure 8.2. Summary of CERF functionality

General Notes Acquired in Early Discussions:

In discussions with the IM3 developing team for CERF, the following modeling characteristics of CERF were noted:

- Does not yet model off-shore wind or roof-top solar
- Renewable placement is harmonized with NREL assumptions
- System topology, including all nodes eligible for interconnection, is required as an input
    - Synthetic datasets (Texas A&M) are currently used for substation locations
    - A polygon is associated to each substation, and the same LMP is reflected throughout that polygon
    - Polygon creation is automated using Voronoi logic
- Decision for locational placement is based on Net Locational Cost calculation, which considers cost to interconnect to nearest substation minus the net operating value, as shown in Figure 8.2
- To calculate the net operating value, LMP input is required for every 5 years, which is attained by running GO at every 5 year time-step
- Generation capacity of power plants will be sized similarly, using capacity of typical size for that generating technology
- Only one generator can be selected per 1km grid location
- Timeseries data for renewables is not an output of CERF
- Only existing substation locations are considered as plausible interconnection locations
- CERF does not consider how climate impact could change the value placing renewables at certain locations
- Does not consider transmission capacity constraints

Observations with Respect to Applicability to LDRD:

To leverage CERF at the level of granularity needed for LDRD downscaling, the following questions need to be addressed:

- CERF requires **system node locations** as an input. The LDRD team would ideally like to feed WECC's full nodal topology as an input into CERF. This will be significantly more buses than IM3 topologies. Questions to consider:
    - Can we swap datasets easily?
    - Do we filter WECC buses? For example:
        - Do we only submit 115kV substations and above?
        - Are there other types of filters we want to apply?
        - Are interconnection costs calculated appropriately in CERF?
    - Do we have geolocation information for WECC buses readily available? Can we use this data right away?
    - Is CERF's reliance on LMP appropriate? Could it result in odd siting behavior, especially when congestion in other parts of the system cause negative prices?

- CERF requires LMP input at every 5-year timestep, considering new system additions predicted by GCAM. In IM3, GO will be used to acquire this hourly LMP, in which GO is a simplified PCM. Running GridView simulations for the WECC footprint every 5 years may be impractical, as each full-year simulation can take 8+ hours to complete. Questions to consider:
    - How to run GridView every 5 years to acquire LMP input into CERF? Can we jump straight to year 2050? No, we will run GridView every 5 years.
    - Should we use every 10 years?
        - The reason for running every 5 years is to avoid stranded investments.
    - Can and should we consider running GridView every 5 years?
        - The benefit of this is that high LMP's caused by transmission congestion could be considered – and therefore generation placement will better accommodate system topologies modeled in GridView.

- CERF does not supply timeseries data for renewables. This will need to be acquired using NREL datasets. Questions to consider:
    - What NREL datasets or tools can we readily use?
    - Can we write a script to acquire timeseries data in the file format desired?

8.2 Total Electricity Loads (TELL)

The Total Electricity Loads (TELL) model generates predictions of hourly total electricity load for every county in the continental U.S. using historical meteorology for a machine learning model training. Predictions from TELL are scaled to match the annual state-level total electricity loads predicted by GCAM.

The high-level overview of how TELL works is:

1. Formulate empirical models that relate the historical observed meteorology to the hourly time-series of total electricity load for each of the 68 BAs that report their hourly loads in the EIA-930 dataset.

2. Use the models to predict future hourly loads for each BA based on IM3 climate forcing.

3. Distribute the hourly loads for each BA to the counties that BA operates in and then aggregate the county-level hourly loads from all BAs into annual state loads.

4. Calculate state-level scaling factors that force the bottom-up annual state-level total loads to match the future annual state-level total loads from GCAM-USA.

5. Apply the state scaling factors to each county-level time-series of hourly total electricity loads.

6. Output yearly (8760 hr) time-series of total electricity demand at the state, county, and BA level that are conceptually and quantitatively consistent with each other, as shown in Figure 8.3.

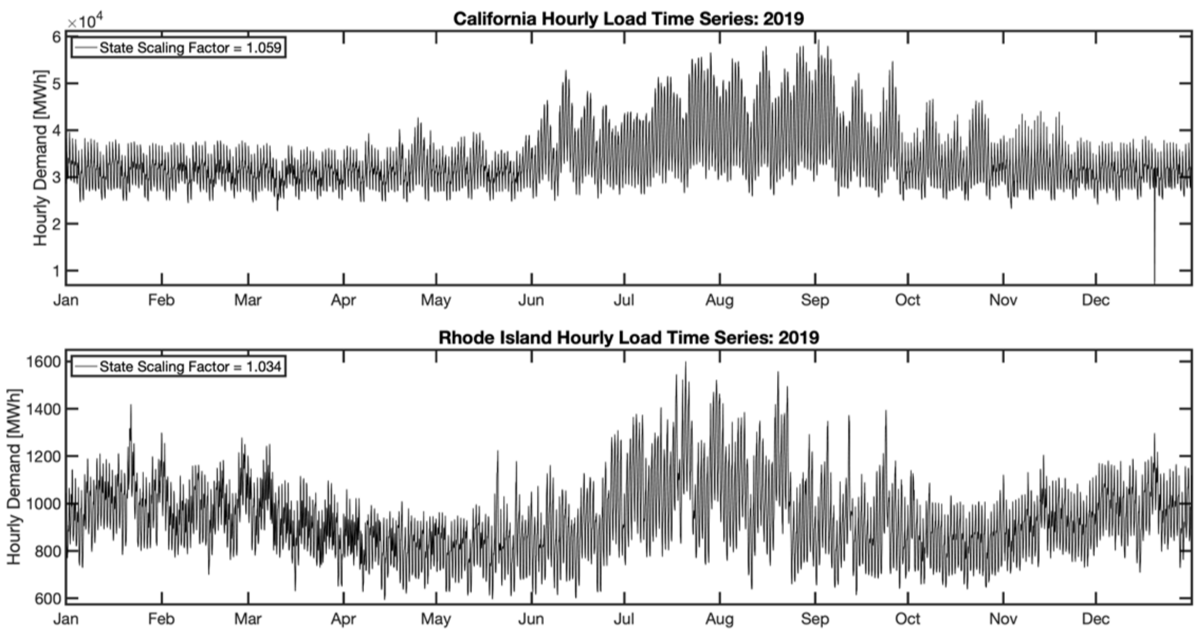

Figure 8.3. Example hourly load predictions for California and Rhode Island scaled to match GCAM energy output

General Notes Acquired in Early Discussions:

In discussions with the IM3 developing team for TELL, the following modeling characteristics of TELL were noted:

- TELL provides hourly estimates of total electricity demand in one-year (8760) increments through year 2100 using the Thermodynamic Global Warming dataset (Jones et al 2022).
- TELL produces demand projections that respond to changes in meteorology/climate and population
- Electricity demand is produced by county, state, and balancing authority scale
- Spatial mapping of BAs is uncertain and imperfect, and occasional bad historical data can impact accuracy of predictions
- A heat map of number of BA's per county was developed, as some counties may map to multiple BAs
- City population weight and their primary utility is used to account for county to BA mapping
- Meteorological predictions are reconciled with GCAM predictions (scaled up to match GCAM predictions by state)

- Model performance has only been calibrated with 2019 actual and predicted data
- Comparisons of predictions in 2100 to GCAM predictions have not yet been completed
- Does not consider load shape changes with newer technologies expected in the future
- Could be enhanced with additional downscaling methodologies to better overlay demand caused by new technologies on top of rest of load

Observations with Respect to Applicability to LDRD:

To leverage TELL at the level of granularity needed for LDRD downscaling, the following questions need to be addressed:

- How can we enhance TELL to incorporate impact to load profiles by technology type (residential, industrial, transportation, etc.)?
- How can we enhance TELL to incorporate impact to load profiles from increased levels of DER penetration?

8.3 Grid Optimization model (GO)

GO is a publicly available production cost model. The computation speed is significantly faster when compared to GridView due to its reduced nodal topology. We do not intend to use this tool in this LDRD project, as we are replacing the PCM functionality with WECC's GridView model which will more accurately represent the system topology. Visualizations of GO are shown in Figure 8.4 and Figure 8.5.

General Notes Acquired in Early Discussions from the GO team:

- Uses datasets from Texas A&M
- Reduced topology algorithm from Texas A&M
    - User specifies the nodes that need to remain
    - Allocates the load in the main notes such that the power flow in the lines is retained
    - Evaluates the trade-off between LMP in 250 nodes vs 100 nodes vs 75 nodes in a simplified network using the 2019 generation mix
- CERF can use a higher fidelity topology network – geospatial layer of all substations in the United States
    - It takes LMPs from GO in each iteration as shapefiles – one LMP value per polygon – 8760 – averaged LMP is used for CERF (to calculate relative value)
    - CERF is looking at LMPs from the next time period by looking at demand increment in the 5 years, with the old generation
- Changes to transmission capacity not yet considered – the addition and retirement of capacity at each timestep can cause transmission capacity limits to be exceeded
    - 80% of transmission lines are anticipated to be congested in future
    - Could incorporate planned transmission expansions
- GO is optimizing day ahead

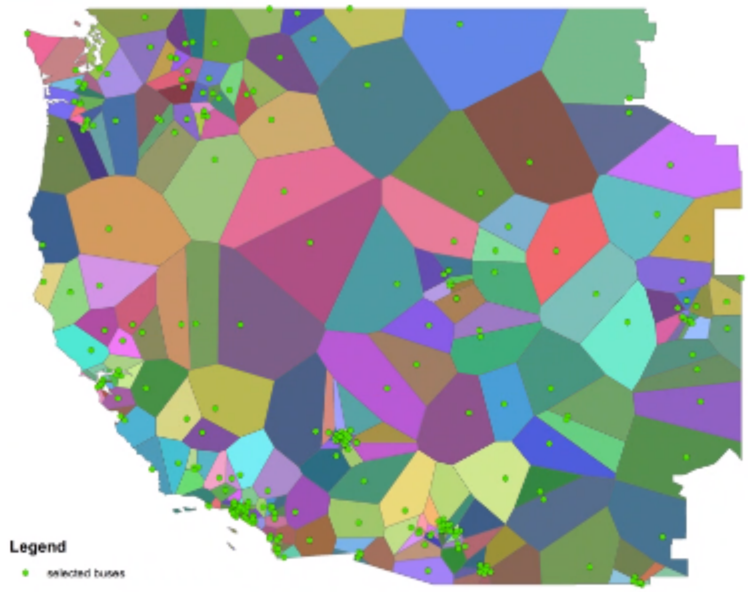

Figure 8.4. Sample of simplified topology to be used in GO simulations in IM3

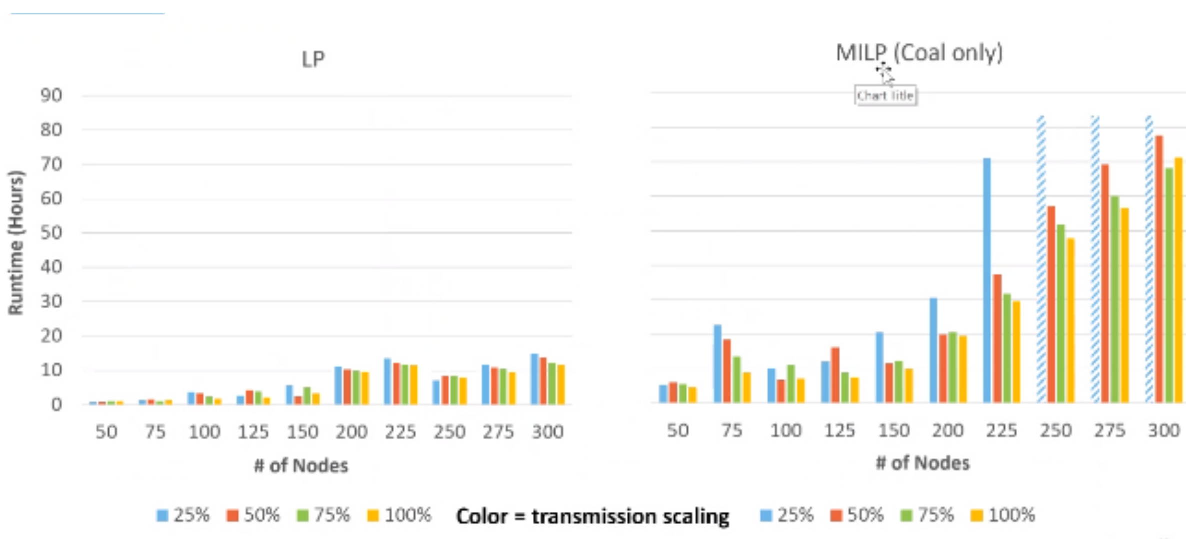

Figure 8.5. Preliminary analysis of GO runtimes without high performance computing

8.4 MOSART-WM

The large-scale water management (WM) model (Voisin et al. 2013), coupled with the river routing model MOSART (Li et al. 2013), simulates “regulated” river flow with spatially distributed water extraction that draws on knowledge of localized demand for water and dam flow regulations and represents seasonal variations (see Figure 8.6). The model is an integration platform for land surface models to interact with integrated assessment models such as GCAM and assess the sustainability and resilience of environmental systems to changes in natural conditions and human activities (sectoral water demands).

Under this project, we intend to use MOSART-WM in its current state of development, as it already delivers the functionality needed to downscale hydro projects into GridView inputs.

General Notes Acquired in Early Discussions:

- Climate forcing is gridded
- How much water is available in each grid cell is determined
- Timeseries of demand at each grid cell come from USGS
- MOSART is a river-routing model
- Output is the regulated flow at every grid-cell (not just at the hydro-facilities)
- MOSART-WM is already coupled with GCAM
- Can be directly used with GridView currently
- GCAM assumes hydro in a fixed amount
- MOSART-WM will consider climate impact to river flow
- Post-process the storage level – will provide weekly energy targets, and min/max generation for GridView
- WECC is switching from monthly to weekly modeling; MOSART-WM supplies both

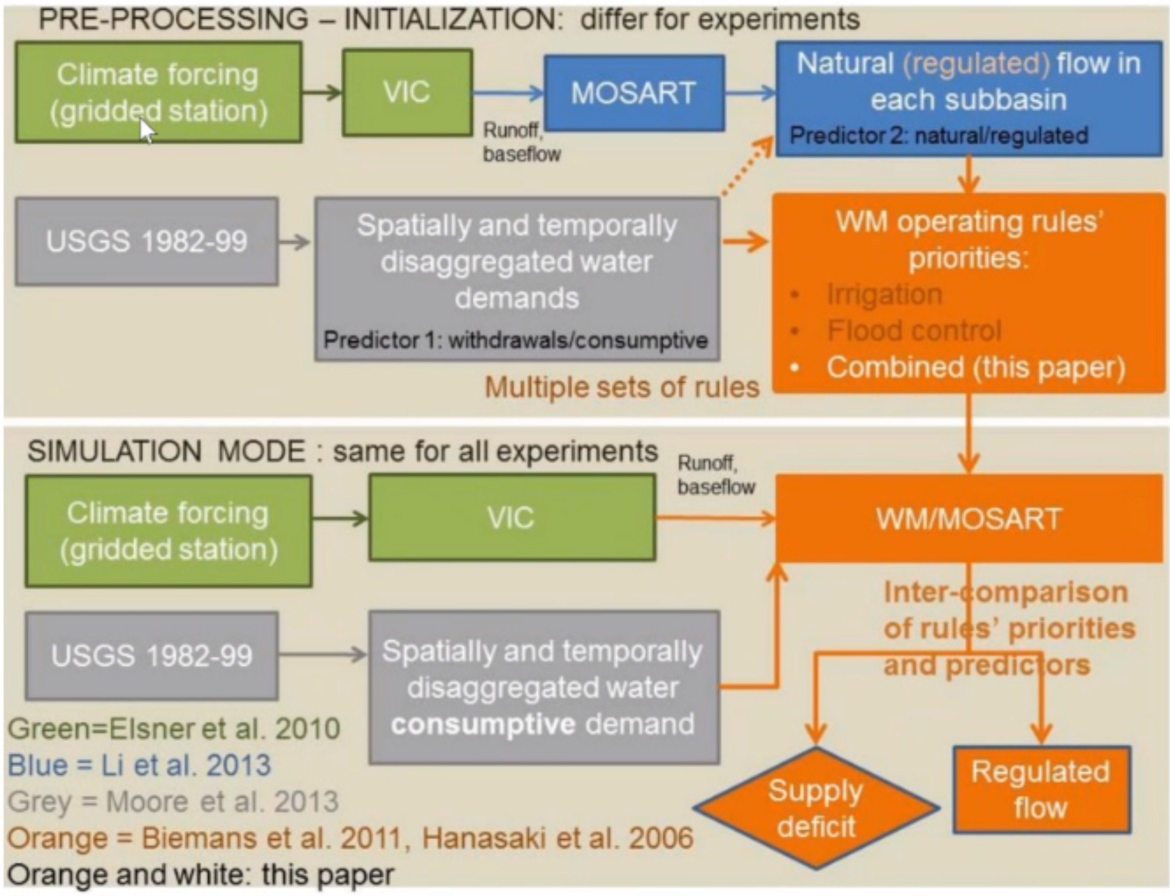

Figure 8.6. Diagram showing the inputs and outputs of MOSART-WM (Source: Voisin et al. 2013)

8.5 LDRD Downscaling Gaps that Remain

There were two major gaps identified after exploring the IM3 tools in development. One of these gaps is downscaling load to hourly profiles such that new technology adoption and its impact to load shapes are captured. And secondly, modeling hourly renewable output required for PCM in GridView.

8.5.1 Modeling Load Shape Impacts of Technology Adoption

GCAM-USA provides state-level energy consumption information for residential, transportation, industrial, and commercial loads. Despite this level of detail, the IM3 TELL tool does not consider the technology adoption impacts within these GCAM load classifications to inform the hourly load shape. IM3 TELL tool primarily relies on historical load correlation with weather to inform a machine learning model that can develop hourly load profiles for future conditions. The gap this LDRD would like to address is improving hourly load projections by considering how technology adoption (such as transportation electrification and DER penetration) will adversely impact load shapes in the day, especially if left unmanaged.

The following discussion will focus on a methodology for incorporating transportation electrification charging into the creation of regional hourly load profiles. Incorporating DER is also to be considered in impacting load profiles (as shown in Figure 5.1), however, generating DER solar profiles will follow similar methodology discussed in 8.5.2 (to be netted with load profiles).

As a minimum viable product for the next fiscal year, the goal of transportation load downscaling will be to obtain BA level load hourly profiles contributed by electrification such that the yearly energy projected by GCAM is matched. Since energy projection is already accounted for in GCAM, the purpose of transportation load downscaling is only focusing on various types of a vehicle's charging behavior. One of the most important aspects of the downscaling is to ensure the GCAM assumption and the load shape generator assumptions are the same. We would identify such factors that need to be matched with assumptions made during GCAM scenario creation as we outline the downscaling methodology.

To project transportation load shape for a future year, there are several assumptions that need to be made. These are: 1) BA level projected fleet size based on vehicle class, 2) charging behavior based on each class, 3) battery size projection.

The model is simplistic and does not account for factors such as major technology disruption, since there is large uncertainty around such possible disruptions. However, several scenarios would be considered corresponding to various decarbonization levels. In this subsection, these assumptions and how those would be translated into load shape are discussed at length. Most of the assumptions are borrowed from our previous study reported in a DOE funded transportation electrification PNNL study conducted in 2020 (Kintner-Meyer et al. 2020).

8.5.1.1 Balancing area level projected fleet size based on vehicle class:

There are two components of projecting the future fleet size- a) the growth of class wise total vehicle fleet (EV and non-EV) for a future year and b) portion of the fleet converted to EV. There are several sources of recent total registered vehicles in USA. One such source is the US Federal Highway Administration dataset ("Table MV-1 - Highway Statistics 2017 - Policy | Federal Highway Administration" n.d.). GCAM uses similar datasets along with future projections to report total class wise fleet size for a given year. This projection would be used to obtain BA level fleet size based on vehicle classes, light duty vehicles (LDV), medium duty vehicles (MDV) and heavy-duty vehicles (HDV).

8.5.1.2 Charging behavior based on each class:

To obtain the charging behavior, we need to look at the vehicle types light duty vehicles (LDV), medium duty vehicles (MDV) and heavy-duty vehicles (HDV) separately.

Light Duty Vehicles (LDV): To obtain LDV load shape, we would utilize the EVI-Pro tool ("Alternative Fuels Data Center: Electric Vehicle Infrastructure Projection Tool (EVI-Pro) Lite Assumptions and Methodology" n.d.; Lee et al. 2021) developed by NREL. EVI-Pro has been parameterized for the simulation of nearly 50,000 real-world travel days from the 2012 California Household Travel Survey across 14 unique ambient temperatures (-20°C to 40°C), 8 representative PEV types (PHEV20, PHEV50, BEV100, BEV250 with sedan and SUV variants for both), and 36 combinations of charging behavior and technology types (including varying home/work charging power and availability, and home or work as the preferred charging location).

The resulting product is a database of over 200 million unique simulated charging events. Each simulated charging event is described in the database based on simulated vehicle ID, arrival and departure time at the charging location, start and end time for the charge event, destination type (home, work, or public), charge level (L1, L2, or DCFC/level 3), and energy dispensed during the charging event (expressed in kWh).

The tool generates sub-hourly load shapes for a given fleet of LDVs based on temperature, fleet size, average miles traveled etc. Other configurable parameters are the distribution of PHEV, sedan vs. SUVs, share of levels of chargers, charger access at home and work, and charging strategies. Among these, fleet size, vehicles miles traveled, sedan vs. SUV share assumptions could be obtained from GCAM. The tool also distinguishes between weekday and weekend charging. In order to obtain various scenarios for load shapes, managed and unmanaged charging strategies are developed. EVI-Pro allows for the charging strategies ranging from no delay to maximum delay at plugging in after arriving at the charging station, which could be utilized to create such scenarios.

Figure 8.7 shows what EVI-Pro would need as input to generate load profile. Only major inputs are identified which would be used for this study.

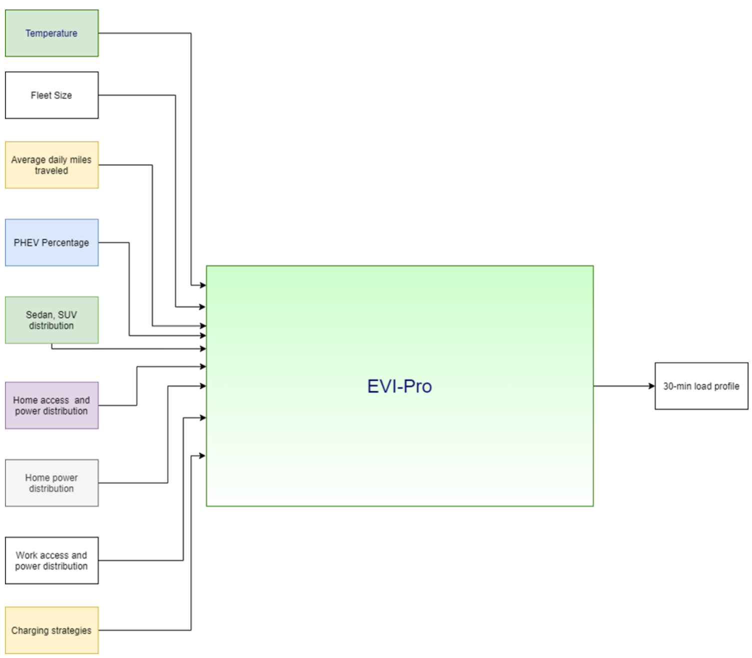

Figure 8.7. LDV load shape generation using EVI-Pro

Medium Duty Vehicles (MDV): For medium duty vehicle charging profile, typical driving patterns and energy use would be mapped from GCAM. The charging strategies heavily depend on battery size, weight and preference of slow depot charging vs. enroute fast, opportunity charging. We would obtain data for typical down time of various MDVs during the day. One source of various MDVs driving behavior is found in NREL's FleetDNA dataset ("Fleet DNA: Commercial Fleet Vehicle Operating Data" n.d.). The core requirements of development of the MDV load shapes are shown in Figure 8.8.

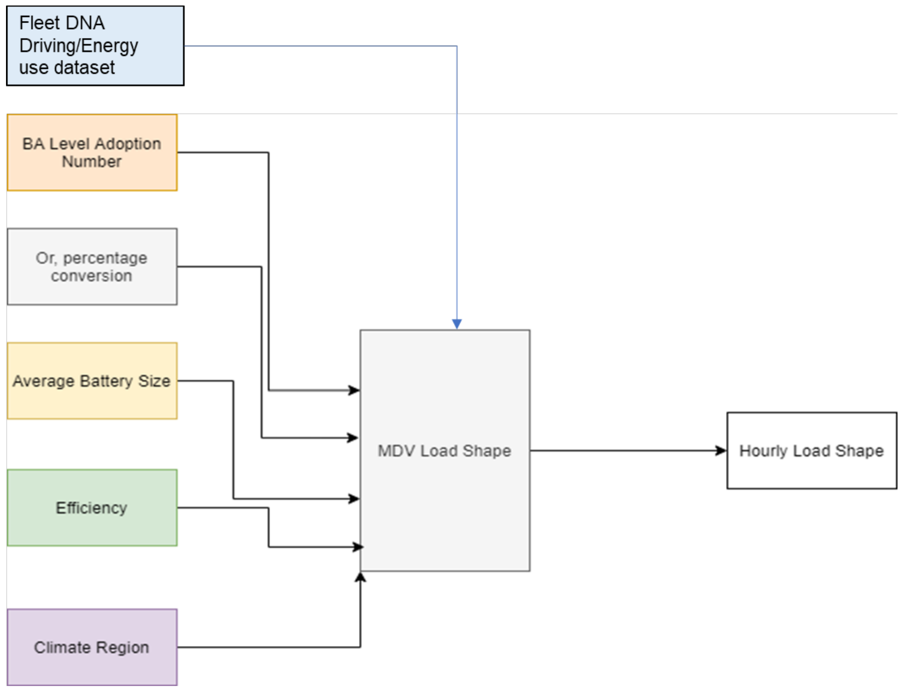

Figure 8.8. MDV load shape generation using PNNL methodologies

Heavy Duty Vehicles (HDV): The load profile for HDVs is the most complex as the HDVs move among various states and balancing authorities. Therefore, static, registration-based estimation might not work. For this effort, we would leverage our previous HDV load shape generation done in a PNNL study (Kintner-Meyer et al. 2020), whose charging station assumptions are shown in Figure 8.9.

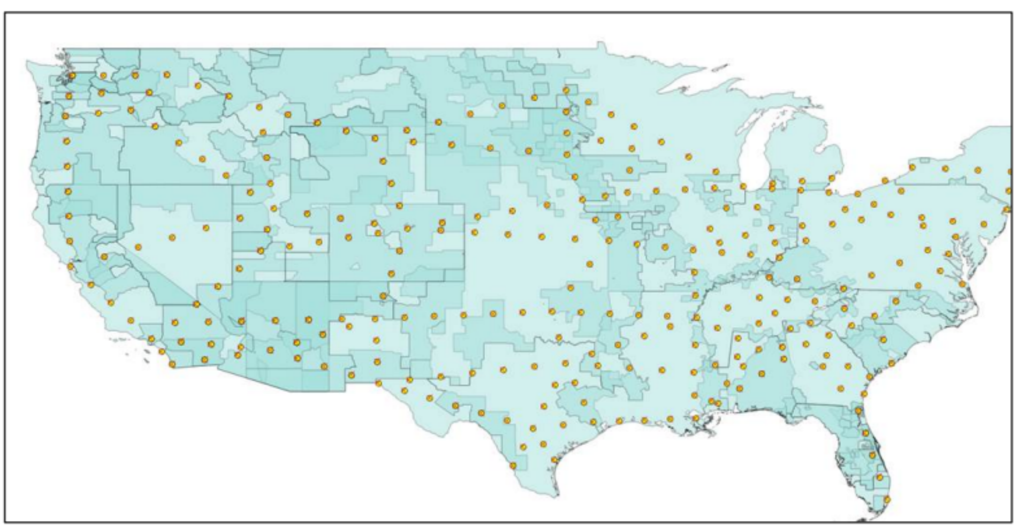

Figure 8.9. HDV charging station assumptions used in PNNL agent-based model

8.5.2 Modeling Hourly Renewable Output that Correlates with Meteorology

PCMs, such as GridView, rely on hourly production from individual generators as input. In the 10-year horizon planning studies that the WECC undertakes, the synthesis of these production time series draws largely from the NREL Solar and Wind Power Data for Integration Studies ("Solar Power Data for Integration Studies" n.d.; "Eastern and Western Data Sets" n.d.). These datasets are the result of an enormous effort to create synthetic solar and wind hourly production time series. The primary limitation to use of these datasets is their reliance on a handful of meteorological years (i.e., 2004 through 2006) which do not coincide with the meteorological inputs to the simulated hourly load from TELL. This section presents three methodologies of increasing sophistication for the development of renewable energy generation time series at horizons congruent with the results from GCAM.

The methodologies for calculating the hourly resource generation are shown in Figure 8.10. The simplest approach is to harness the existing NREL Solar and Wind Power Data for Integration Studies. This approach, "NREL Datasets," will require the least amount of effort. Existing datasets will be leveraged with minimal post-processing, requiring roughly 30 labor hours. The second methodology, "Generate Production from WRF," will leverage the power production methodologies developed for the NREL Integration Studies and incorporate the same meteorological inputs as those used in TELL. This effort will be more involved, requiring roughly 100 labor hours. The third methodology, "GP+GBM," which is an abbreviation of "Gaussian Process plus Gradient Boosting Machine," is the most sophisticated, aiming to bridge a gap in greenhouse gas abatement studies of this scale through building on lessons learned from machine learning in the operational energy forecasting literature. This effort will be state of the art, requiring roughly 250 labor hours.

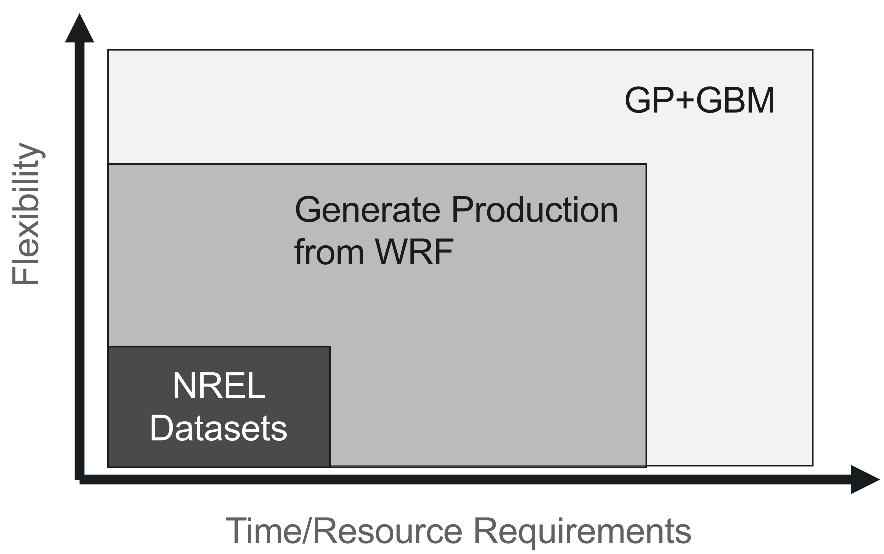

Figure 8.10. Three different methodologies and their increasing complexity that could be developed under this project to create hourly renewable output

The following sub sections detail the methodologies. Figure 8.11 demonstrates the element within the larger framework (Figure 5.1) which will be developed and includes a new external dataset, the WECC 2030 Anchor Dataset (Item 4). In all the methodologies detailed below, capacities for wind and solar are provided by GCAM and siting (latitudes and longitudes) for wind and solar are provided by CERF.

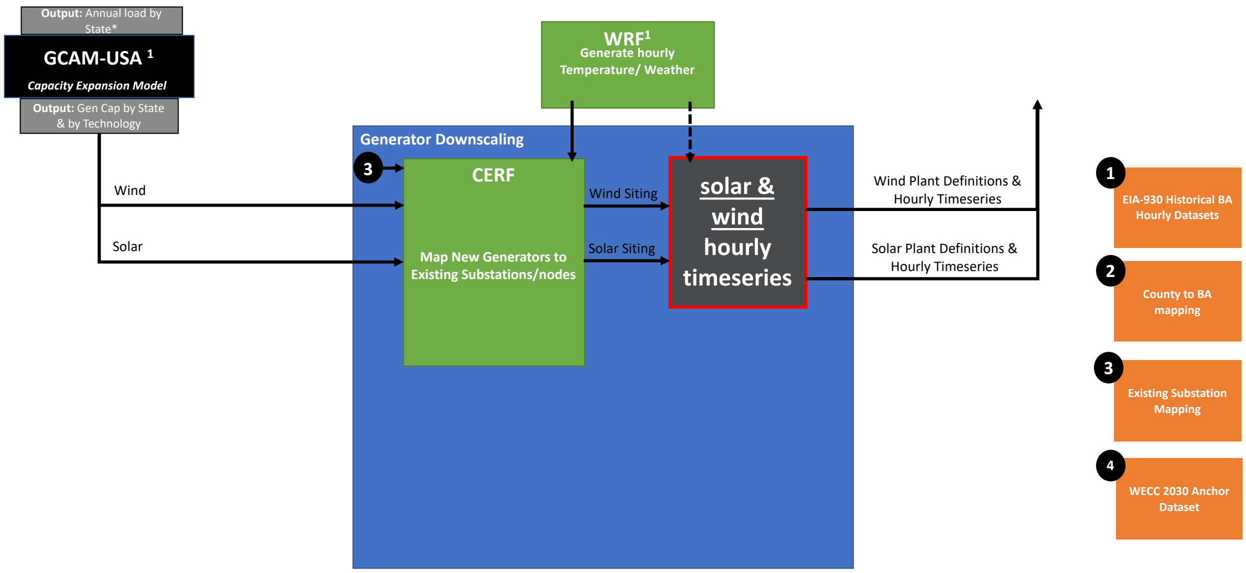

Figure 8.11. Renewable hourly timeseries element in the larger downscaling framework to be developed

8.5.2.1 Methodology Option 1: NREL Datasets

The simplest methodology leverages datasets created by NREL for the Western Wind and Solar Integration Study (WWSIS). A schematic of this methodology is shown in Figure 8.12. For each resource type (wind, PV, CSP/concentrated solar), a representative WWSIS site is selected by the minimum Euclidean distance to the latitude and longitude of the new generator identified by CERF. Then the production time series is normalized to the maximum production and scaled to the generation identified by GCAM. These steps rely on WWSIS site macro data for generator locations and capacities (Draxl and Mathias-Hodge 2016).

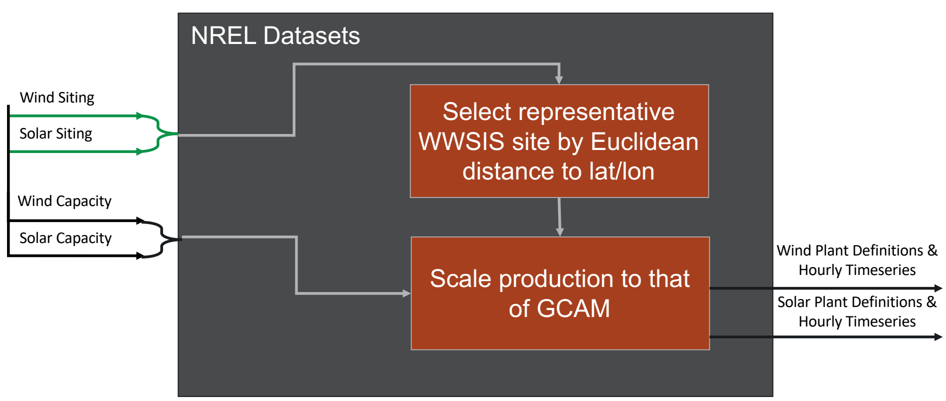

Figure 8.12. “NREL Datasets” methodology for generating renewable hourly timeseries

While leveraging existing NREL wind and solar hourly production datasets reduces the cost burden for this project, this approach is crippled by its inability to reflect meteorological conditions in the target forecast horizon. This is especially crucial to wind and solar production time series, which are largely dependent on meteorological physics, such as wind speed and global horizontal irradiance (resulting from cloud cover). The NREL WWSIS Datasets provide hourly data for solar for the year 2006 and for wind for the years 2004 through 2006. While these datasets are expertly calibrated to the meteorological conditions of these years, they will not be able to represent the weather conditions of target forecast years (such as 2035 through 2050 and 2100) and they will not correlate with the weather conditions built for the TELL model.

8.5.2.2 Methodology Option 2: Generate Production from WRF

The NREL WWSIS Datasets include instructions for reproducing hourly production time series with the Weather Research and Forecasting (WRF) model from 2004 through 2006 and resource-specific tools (King, Clifton, and Hodge 2014; Hummon et al. 2012). This methodology can be extended to rely on the WRF model (called the thermodynamic global warming dataset) which is used to create load forecasts with TELL in future year scenarios (2035, 2040, 2045, 2050, …, 2100). This adherence to congruent physical inputs will produce hourly renewable energy production time series which correlate with load time series. Without this correlation, PCM tools will be calibrated to differing underlying physical models.

The process for generation of renewable energy production time series from WRF is shown in Figure 8.13. For each resource type, the same WRF year is used as in the TELL model. The location of the generator (provided by CERF) is used to extract the hourly meteorological time series from WRF. Wind hourly time series for a given location requires this meteorological time series, the NREL Wind Toolkit Power Curves, and the generator capacity (provided by GCAM). These Power Curves, shown in Figure 8.14, provide a normalized output for a given turbine class (King, Clifton, and Hodge 2014). The turbine class is identified by the wind speed at the generator location using the NREL WIND Toolkit (Draxl et al. 2015). In each hour, the wind speed provided by WRF is traced to a normalized output and scaled to the installed capacity.

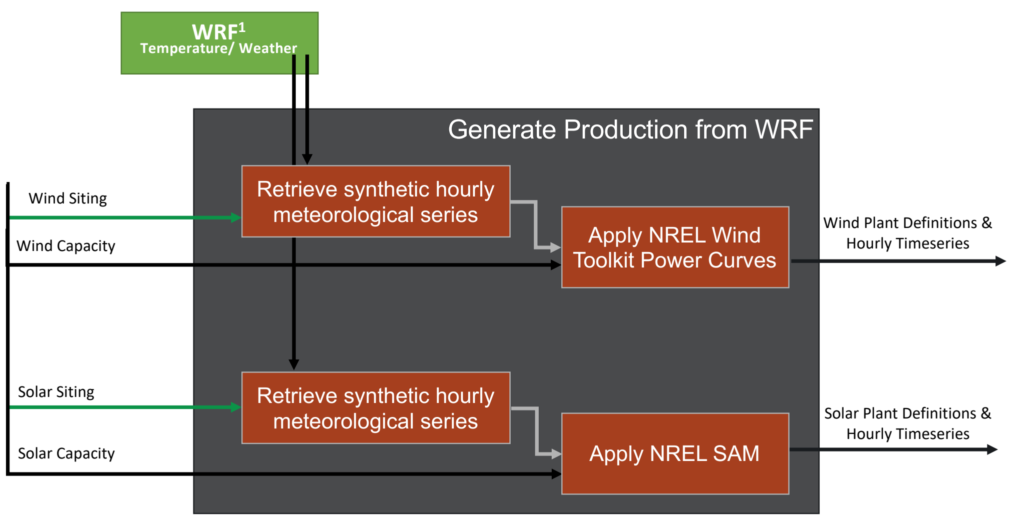

Figure 8.13. “Generating production from WRF” methodology for generating renewable hourly timeseries

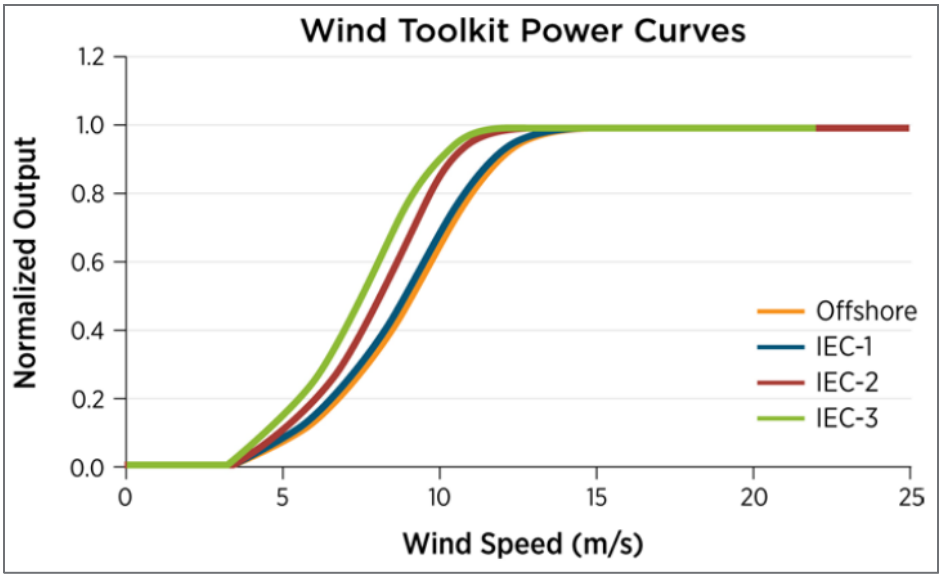

Figure 8.14. Wind toolkit power curves reprinted from https://www.nrel.gov/docs/fy14osti/61714.pdf

As with the wind production time series, the location of the solar power generator informs the hourly meteorological time series from WRF. This time series is imported into NREL's System Advisor Model for each generator ("Home - System Advisor Model (SAM)" n.d.). This process can be automated with software such as pvlib (Holmgren, Hansen, and Mikofski 2018). SAM is a highly flexible tool; the PVWatts module within SAM can be used to simulate solar production with generic inputs and site-specific installed capacities.

8.5.2.3 Methodology Option 3: GP+GBM

The methodology options presented above hinge on the renewable energy production algorithms developed by NREL for the Western Wind and Solar Integration Study, conducted in the 2010s to create synthetic plant-level time series for as-yet unbuilt generators. This dataset is built from industry knowledge of wind turbine power curves and PV panel components at a snapshot in time. Reliance on this technique alone ignores the actual production of generators, which can often include curtailment, unplanned outages, and locally correlated weather events.

Methodology option 3 proposes a Grey-Box solution to incorporate elements not captured directly by the NREL renewable energy production algorithms: Extend the White-Box physics-based approach of the NREL WWSIS results with a Black-Box machine learning algorithm. By leveraging advances in machine learning developed for operational forecasting of energy production, this approach will allow the project to estimate the bias in these decades-ahead forecasts and constrain the variance in estimates. Not only will the project have more accurate hourly forecasting of production, but it will also be able to measure its accuracy. This component is vital, as it provides knowledge about the generalizability of the methodology to future years (future weather) and new generation sites (future technology).

There are two challenges well-suited to the inclusion of machine learning techniques: (1) spatial interpolation of existing generators to synthetic new generation, and (2) integration of actual historical production with synthetic new generation. These two challenges will be addressed with Gaussian Processes (GP) and Gradient Boosting Machines (GBM), lending the name GP+GBM. A Gaussian Process is a collection of joint Gaussian distributions of functions, where the covariance function constrains the flexibility of a proposed formulation to observations. GPs have proven highly capable of generating a probabilistic surface with fidelity to spatial and temporal data (Gelfand and Schliep 2016; Datta et al. 2016), as well as high-resolution interpolation applied to temperature and electricity load forecasting (Lloyd 2014). A Gradient Boosting Machine (Friedman 2001) is an ensemble method of function approximation by iterative and additive resolution of outliers selected by the loss function. This machine learning technique has consistently outperformed other tools in forecast accuracy of both load and power generation (Persson et al. 2017; Lloyd 2014).

GP+GBM will be used to calculate future hourly plant-level generation that incorporates fidelity to historical actual production and concurrent meteorological conditions. Figure 8.15 demonstrates the necessary inputs and order of operations for processing. The GP will be used to augment the training dataset for the GBM, where the regressors will be plant-level capacity and hourly meteorological conditions from WRF.

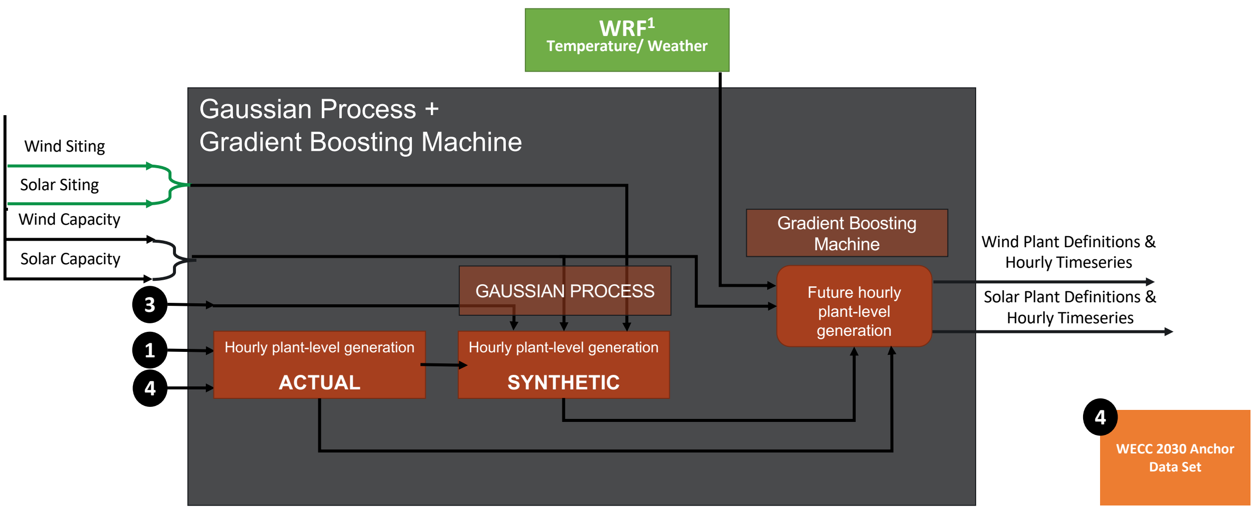

Figure 8.15. "GP + GBM" methodology for generating renewable hourly timeseries

The process diagram shown in Figure 8.15 demonstrates two steps: (1) development of the training set, and (2) application of GBM predictive model. The training set will be developed with actual historical production time series and actual historical meteorological conditions provided as inputs to the WRF model. The hourly actual plant-level generation is produced by combining the hourly BA-level generation from EIA-930 (dataset 1) with plant-level disaggregation factors from the WECC 2030 Anchor Data Set (dataset 4). The historical hourly plant-level generation is the regressand, and the plant-level capacities and WRF historical hourly weather conditions are the regressors fed to the GBM. This is demonstrated in Equation (1), where *cap* refers to the capacities of each *n* generators, there are *m* variables used from the *WRF* model, and *gen* is the historical hourly plant-level generation. This formulation will be applied to existing generation; these datasets for actual generation span the years 2015 through 2021 (union of EIA-930 and WRF time series).

$$\beta_1 \begin{bmatrix} \text{cap}_1 \\ \vdots \\ \text{cap}_n \end{bmatrix} + \beta_2 \begin{bmatrix} \text{WRF}_{1,1} & \cdots & \text{WRF}_{1,m} \\ \vdots & \ddots & \vdots \\ \text{WRF}_{n,1} & \cdots & \text{WRF}_{n,m} \end{bmatrix} = \begin{bmatrix} \text{gen}_1 \\ \vdots \\ \text{gen}_n \end{bmatrix} \qquad (1)$$

Equation (1) will be fit with a GBM and cross-validated with an expanding training set rolling forward in time, also called sliding window cross-validation. This means that unique models will be fit for each train and validation pair. Figure 8.16 demonstrates that seven models will be fit, where the set on the left indicates the training set, and the set on the right the validation set. The error in the validation set averaged across all seven models will provide an estimate of the error in using these regressors for prediction beyond 2021.

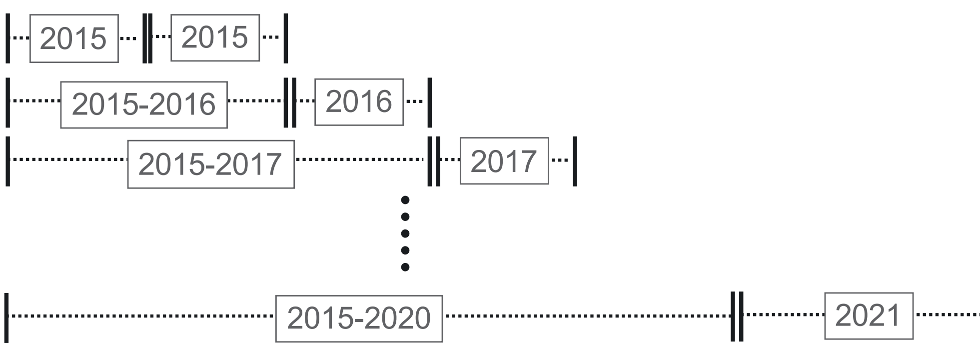

Figure 8.16. Sliding window cross-validation for GBM

The formulation presented in Equation (1) assumes the training set only contains actual historical generation. Prediction of new generation with this formulation requires synthesis of historical time series for training. A nearest-neighbor Gaussian Process (GP) model will be developed to spatially interpolate historical actual production to synthetic historical production from new generators. This model will require as inputs the actual hourly plant-level generation, locations (lat/lon) of these existing plants (dataset 3), and capacities and locations of future plants provided by CERF.

The GBM model derives its strength from training on actual historical time series. To test the impact of the inclusion of synthetic generation on prediction accuracy, the GBM framework in Equation (1) will be re-evaluated with an increasing percentage of actual historical generators excluded. This is intended to mimic the modeling error of introducing a large percentage of new generation by CERF and will serve to inform the team about potential bias in results.

There will be two levels of complexity to the GBM. The first will develop a single GBM predictor for each technology type (wind, PV, CSP/concentrated solar), where this generalized predictor will be fed the individual generator capacities for each location and the WRF meteorological time series at a future year. The second will drop the assumption of static coefficients to the GBM predictor for each technology class, which assumes complete pooling of information within a class. The alternative to complete pooling is no pooling of information – a single predictor is built for each generator. Instead of either, the coefficients will be modeled using prior knowledge of their distributions – one predictor will be built for each technology class and the coefficients tuned by the GBM will be drawn from an adaptive regularizing prior. In this way, the GBM will provide unique coefficients for each generator while sharing information within technology types, building on recent cutting-edge work in machine learning (Miller, McArtor, and Lubke 2017; Griesbach, Säfken, and Waldmann 2021).

9.0 Value by End-User

The outcome of this proposal, linking GCAM and GridView (and other PCMs), can be used by multiple DOE sponsors. For EERE this tool can be used to assess the electricity sector transformation under multiple energy efficiency and renewable energy pathways. Particularly, the Strategic Programs office is interested in the interaction between the demand and supply sectors. We will design a flexible linking tool that can be used to connect to other existing PCMs such as PLEXOS.

The stakeholders we intend to collaborate with will also benefit from such efforts and the resulting longer-term vision. Working with the WECC to create a proof of concept helps demonstrate value added to current electricity planning practices and important industry input and collaboration. With this effort, WECC members will be able to do advanced studies on new decarbonization scenarios not yet considered. Working closely with EERE and OE will allow increased transparency and understanding of how such a tool can be utilized and its future implications to industry and modeling practices. Working with industry organizations like Northwest Power Pool and utility members to gain feedback along the way will increase their confidence in utilizing such a tool that is really pushing the boundaries of traditional long-term planning.

Ultimately, the results of these collaborations and the refinement of this capability will benefit state energy offices, utility commissions and state and federal lawmakers by providing multisectoral insights to help them establish informed policy and aggressive, but achievable decarbonization goals and targets. These insights might include a better understanding of the inter- and intra-state regional challenges that could be realized under different decarbonization pathways or an evaluation of the effectiveness of different electrification policies, carbon pricing regimes, or even environmental policy that impacts industrial sectors.

10.0 Long-Term Vision

By establishing the capability to downscale GCAM-USA output to PCM inputs, high impact analytical capabilities will be unlocked for PNNL. By enabling the ability to extract hourly snapshots of system dispatch and flow from a PCM, advanced reliability, and resiliency studies can be performed to better understand the physical constraints and challenges with operating under conditions that are vastly different from today's operation. For example, these advanced studies will allow PNNL to support end-users in identifying regional tipping points at which renewables and technology adoption will cause significant reliability and stability concerns under various decarbonization pathways.

Additionally, the methodology to link GCAM and GridView in this LDRD will be developed in such a way for future integration with detailed distribution system modeling tools (such as GridLAB-D) for hosting capacity analysis and distribution planning.

These examples are portrayed in the simulation framework (Figure 5.1), a subset of which is shown in Figure 10.1 below.

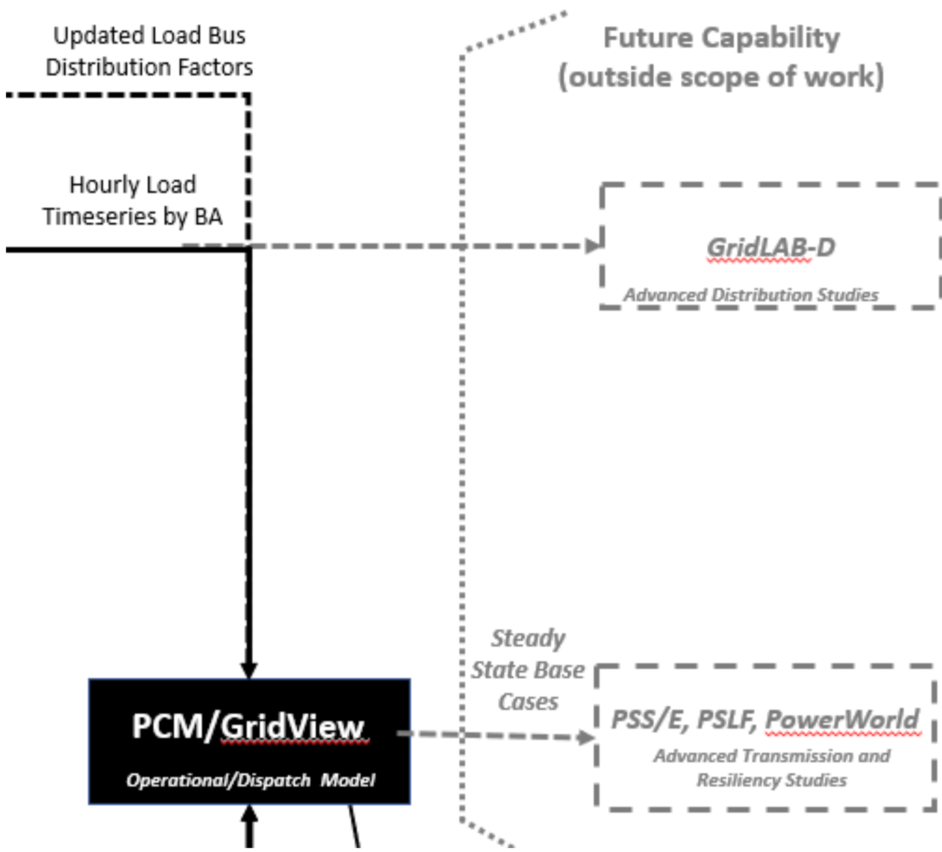

Figure 10.1. Subset of the larger framework in Figure 5.1 illustrating future capabilities that are enabled after accomplishing successful downscaling and production cost modeling simulations of decarbonization pathways

11.0 Conclusion

This report demonstrates the initial effort in documenting the scope of work for the LDRD titled: "Integrated Assessment Modeling of Grid Resilience: GCAM-to-PCM Scoping." IM3 tools are to be leveraged and enhanced to downscale GCAM-USA outputs to the level of detail needed for high-fidelity/industry-grade PCM simulations. This downscaling will be demonstrated on the WECC footprint using WECC GridView PCM as base models. A simulation framework, proposed project schedule, and the minimum viable product in establishing downscaling methodologies were addressed in this report.

12.0 References

"Alternative Fuels Data Center: Electric Vehicle Infrastructure Projection Tool (EVI-Pro) Lite Assumptions and Methodology." n.d. Accessed February 18, 2022. https://afdc.energy.gov/evi-pro-lite/load-profile/assumptions#section-3-1.

Datta, A., S. Banerjee, A.O. Finley, and A.E. Gelfand. 2016. "On Nearest-neighbor Gaussian Process Models for Massive Spatial Data." *WIREs Computational Statistics* 8: 162–71.

Draxl, Caroline, Andrew Clifton, Bri-Mathias Hodge, and Jim McCaa. 2015. "The Wind Integration National Dataset (WIND) Toolkit." *Applied Energy* 151 (August): 355–66. https://doi.org/10.1016/j.apenergy.2015.03.121.

Draxl, Caroline, and Bri Mathias-Hodge. 2016. "WIND Toolkit Power Data Site Index." National Renewable Energy Laboratory - Data (NREL-DATA), Golden, CO (United States); National Renewable Energy Laboratory. https://doi.org/10.7799/1329290.

"Eastern and Western Data Sets." n.d. Accessed February 18, 2022. https://www.nrel.gov/grid/eastern-western-wind-data.html.

"Fleet DNA: Commercial Fleet Vehicle Operating Data." n.d. Accessed February 18, 2022. https://www.nrel.gov/transportation/fleettest-fleet-dna.html.

Friedman, Jerome H. 2001. "Greedy Function Approximation: A Gradient Boosting Machine." *The Annals of Statistics* 29 (5): 1189–1232. https://doi.org/10.1214/aos/1013203451.

Gelfand, Alan E., and Erin M. Schliep. 2016. "Spatial Statistics and Gaussian Processes: A Beautiful Marriage." *Spatial Statistics*, Spatial Statistics Avignon: Emerging Patterns, 18 (November): 86–104. https://doi.org/10.1016/j.spasta.2016.03.006.

"Global Change Intersectoral Modeling System." GCIMS, 2023. https://gcims.pnnl.gov/global-change-intersectoral-modeling-system

"GridView Brochure." ABB, 2023. https://library.e.abb.com/public/6676a5b70f5b4e148c03b74f77056318/GridView-brochure-9AKK106930A8192.pdf?x-sign=Su4cZgpFEFNfBltLZfDqzYrAEp7AUvJfbjtjFPZ0WJMSjOFjGPoBbGGfsLiO19XW

Griesbach, Colin, Benjamin Säfken, and Elisabeth Waldmann. 2021. "Gradient Boosting for Linear Mixed Models." *The International Journal of Biostatistics* 17 (2): 317–29. https://doi.org/10.1515/ijb-2020-0136.

"GridView." Hitachi, 2023. https://www.hitachienergy.com/us/en/products-and-solutions/energy-portfolio-management/enterprise/gridview

Holmgren, William F., Clifford W. Hansen, and Mark A. Mikofski. 2018. "Pvlib Python: A Python Package for Modeling Solar Energy Systems." *Journal of Open Source Software* 3 (29): 884. https://doi.org/10.21105/joss.00884.

"Home - System Advisor Model (SAM)." n.d. Accessed February 18, 2022. https://sam.nrel.gov/.

Hummon, Marissa, Eduardo Ibanez, Gregory Brinkman, and Debra Lew. 2012. "Sub-Hour Solar Data for Power System Modeling From Static Spatial Variability Analysis: Preprint." NREL/CP-6A20-56204. National Renewable Energy Lab. (NREL), Golden, CO (United States). https://www.osti.gov/biblio/1059579.

IAEA. uranium: resources, Production and demand. International Atomic Energy Agency, Vienna, Austria (2011)

"Integrated Multisector Multiscale Modeling." IM3, 2023. https://im3.pnnl.gov/.

Iyer, G., Ledna, C., Clarke, L. et al. Measuring progress from nationally determined contributions to mid-century strategies. Nature Clim Change 7, 871–874 (2017a). https://doi.org/10.1038/s41558-017-0005-9

Iyer G.C., L.E. Clarke, J.A. Edmonds, G.P. Kyle, C.M. Ledna, H.C. McJeon, and M.A. Wise. 2017b. GCAM-USA Analysis of U.S. Electric Power Sector Transitions. PNNL-26174. Richland, WA: Pacific Northwest National Laboratory.

Iyer, G. C., M. Brown, S. M. Cohen, J. Macknick, P. Patel, M. Wise, M. Binsted, and N. Voisin. 2019. “Improving Consistency among Models of Overlapping Scope in Multi-Sector Studies: The Case of Electricity Capacity Expansion Scenarios.” *Renewable and Sustainable Energy Reviews* 116 (December): 109416. https://doi.org/10.1016/j.rser.2019.109416.

Jones, Andrew D., Deeksha Rastogi, Pouya Vahmani, Alyssa Stansfield, Kevin Reed, Travis Thurber, Paul Ullrich, and Jennie S. Rice. 2022. “IM3/HyperFACETS Thermodynamic Global Warming (TGW) Simulation Datasets.” MultiSector Dynamics - Living, Intuitive, Value-adding, Environment. https://doi.org/10.57931/1885756.

King, J., A. Clifton, and B.-M. Hodge. 2014. “Validation of Power Output for the WIND Toolkit.” NREL/TP-5D00-61714. https://www.nrel.gov/docs/fy14osti/61714.pdf.

Kintner-Meyer, Michael, Sarah Davis, Siddharth Sridhar, Dhruv Bhatnagar, Shant Mahserejian, and Malini Ghosal. 2020. “Electric Vehicles at Scale -- Phase I Analysis: High EV Adoption IMpacts on the Western U.S. Power Grid.” PNNL-29894. https://www.pnnl.gov/sites/default/files/media/file/EV-AT-SCALE_1_IMPACTS_final.pdf.

Lee, Dong-Yeon, Nicholas Reinicke, Eric Wood, Yanbo Ge, and Erin Burrell. 2021. “Electric Vehicle Infrastructure Projection Tool (EVI-Pro).” NREL/PR-5400-77651. National Renewable Energy Lab. (NREL), Golden, CO (United States). https://www.osti.gov/biblio/1764904-electric-vehicle-infrastructure-projection-tool-evi-pro.

Lloyd, James Robert. 2014. “GEFCom2012 Hierarchical Load Forecasting: Gradient Boosting Machines and Gaussian Processes.” *International Journal of Forecasting* 30 (2): 369–74. https://doi.org/10.1016/j.ijforecast.2013.07.002.

Lopez, Anthony, Billy Roberts, Donna Heimiller, Nate Blair, and Gian Porro. 2012. “U.S. Renewable Energy Technical Potentials. A GIS-Based Analysis.” Office of Scientific and Technical Information (OSTI). https://doi.org/10.2172/1219777.

Miller, Patrick, Daniel McArtor, and Gitta Lubke. 2017. “A Gradient-Boosting Machine for Hierarchically Clustered Data.” *Multivariate Behavioral Research* 52 (1): 117. https://doi.org/10.1080/00273171.2016.1265433.

Persson, Caroline, Peder Bacher, Takahiro Shiga, and Henrik Madsen. 2017. “Multi-Site Solar Power Forecasting Using Gradient Boosted Regression Trees.” *Solar Energy* 150 (July): 423–36. https://doi.org/10.1016/j.solener.2017.04.066.

Rogner, H-H. 1997. “AN ASSESSMENT OF WORLD HYDROCARBON RESOURCES.” Annual Review of Energy and the Environment. Annual Reviews. https://doi.org/10.1146/annurev.energy.22.1.217.

Schneider, Erich A., and William C. Sailor. 2008. “Long-Term Uranium Supply Estimates.” Nuclear Technology. Informa UK Limited. https://doi.org/10.13182/nt08-a3963.

“Solar Power Data for Integration Studies.” n.d. Accessed February 18, 2022. https://www.nrel.gov/grid/solar-power-data.html.

“System Stability Planning Anchor Data Set (ADS).” n.d. Accessed February 18, 2022. https://www.wecc.org:443/SystemStabilityPlanning/Pages/AnchorDataSet.aspx.

“Table MV-1 - Highway Statistics 2017 - Policy | Federal Highway Administration.” n.d. Accessed February 18, 2022. https://www.fhwa.dot.gov/policyinformation/statistics/2017/mv1.cfm.

Vernon, Chris R., Nino Zuljevic, Jennie S. Rice, Timothy E. Seiple, Michael C. W. Kintner-Meyer, Nathalie Voisin, Ian P. Kraucunas, et al. 2018. “CERF – A Geospatial Model for Assessing Future Energy Production Technology Expansion Feasibility.” *Journal of Open Research Software* 6 (1): 20. https://doi.org/10.5334/jors.227.

Zhou, Yuyu, Patrick Luckow, Steven J. Smith, and Leon Clarke. 2012. “Evaluation of Global Onshore Wind Energy Potential and Generation Costs.” Environmental Science & Technology. American Chemical Society (ACS). https://doi.org/10.1021/es204706m.

Pacific Northwest National Laboratory

902 Battelle Boulevard
P.O. Box 999
Richland, WA 99354

1-888-375-PNNL (7665)

www.pnnl.gov