Generating Value from Detailed, Realistic Synthetic Electric Grids

Final Project Report

S-91

Power Systems Engineering Research Center
Empowering Minds to Engineer
the Future Electric Energy System

Generating Value from Detailed, Realistic Synthetic Electric Grids

Final Project Report

Project Team
Thomas Overbye, Project Leader
Katherine Davis
Texas A&M University

Bernie Lesieutre
Line Roald
University of Wisconsin-Madison

Graduate Students
Hanyue Li
Jessica Wert
Seri Kang
Shashwat Tripathi
Texas A&M University

Jonathan Snodgrass
Noah Rhodes
Sofia Taylor
Aditya Rangarajan
University of Wisconsin-Madison

PSERC Publication 22-03

October 2022

For information about this project, contact:

Thomas J. Overbye
Texas A&M University
Department of Electrical and Computer Engineering
308C, Wisenbaker Engineering Building
College Station, Texas 77843-3128
Phone: 979-458-5001
Email: overbye@tamu.edu

Power Systems Engineering Research Center

The Power Systems Engineering Research Center (PSERC) is a multi-university Center conducting research on challenges facing the electric power industry and educating the next generation of power engineers. More information about PSERC can be found at the Center's website: http://www.pserc.org.

For additional information, contact:

Power Systems Engineering Research Center
Arizona State University
527 Engineering Research Center
Tempe, Arizona 85287-5706
Phone: 480-965-1643
Fax: 480-727-2052

Acknowledgements

This is the final report for the Power Systems Engineering Research Council (PSERC) research project titled “Generating Value from Detailed, Realistic Synthetic Electric Grids”. We wish to express our appreciation for the support provided by PSERC’s industry advisors:

Harvey Scribner (SPP)
Jim Price (CAISO)
Bryan Palmintier (NREL)
Di Shi (GEIRI)
Evangelos Farantatos (EPRI)
Cho Wang (AEP)
Steven Judd (ISO-NE)
Melvin Schoech (CenterPoint)
Al Engelmann (ComEd)
Yazhou Jiang (GE)
Anil Jampala (GE)
Mahesh Morjaria (First Solar)
Patrick Panciatici (RTE)
Jianzhong Tong (PJM)
Baj Agrawal (APS)
Felicia Ruiz (MISO)
Casey Cathey (SPP)
Atena Darvishi (NYPA)
Paul Myrda (EPRI)

The authors and PSERC industry advisors would like to recognize the research teams at Texas A&M University (Tom Overbye, Kate Davis, Komal Shetye, Jonathan Snodgrass, Hanyue Li, Jess Wert, Seri Kang) and the University of Wisconsin-Madison (Bernie Lesieutre, Line Roald, Noah Rhodes, Sofia Taylor, Jonathan Snodgrass (graduated 2021)) for their research contributions to this project:

Executive Summary

The goal of this research project was to generate value from large-scale detailed and realistic synthetic electric grids. This was accomplished by completing four sub-tasks: 1. Developing customized grids, 2: Developing specific grid scenarios, 3: Exploring decision making with uncertainty, 4: Expanding the scope of synthetic grids for coupling with other infrastructures

Task 1: Developing customized grids

For Task 1 we focused on the development, visualization, and application of large-scale synthetic electric grids, with specific examples are provided in many of the papers given below in the publications section. One of the challenges with doing engineering studies (such as power flow, optimal power flow, contingency analysis and time-domain simulations) on these systems has been maintaining good engineering situational awareness (SA). To address these challenges, we created visualizations that quickly and dynamically reduced an 82,000 bus system to show the overall system flows and the minimum bus voltages. We have also created several new synthetic grids, in some cases extending the complexity of the grids (such as include transformer impedance correction tables). These grids are available at https://electricgrids.engr.tamu.edu/ with details on a new 27,000 bus grid.

Additionally, we explored the application of wide-area transmission grid operation visual storytelling. The gist of this approach is to apply scientific video visualization approaches that have been developing over the last decade in other fields to power system problems, with the synthetic grids used to demonstrate the approach.

Task 2: Developing specific grid scenarios

In task 2 we extended synthetic grid models to consider potential future scenarios. In particular, we used geolocated electric grid models to associate external data such as wildfire risk to electric grid models. Initially, we focused on the small IEEE RTS-GMLC test case (geo-located in California). However, to obtain more interesting and realistic test cases, we developed a test case based on the actual geographical footprint of the California transmission grid. The development of this test case involved gathering data from multiple publicly available sources, including data on the transmission corridors in California obtained from the California Energy Commission and total load data from CAISO. To create a synthetic topology grid based on the location of the transmission lines, we developed methods to identify logical connection points between lines and added simple substation topology. This synthetic grid topology based on the real transmission corridor data was supplemented with realistic, but not real data on generation, load and transmission line parameters.

The benefit of a geographically accurate test case with detailed transmission corridor data is the ability to intersect the grid data with other interesting data sources describing the environmental and social context of the grid. In our case, we again considered integration of wildfire risk data (further described in Task 3), and are currently looking into extending this to account for additional environmental hazards including flooding and sea level rise, as well socio-economic data on population resiliency to disasters.

Task 3: Exploring decision making with uncertainty

To demonstrate the usefulness of the geographically accurate grid scenarios developed in Task 2, we leveraged the same grid data towards the Task 3 objectives on decision making under uncertainty. Specifically, we considered how to integrate electric grid and wildfire data to optimize the response of electric grid operators to increased wildfire risk. Specifically, we developed a model that uses the synthetic grid test case based on realistic geographical locations of transmission lines in California along with data on wildfire risk across an entire wildfire season. We overlapped publicly available GIS data from the actual California grid with the Wind-Enhanced Fire Potential Index (WFPI) published daily by the US Geographical Survey and developed several different wildfire risk metrics for each line. We then solved an optimization problem to identify the lines that are most promising candidates for moving underground, i.e., which lines are likely to have the largest impact on risk reduction if they are moved under ground.

Further, the use of synthetic grids in our testbed allows us to explore user actions taken by different roles (i.e., operators) in a "sandbox" environment. Using our cyber-physical testbed, we explored ways to model and mimic operator actions, i.e., using deep reinforcement learning. We also used our cyber-physical testbed capabilities to help develop a sharable/remotely accessible platform for testing and comparison of these types of approaches for the community.

Task 4: Expanding the scope of synthetic grids for coupling with other infrastructures

For Task 4 we leveraged work done on a separate ARPA-E project (involving IAB partner NREL) to develop synthetic electric grids that combine both transmission and distribution models down to the parcel level with models of other infrastructures such as transportation now being included. One example examined the impact of electric vehicle (EV) charging on the grid. Both the Texas 2000-bus and 7000-bus synthetic networks with their detailed distribution network topology were used extensively for this. The charging load profile by the hour and location was generated using dynamic traffic assignment simulations of transportation networks, incorporating parameters such as time of day, trip characteristics, and drivers' range anxiety. This spatio-temporal EV load was then mapped to distribution or transmission nodes (as available) in the synthetic power grid network by determining which substation service areas the transportation nodes lie in using graph-theory techniques such as Voronoi tessellation. Detailed EV profiles and studies have been completed for the Austin and Houston regions in Texas using passenger, light duty vehicle (LDV) models. Charging impacts examined system/equipment loading and generator emissions. Our research has also shown the importance of using large, regional, transmission system models in performing multi-city EV load grid impact studies from the point of view of generation dispatch and emissions.

Publications Fully or Partially Supported by the Project:

All publications are available online at either
Tom Overbye's website, https://overbye.engr.tamu.edu/publications/
Kate Davis's website, https://katedavis.engr.tamu.edu/publications/
Arxiv (Line Roald) https://arxiv.org/search/eess?searchtype=author&query=Roald%2C+L
Or by emailing the authors

[1] N. Rhodes, L.A. Roald, “Balancing Wildfire Risk and Power Outages Through Optimized Power Shut-Offs”, IEEE Transactions on Power Systems, vol. 36, July 2021, pp. 3118-3128

[2] S. Taylor, and L. A. Roald, “A framework for risk assessment and optimal line upgrade selection to mitigate wildfire risk,” Electric Power Systems Research, vol. 213, pp. 108592, 2022.

[3] N. Rhodes, and L. Roald, “Co-optimization of power line shutoff and restoration for electric grids under high wildfire ignition risk,” arXiv preprint https://arxiv.org/abs/2204.02507, 2022.

[4] J.L. Wert, K.S. Shetye, H. Li, J.H. Yeo, X. Xu, A. Meitiv, Y. Xu, T.J. Overbye, “Coupled Infrastructure Simulation of Electric Grid and Transportation Networks,” 2021 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT), February 2021, Washington DC, USA.

[5] T.J. Overbye, J.L. Wert, K.S. Shetye, F. Safdarian, A.B. Birchfield, “The Use of Geographic Data Views to Help With Wide-Area Electric Grid Situational Awareness,” 2021 IEEE Texas Power and Energy Conference, College Station, TX, Feb. 2021.

[6] T.J. Overbye, K.S. Shetye, J.L. Wert, W. Trinh, and A. Birchfield, “Techniques for Maintaining Situational Awareness During Large-Scale Electric Grid Simulations,” IEEE Power and Energy Conference at Illinois (PECI), Champaign, IL, April 2021.

[7] K.S. Shetye, T.J. Overbye, H. Li, and J. Thekkemathiote, “Considerations for Interconnection of Large Power Grid Networks,” PECI, Champaign, IL, April 2021.

[8] C. Haseltine and L. Roald, “The Effect of Blocking Automatic Reclosing on Wildfire Risk and Outage Times”, 52^{nd} North American Power Symposium (NAPS), Tempe, AZ, April 2021.

[9] Y. Liu, M. Gaskamp, Z. Mao, D. Wallison, K. Shetye, K. Davis and T.J. Overbye, “Evaluation of Performance Metrics for Electric Grid Operational Scenarios,” 52^{nd} NAPS, Tempe, AZ, April 2021.

[10] W. Trinh, Z. Mao, T. J. Overbye, J. D. Weber, and D. J. Morrow, “Considerations in the Initialization of Power Flow Solutions from Dynamic Simulation Snapshots”, 52^{nd} NAPS, Tempe AZ, April 2021.

[11] D. Wallison, M. Gaskamp, Z. Mao, Y. Liu, K. S. Shetye, and T. Overbye, “Design Considerations for Operational Power Systems Scenarios,” 52^{nd} NAPS, Tempe, AZ, April 2021.

[12] B. Allison, T. Overbye and J. Weber, “Improved Generator Voltage Control in Power Flow Solutions”, 52^{nd} NAPS, Tempe, AZ, April 2021.

[13] P. Dehghanian, J.H. Yeo, J. Wert, H. Li, K. Shetye, and T. Overbye, “Application of Transformer Impedance Correction Tables in Power Flow Studies,” 52^{nd} NAPS, Tempe, AZ, April 2021.

[14] T.J. Overbye, J. Wert, K.S. Shetye, F. Safdarian, A.B. Birchfield, “Delaunay Triangulation Based Wide-Area Visualization of Electric Transmission Grids,” Kansas Power and Energy Conference, Manhattan, KS, April 2021.

[15] F. Safdarian, A. B. Birchfield, K.S. Shetye, "'Additional Insights in Creating Large-Scale, High Quality Synthetic Grids: A Case Study," Kansas Power and Energy Conference, Manhattan, KS, April 2021.

[16] A. Birchfield, T. J. Overbye, "Graph Crossings in Electric Transmission Grids," 2021 North American Power Symposium, College Station, TX, November 2021.

[17] K. Shetye, H. Li, J. Wert, X. Xu, A. Meitiv, Y. Xu and T. J. Overbye, "Generation Dispatch and Power Grid Emission Impacts of Transportation Electrification," 2021 North American Power Symposium, College Station, TX, November 2021.

[18] K. Shetye, B. Yancey and T. J .Overbye, "Benchmarking Standard Library and User Defined Models of Renewable Generation," 2021 North American Power Symposium, College Station, TX, November 2021.

[19] I. Idehen, K. Shetye, T.J. Overbye and D. Morrow, "Large-Scale Electric System Research Center at Texas A&M – Description and Associated Research Activities," 2021 North American Power Symposium, College Station, TX, November 2021.

[20] T.J. Overbye, K. Shetye, J. Wert, H. Li, C. Cathey, H. Scribner, "Stability Considerations for a Synchronous Interconnection of the North American Eastern and Western Electric Grids," Proc. 55th Hawaii International Conference on System Sciences (HICSS), virtual, January 2022.

[21] W. Trinh, S. Kunkolienkar, Y. Liu, T. J. Overbye, "On the Sensitivity of Ring-Down Observed Inter-Area Modes in Large-Scale Electric Grids Using a Simulation Approach," IEEE. Power and Energy Conference at Illinois (PECI), Champaign, IL, March 2022.

[22] L. Liu, T.J. Overbye, W. Wang, X. Fang, J. Wang, H. Cui, F. Li, "Transmission-Distribution Dynamic Co-simulation of Electric Vehicles Providing Grid Frequency Response," IEEE Power & Energy Society General Meeting (PESGM), Denver, CO, July 2022

[23] T. J. Overbye, J. Snodgrass, A. Birchfield, M. Stevens, "Towards Developing Implementable High Altitude Electromagnetic Pulse E3 Mitigation Strategies for Large-Scale Electric Grids," IEEE. Texas Power and Energy Conference (TPEC), College Station, TX, March 2022.

[24] J. Snodgrass, S. Kunkolienkar, U. Habiba, Y. Liu, M. Stevens, F. Safdarian, T. Overbye, R. Korab, "Case Study of Enhancing the MATPOWER Polish Electric Grid," IEEE Texas Power and Energy Conference, College Station, TX, College Station, TX, March 2022.

[25] J.L. Wert, J. Yeo, F. Safdarian, T.J. Overbye, "Situational Awareness for Reactive Power Management in Large-Scale Electric Grids," IEEE Texas Power and Energy Conference (TPEC), College Station, TX, College Station, TX, March 2022.

[26] J. Wert, F. Safdarian, T.J. Overbye, D.J. Morrow, "Case Study on Design Considerations for Wide-Area Transmission Grid Operation Visual Storytelling," IEEE Kansas Power and Energy Conference (KPEC), virtual, April 2022.

[27] S. Tripathi, "A study of cyber-attack resilience in a DER-integrated synthetic grid based on industry standards and practices," MS, Texas A&M, May 2022 (this thesis isn't currently available online, but IAB members can obtain it by emailing Kate Davis at katedavis@tamu.edu).

Table of Contents

List of Figures

List of Tables

1. Introduction

1.1 The Motivation for the Creation and Utilization of Synthetic Power System Networks

A gap between academia and industry in the field of power electrical engineering comes from the fact that data of actual power systems are protected and access to those are restricted even for research and innovation purposes. In the field of power electrical engineering there is often a lack of access for many researchers to some information concerning the actual electric grid and its associated data. For example, the US power flow cases and several structural information of the actual grid are considered to be critical energy/electricity infrastructure information (CEII) [3] with restricted availability. However, in order to improve modern power system models, operation and planning optimization problems such as power flow, economic dispatching, unit commitment and generation expansion planning, complex electricity market models with emerging distributed energy resources, dynamics and transient stability studies, geomagnetic disturbance studies, and advanced algorithms, there is a strong need for access to diverse, large and complicated power systems that are available for research and publications.

Several IEEE test cases are established and widely used to represent a portion of the American Electric Power System (in the Midwestern US) [4]. A test system proposed in [5] is based on structural attributes and data from the ISO - New England. Reference [6] develops an approximate model of the European interconnected system using actual transmission networks to study the effects of cross-border trades. However, until recently, there was limited work focusing on the creation of complicated and realistic synthetic large-scale power system models using publicly available data that can mimic the full complexity of modern electricity grids for more accurate power system studies.

The primary method for designing synthetic system models is to create them "from scratch", i.e. constructing a greenfield power system network model. This is computationally difficult, since it is equivalent to running a transmission expansion study with at least $1.4n$ candidate transmission paths, where n is the number of buses in the network. In reality, the number of candidate paths is much higher, up to $\frac{n^2}{2}$. Our previous work [7–10] explained algorithms for tractably constructing electric grid network models. The methods use census data [11, 12] and U.S. Energy Information Association (EIA) generation data [13], which span on the actual geographic footprints and provided realistic test cases for power system studies without revealing any sensitive information. Additional complexities can still be added into synthetic models to extend their applications.

1.2 Final Report Organization

The goal of this PSERC final report is to build on the research teams' prior experience to construct new synthetic power system network models and make improvements to existing synthetic grids.

These efforts are broken down into the 4 tasks detailed in this report. Task 1 (Chapter 2) focuses on developing customized grids and presents the creation of a new 27,000 bus network, improvements to one of the MATPOWER Polish network cases, development of new visualization techniques to maintain situational awareness and for wide-area visualization, and the results of a study that examined the synchronous interconnection of the Eastern and Western United States electric grids. In Chapter 3, we developed specific grid scenarios by creating a geographically accurate synthetic transmission grid over the geographic footprint of California. Task 3 in Chapter 4 explores decision making with uncertainty by first creating a framework for assessing the wildfire risk of a transmission line and determining optimal sets of lines to move underground. Additionally, we performed dynamic transmission and distribution co-simulations examining electric vehicles providing grid frequency support, created operational power system simulation scenarios, and studied the cyber attack resilience of a system with distributed energy resources. Finally, Chapter 5 improved existing synthetic grid models by coupling them to other infrastructures. Specifically, we examined the impacts of electric vehicles on generation dispatch and grid-related emissions.

2. Task 1: Developing Customized Grids

Task 1 is focused on developing customized grids to be used for simulations, research, and education. This was accomplished by creating a new 27,000 bus synthetic network geo-located in central United States [14] (Section 2.1). Additionally, one of the Polish grid Matpower cases was updated to include approximate geographic coordinates and synthetic dynamic models as explained in Section 2.2 [15]. New visualizations techniques were developed for maintaining situational awareness during power grid simulations [16] (Section 2.3). Additional techniques were developed and implemented for visualizing wide-area grid behavior and simulations as described in Section 2.4 [17]. Finally, these visualization techniques were used in a study that examined the dynamic behavior of a unified, synchronously connected Eastern and Western United States power grid [18] (Section 2.5).

2.1 A Large 27,000 Bus Network Example

This section describes a 27,000 bus synthetic network that has been developed to generate a fictitious but realistic power system model based on the actual generation data without revealing any protected information. The network is statistically similar to the actual U.S. power system on the Midwest U.S. footprint with capability to represent characteristic features of actual power grids. This synthetic network model is available at [19] and can be shared freely for teaching, training, and research purposes.

2.1.1 Methodology: Creating Synthetic Grid Models

The synthetic grid models are created using metrics derived from the North American Eastern Interconnect (EI) and publicly available data, provided by the U.S. Census and Energy Information Administration. Reference [8] outlines fundamental steps for the creation of synthetic power system models including geographic load, generator substations, and assignment of transmission lines. The overall approach for building these networks is summarized below, and described more in [8].

Substation Planning

This step includes locating and sizing load and generator substations using public data and statistics derived from the U.S. electric grids. Considering humans as the primary consumers of electricity, population data in geographic latitude and longitude coordinates is the main base of synthetic loads. Additionally, publicly available generator data from the U.S. EIA, which includes data for all generators in the U.S., is used to site and size generators in the synthetic case.

Then, a clustering technique is employed, which ensures the synthetic substations meet realistic proportions of load and generation, among other constraints. Within each substation, loads are usually connected to the lowest voltage level, and generators are often connected to the highest

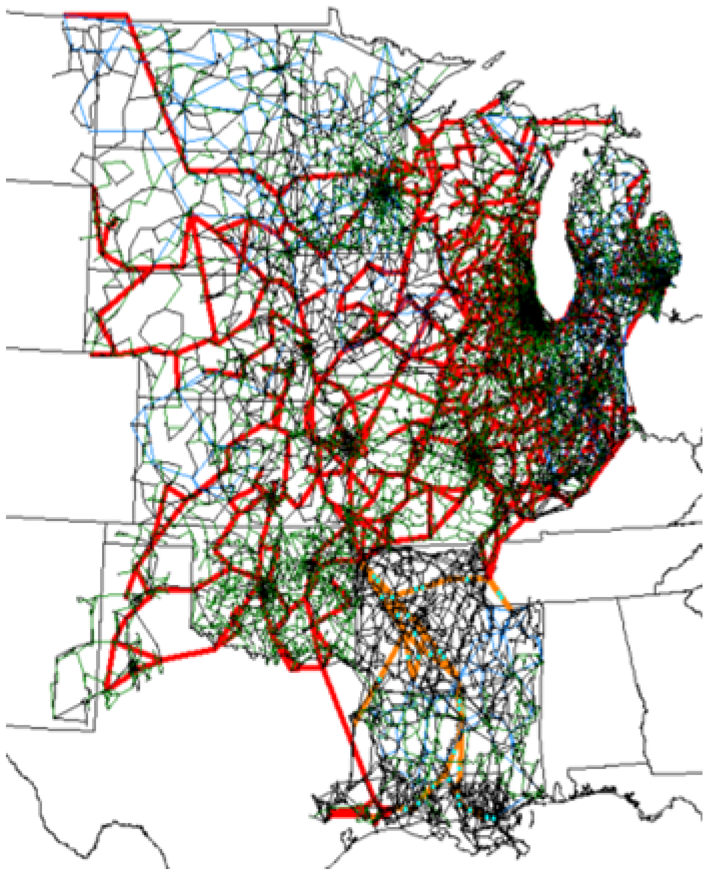


Figure 2.1: One-line Diagram of the 27K-Bus Case

voltage level through a generator step up (GSU) transformer. In addition, transformers are added in each substation to connect multiple nominal voltage levels.

Transmission Planning

This step is the most challenging and computationally expensive step. First, transmission line electrical parameters are calculated based on the assigned voltage levels and the percentages of substations with each voltage level as well as line length, which are determined based on the distance between substations. Conventions employed by grid planners are considered as metrics to make the parameter selections more realistic. Reference [9] presents a methodology to generate synthetic line topologies with realistic parameters. This step includes several structural statistics to be used in characterizing real power system networks, including connectivity, Delaunay triangulation overlap, DC power flow analysis, and line intersection rate and considers N-1 contingencies to improve the system reliability.

Reactive Power Planning

At this stage, AC power flow solvable synthetic cases are created, with varying network sizes and complexities. Then, the test cases are augmented with additional complexities like voltage control devices such as switched shunts.

Key Considerations and Challenges

Key challenges include geography constraints such as lakes, mountains, and urban areas, as well as network topology parameters, increasing power flow feasibility in base and N-1 contingency conditions, intractability of n^2 possible combinations of branches (where n is the number of buses), many competing metrics to meet, and consideration of contingency conditions that increases computation even more. References [20, 21] present some metrics for validating synthetic grids for achieving realistic data sets.

2.1.2 Case Description and Visualizations for the 27,000 Bus Network

Table 2.1: A Summary of the Statistics for the 27,000 Bus Case

Metric or Statistic	Quantity
Number of substations	13,074
Number of buses	27,163
Number of areas	19
Number of transmission lines	28,550
Number of transformers	10,651
Number of loads	14,054
Number of generators	4,224
Number of shunts	1,961
Total design load (GW)	154 GW

A synthetic grid is created in the US Midwest footprint mainly on Midcontinent Independent System Operator (MISO) and Southwest Power Pool (SPP) coverage area based on the method that is explained in the previous section. All simulations are carried out using PowerWorld [22] on an Intel(R) Core(TM) i7 2.59 GHz laptop with 32GB of RAM. The case is built as a 27,163-bus case for power flow studies and general analysis and research purposes. The number of buses are estimated based on the size of the grid. The geography is complex and diverse in terms of vegetation and civilization, and includes 19 areas, divided by U.S. states. The transmission network is built with seven voltage levels: 500 kV, 345 kV, 230 kV, 161 kV, 138 kV, 115 kV and 69 kV. These voltage levels and percentages of substations including each voltage level is approximated from [23]. Figure 2.1 shows a one-line diagram of the case with 500 kV and 345 kV highlighted with thicker lines. Orange shows 500 kV lines and red shows 345 kV lines, blue shows 230 kV, black refers to 161 kV, 138 kV and 115 kV lines and green shows 69 kV lines. Figure 2.2 shows the nominal voltage of transmission lines over the number of buses, and Table 2.1 gives a summary of the case.

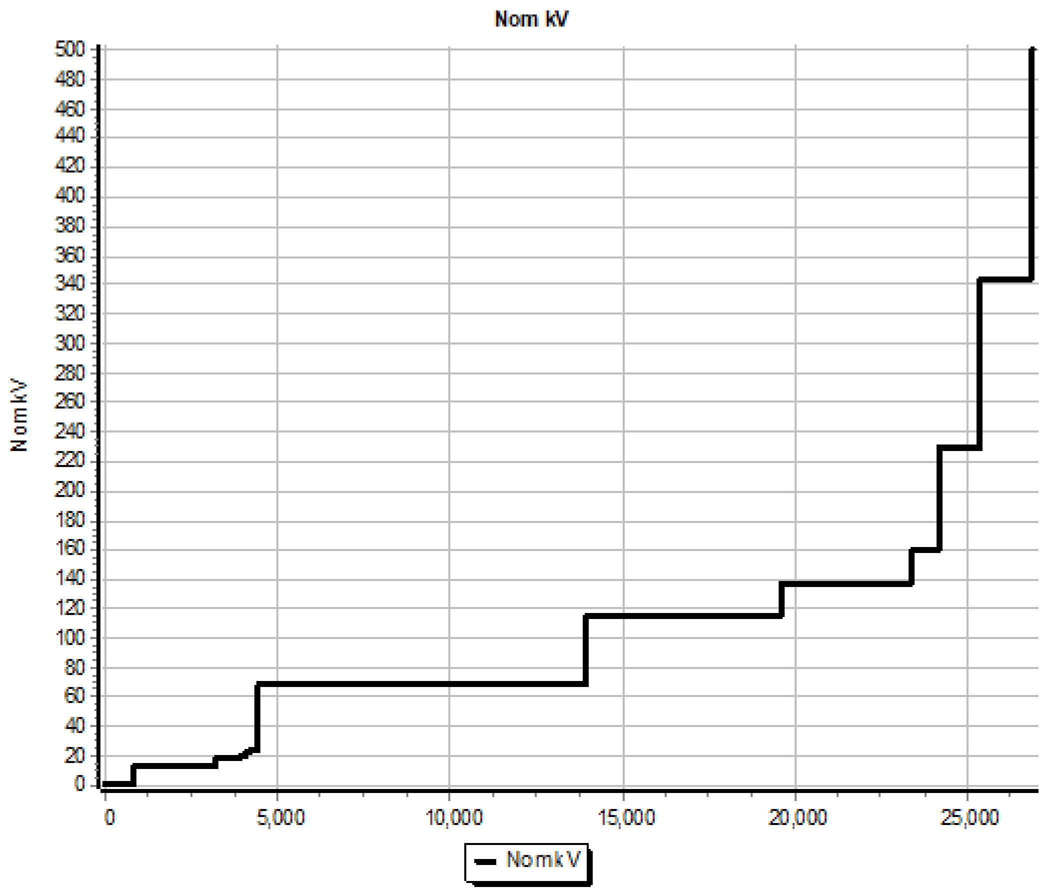


Figure 2.2: Nominal voltage of transmission lines over the number of buses of the 27K-bus case

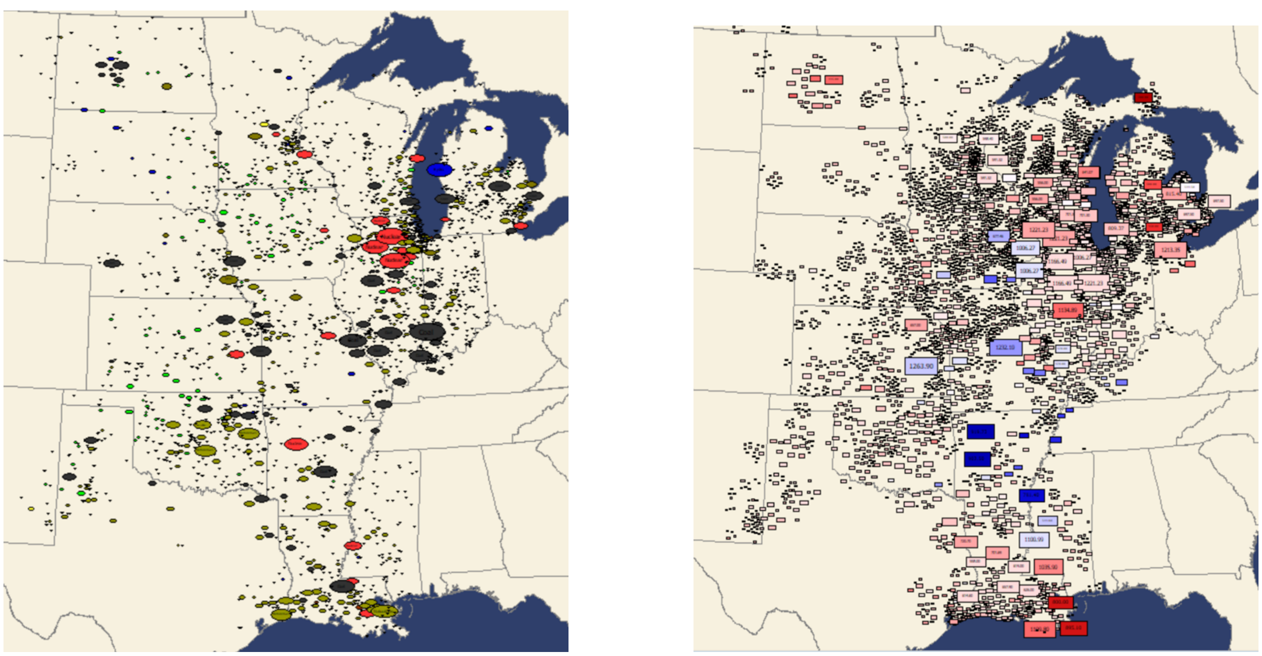
The types and number of generators in the case is extracted from EIA-860. All generator with a capacity larger than 8 MW are included in the case. Table 2.2 shows a summary of the number and overall capacity of generators grouped by fuel types.

Figure 2.3a shows the geographic data view (GDV) of generators. The size of ovals is proportional to the MW capacity of the units and the colors show the fuel types of generation substations where black refers to coal, brown to gas, green to wind, yellow to solar, dark blue to hydro, red to nuclear, dark magenta to petroleum and brown to other types.

Figure 2.3b shows the set points of generating units where the size of rectangles is proportional to the output active power in MW and color is proportional to output reactive power in MVAR. Dark red shows reactive power set points near maximum and dark blue shows reactive power of generators near minimum. As it can be observed from Figure 2.3b, most generators are not at their maximum or minimum MVAR limits. This gives the generators more freedom to regulate the voltages of nearby buses by injecting or consuming reactive power.

Table 2.2: Number and Type of Generators in 27k Bus Case

Fuel Type	**Number of Units**	**MW Capacity**
Natural Gas	1,268	128,513
Coal	187	65,296
Nuclear	27	27,106
Wind	427	46,398
Solar	30	735
Hydro	163	7684
Petroleum	101	3,917
Other	204	34,130
Total	2,407	313,779



(a) Generation substation MW capacity and fuel type

(b) Generator real and reactive power output

Figure 2.3: 27k-bus system generator displays

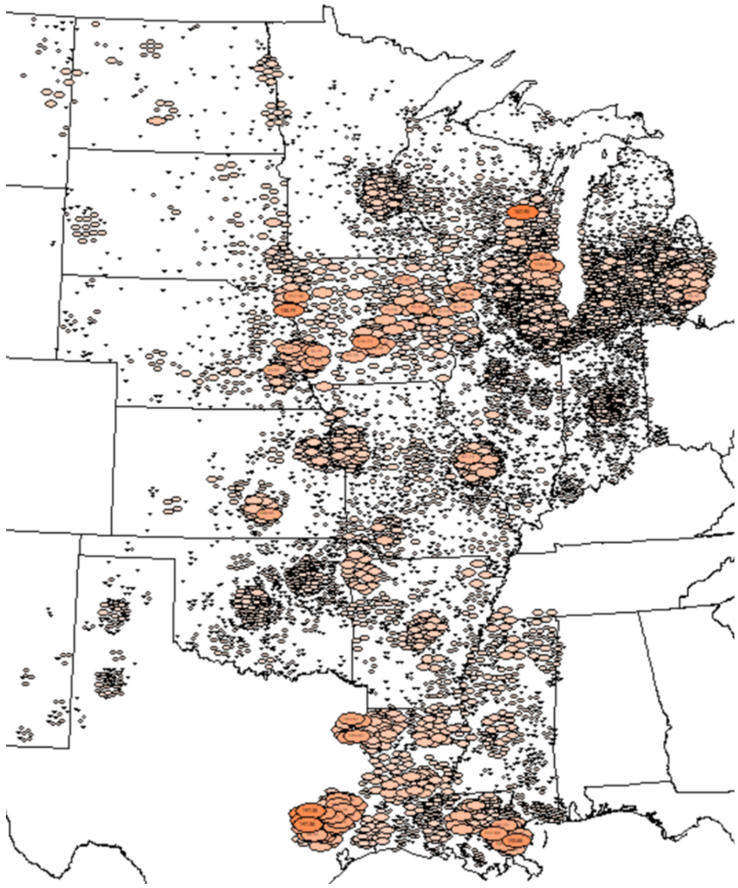


Figure 2.4: Load substations in the 27k-bus system

As explained in the previous section load is placed at every location, with a MW and MVAR amount proportional to the population. Statistics derived from the Eastern Interconnect are used to calculate values for the amount of load consumed per person. Also, industrial load is added to the locations with a high share of industrial load. Finally, a load substation is placed for every load and sized using the same MW and MVAR values for the load it is connected to. To complete the substation, a voltage level is assigned based on typical values for the region the substation resides in and its MW load. Additionally, a step-down transformer is assigned with electrical parameters that are based on median values for Eastern Interconnect transformers that share the same voltage level. Figure 2.4 shows load substations where their size is proportional to the load in MW. As it can be observed, the load size varies in different areas and the goal of this GDV is to match the load size and its variation to the actual cases.

Another important challenge with improving this case was to match the overall power flow patterns across the geographical footprint of U.S. Midwest. In order to have an organized view of the flows, GDV summary objects are used.

Figure 2.5 shows the net flows between areas. Size of arrows is proportional to the size of power

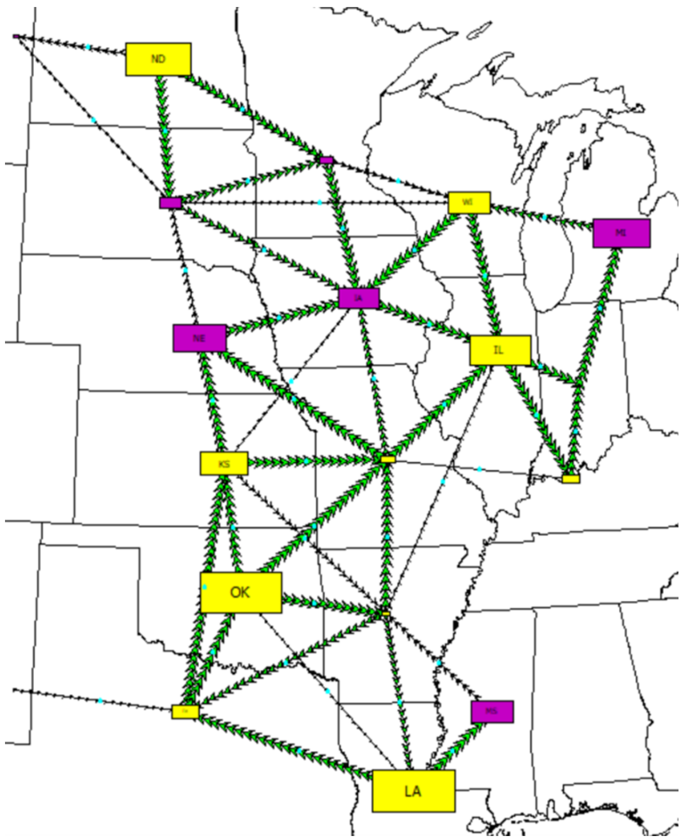


Figure 2.5: Inter-Area flows in the 27k Bus System

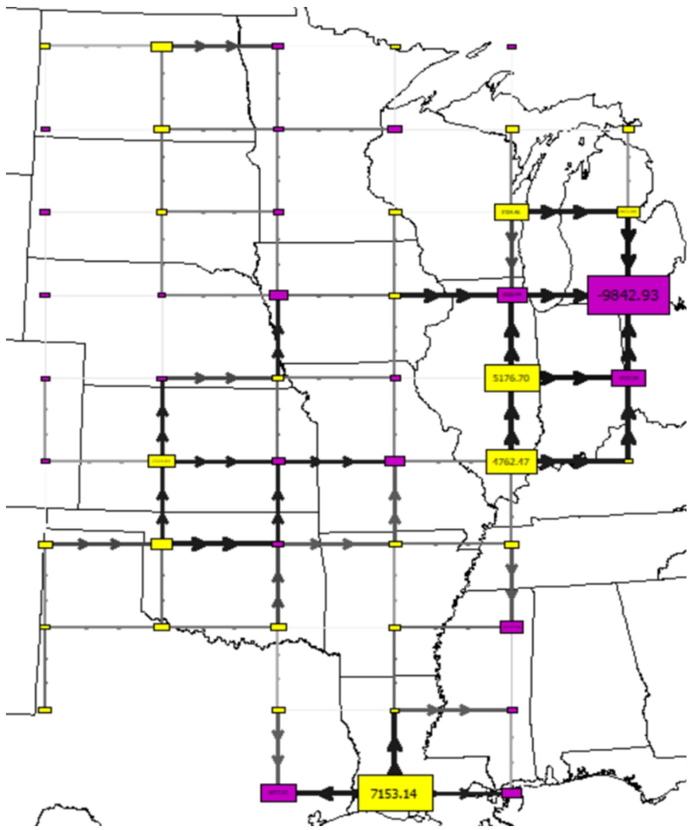


Figure 2.6: Grid-based visualization of 27k-bus system

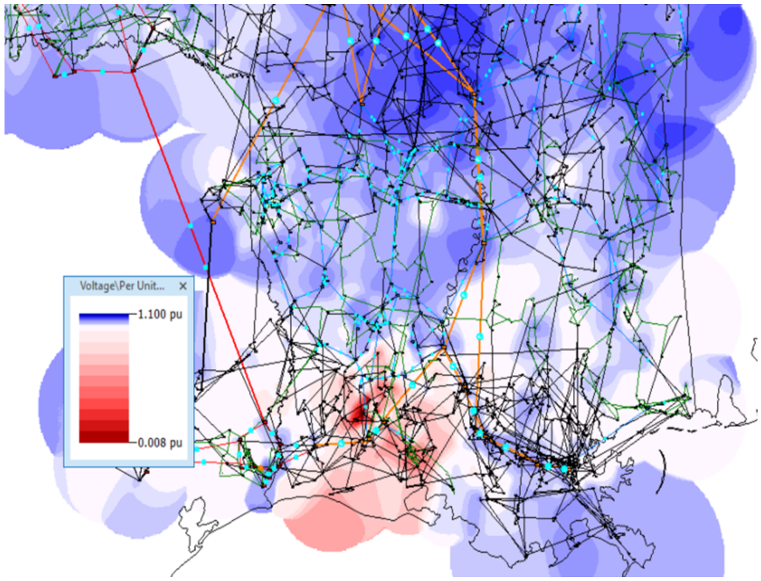


Figure 2.7: Bus voltage contour highlighting low voltage solutions

flow, and size of rectangles are proportional to the net injection to the areas in MW, where magenta refers to net import and yellow refers to net export. In order to have a more organized view of the flows, GDV summary objects are used. These GDVs group the electric grid objects geographically and show the summary GDVs based on a vertical and horizontal grid covering the entire system footprint. Such summaries could be used with actual values or with the previously mentioned differences between solutions. Figure 2.6 shows the overall flow pattern of electricity using a 6 by 10 grid of GDV summary objects on the 27K bus system. The size of each GDV is proportional to the net injection where magenta refers to net import and yellow refers to net export. The arrows show the direction of flow and the thicker and darker they are, the more active power flow between GDV summary objects.

An important challenge with adjusting load and flow patterns was to have a solvable AC power flow problem without any mismatches in the nodal power balance and avoid getting alternative solutions. In general, the power flow equations can have a normal solution, and potentially one or more alternative solutions. Alternative solutions (also called low-voltage solutions) may occur when initial state of voltage magnitudes of some buses are close to the other side of the P-V curve and this is not a desired solution [24, 25]. Figure 2.7 shows very low voltages in red.

The initial state of the system is very important on the feasible solution. If initial voltage levels are violated from lower than 0.9 per unit (PU) or higher than 1.1 PU, alternative solutions may occur. In order to avoid low voltage solution, the initial state can change to a flat start where all voltage magnitudes are changed to 1 PU and if any change need to be applied in the grid, such as increase or decrease in the load, it should be applied in small steps and voltage magnitude of buses should be regulated after each step with the use of reactive power control devices such as switched shunts.

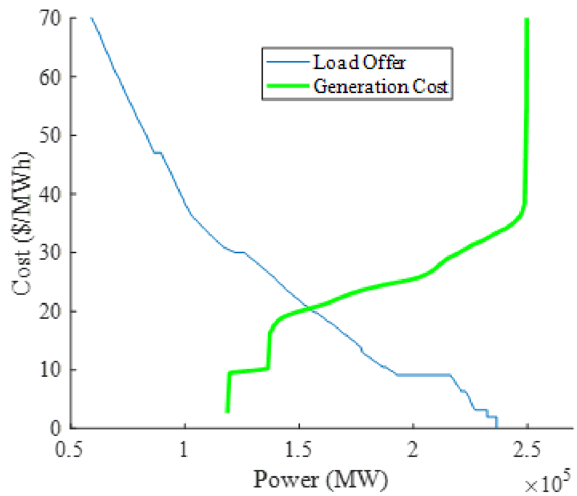


Figure 2.8: Super area incremental cost and marginal benefit curves of 27k-bus system.

Another development to the 27K grid is considering a price responsive demand model, as seen in Figure 2.8. The load benefit model is considered a piece-wise linear model, and each load entity has four offer steps. Up to 3-10% of the load is assumed to have high prices up to 3000 $/MWh which may be shed. From 3-10% to around 25-40% of each load has a price up to 180$/MWh. From 25-40% to around 50-75% of each load has a price up to 140 $/MWh and the last step has a price up to 10 $/MWh. Generators' cost is determined based on actual costs depending on their fuel types and locations. The Following curves in Figure 2.8 show the load marginal benefit curve and incremental cost curve of the whole area.

Table 2.3 shows some validation metrics of grid proportions, generators, load, and substations of the 27K bus case. Validation metrics are derived from the North American Eastern Interconnect and detailed in [21].

2.2 Modifications to a MATPOWER Polish Network

The Polish Systems consist of eight power system test cases made publicly available as part of the MATPOWER distribution [26]. The cases vary primarily in load levels and generation capacity based both on time of day and seasonal changes. These cases contain the required data for an AC Optimal Power Flow (ACOPF) study, namely bus numbers (without identifying names), network topology, branch parameters, generator cost functions, and necessary constraints: branch MVA, bus voltage, generator real and reactive power limits. The case chosen for updates, case2746wop, is representative of the Polish power system during the winter 2003-04 off-peak load level. Per [27], "Multiple centrally dispatchable generators at a bus have not been aggregated. Generators that are not centrally dispatchable in the Polish energy market are given a cost of zero." Before describing updates to one of the Polish test cases, a brief history is given that describes how these cases were originally created.

Table 2.3: Validation Metrics for the 27k Bus System

Validation Metric	Criteria	27K Bus Case
Buses per Substation	Mean 1.7-3.5	2.08
Substations in	<200 kV, 85-100%	99.80%
kV Range	>201 kV, 7-25%	13%
Substations with Load	75-90%	89.50%
Load per Bus	Mean 6-18 MW	6 MW
Load Power Factor	Mean 0.93-0.96	0.96
Generator Substations	5-25%	12.90%
Generator MW	25-200 MW, 40+%	36%
Maximum Capacities	200+ MW, 2-20%	8.50%
Committed Generators	60-80%	79%
Shunt Capacitors	10-25% of subs shunts	12%
and Reactors	30-50% above 200 kV	37%

2.2.1 Motivations for the Modifications to the Polish Systems

A salient component of engineering is the modeling of real-world systems as sets of mathematical equations. Using those models, engineers can analyze the systems and devise solutions to problems within those systems. Typically, it is not acceptable to experiment on in-service, critical systems; and there are few systems more critical than the world's electric power grids. Hence, power systems engineers are constantly working to construct and validate models of power system networks and their components including transmission lines, electrical loads, generators, governors, and other equipment.

This section details the modifications made to the mathematical model of the Polish electrical grid created by Dr. Korab for performing optimal power flow analysis of the Polish grid. The case2746wop case was used to perform the work for this section, and will herein be referred to as "the case" or "the Polish grid case". The cases provided with MATPOWER only included an OPF-level model of the system. Power flow and optimal power flow simulations are valuable tools in the study of power systems; however, they only provide steady state information of the system and lack modeling information for transient or dynamics studies, and geographic coordinates.

In real-world systems the dynamic behavior cannot be ignored; the critical necessity of any power grid mandates its stability to potential disturbances [28]. Contingencies that perturb a power system from its steady state occur frequently and must be analyzed to ensure system stability. While line and generator outages are primarily considered, addition and removal of large loads may affect overall system stability. Necessary transient stability studies of a power system require the inclusion of dynamic models in the power system case.

Table 2.4: Polish Grid Statistics

Number of buses	2746
Number of substations	1718
Number of areas	4
Number of zones	6
Number of transmission lines	3340
Number of transformers	174
Number of loads	1997
Number of generators	514
Number of LTC	171
Number of shunts	6
Number of phase shifters	1
Total design load (MW)	18962

Additionally, substation geographic information is required for analyses such as GIC studies [29, 30], the optimal placement of renewable resources [31], or examining the impact of severe weather or hazards such as wildfires on transmission networks [32]. By geo-locating the high voltage substations, the Polish grid case can be used to test improvements to algorithms such as these.

2.2.2 Review of Previous Improvements and Uses of the Polish Grids

Since its inception, the Polish grid has been used in a variety of different projects and studies, strengthening the case for its validity and further demonstrating the importance of the addition of dynamic models and approximate geographic coordinates. In [33], the Polish grid is used to test the design of an efficient wide-area measurement system and the observability of that system. Modifications were made to the grid in [34] where research was performed in devising a unit commitment algorithm; and there are many other references in the literature that cite the use of the other cases of the Polish grid included with MATPOWER.

Though limited, the case has also been subjected to dynamics studies. In 2018, Dr. Cortilla-Sanchez gave a presentation, that summarized the work of adding to the case among other things, GENSAL generator, IEEE Type 1 exciter, Type 2 governor models [35]. Multiple protective relays are modeled in this update to the Polish grid case as well. This research was regarding mitigation of cascading failures in a power grid. This presentation was given to the cascading failure working group at the IEEE PES General Meeting held in Portland, OR during a panel session.

2.2.3 Metrics, Comparison and Analysis of the Polish Grid Case

In this section the important characteristic of Polish grid is compared to the US grid. A summary of important characteristics of Polish grid are mentioned in Tables 2.4 and 2.5.

Table 2.5: Type and Number of Generators in Polish Grid

Fuel Type	Number of Units	MW Capacity
Natural Gas	2	215
Coal	144	25275
Hydro	19	1862
Unknown	341	3921
Other types such as biomass	8	471
Total	514	31744

Table 2.6: Polish Grid Statistics based on Criteria from US Grid Part 1

Validation Metrics	Criteria from US Grids	Polish Grid
Buses per substation	Mean 1.7-3.5	1.6
Percent of substations	<200 kV 85-100%	93.10%
containing buses in kV range	>200 kV 7-25%	6.90%
Substations with load	75-90%	90.90%
Load per bus	Mean 6-18 MW	6.9 MW
Generator capacity / load	1.2-1.6	1.3
Substations with generators	5-25%	1.68%
Generator capacities	25-200 MW 40+%	36.70%
	200+ MW 5-20%	48.30%
Committed generators	60-80%	69.40%
Generators dispatched >80%	50+%	27%
Generator maxQ/maxP	0.4-0.55 for >70%	33.30%

Several specific metrics based on important characteristics of the grids are extracted from US Eastern Interconnect (EI), Western Electricity Coordinating Council (WECC), and Electric Reliability Council of Texas (ERCOT) and introduced in [20, 21]. These metrics are originally created to validate synthetic grids. Texas A&M University (TAMU) synthetic grids are created over different footprints in the US with the incentive to create realistic publicly available grid data for research, teaching and training and are available in [19]. The strategy of creation of the TAMU synthetic grids is explained in [8]. Table 2.6 shows these metrics as the criteria from US grids and then show the statistics from Polish grid compared to these criteria.

From Table 2.6, it is observed that the following sets of metrics are similar between the US grid and the Polish grid case: ratio of buses per substation, load per generator capacity, percentage of substations containing buses, and the percentage of load substations. The average load per bus is in the range of 6 to 8 MW and is 6.9 MW in the Polish grid. However, most generators' statistics

Table 2.7: Polish Grid Statistics based on Criteria from US Grid Part 2

	US Criteria/ Voltage levels	400kV (US 500kV)	220kV (US 230kV)	110kV (US 115kV)
	Number of buses	59	157	2528
Transformer per-unit reactance on own base	80% within [0.05 0.2] pu	94.10%	99.20%	50%
Transformer X/R ratio and MVA limits by kV level	X/R 40% below median	90.20%	67.80%	0%
	MVA 40% below median	100%	100%	0%
	X/R 40% above median	9.80%	32.20%	100%
	MVA 40% above median	0%	0%	100%
	X/R 80% in 10-90 range	58.80%	99.20%	100%
	MVA 80% in 10-90 range	100%	100%	50%
Line per-unit per-distance reactance by kV level	70% within 10-90 range	4.40%	29%	20%
Line X/R ratio and MVA limit by kV level	X/R 70% in 10-90 range	77.80%	61.60%	69.60%
	MVA 70% in 10-90 range	0%	80.40%	72.70%
Lines/Substations by kV level	1.1-1.4	1.29	1.45	1.31

are different in the Polish grid compared to the US grid. For example, based on the US grids, it is expected that more than 70% of the maximum reactive power per maximum real power ratio is in range 0.4 to 0.55 but only around 33% of generators from Polish grid are in this range.

Also, it is observed that US transmission grids usually use 69kV, 115kV, 138kV, 161kV, 230kV, 345kV, 500kV, and 765kV but the Polish grid uses 110kV, 220kV and 400kV voltage levels for the transmission grid. Therefore, the statistics from Polish grid are mostly different compared to the metrics extracted from US grids based on different voltage levels. In Table 2.7, a comparison of Polish grid statistics based on the US metrics for the closest voltage level is given.

2.2.4 Summary of Updates to the Polish Network Model

As is typical of most publicly available power system test cases, the Polish grid test cases include only steady-state generator parameters, and the cases do not include substations with geographic coordinates. To improve the quality of the Polish system networks, geographic coordinates were added to the network, and the generator models were expanded to include parameters such as field voltage impedance, inertial and frictional constants and other parameters. The detailed explanation of these additions is found in [15]. By expanding the Polish systems to include these models and geographic coordinates, the Polish network models can be used for analyses such transient stability, GIC studies, transmission expansion, and other simulations involving geographically-based phenomena such as weather or extreme events.

2.3 A Summary of Techniques for Maintaining Situational Awareness During Large-Scale Electric Grid Simulations

The design and operation of large-scale electric grid require a variety of different engineering studies and simulations. Some of these are static, such as power flow, contingency analysis and security constrained optimal power flow. And some are dynamic, usually involving time-domain simulations to determine the behavior of the electric grid following some disturbance (contingency). In all of these it is important that the person doing the study or simulation understand what is going on. A term that can be used to convey this concept is situational awareness (SA). While defined informally as "knowing what's going on," a more formal definition is "the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future" [36, 37]. The term is now widely used in electric grid operations and increasing with engineering studies [38–41]. The focus of this section is on techniques to help with time-domain simulations of large-scale electric grids.

The SA challenges with these simulations depend upon the electric grid size, the complexity of its models, the simulation contingency scenario complexity, and the desired application. For example in many educational and some research simulations the grid size, model complexity, scenario complexity and desired application are similar to the 96-bus angular stability study presented in [42]; SA can usually be adequately maintained just using a graph or two (e.g., Figure 8 of [42] showing the rotor angles for the 20 generators). Similarly even with a large system with complex models and scenarios, if the goal is just to insure that the results for potentially thousands of different contingencies (perhaps run in parallel) meet some criterion (such as for voltage recovery as given in [43]) then likewise the SA needs would be modest.

In contrast the focus here is on improving SA associated with simulations in which there is a desire is to obtain a rather detailed understanding of the total system response. Example applications include doing simulations to insure all of the system models perform adequately, designing remedial action schemes [44], considering the impact of unusual events on the grid (with one example a high altitude electromagnetic pulse [45]), or a recent study by the authors considering an ac interconnection of the North American Eastern and Western grids in which the associated grid had 110,000 (110K) buses, there were 245 different types of device models and more than 46,000 model instances [46, 47].

Leveraging the authors' extensive experience in doing such simulations and in developing the associated software, this section and two subsections present a number of techniques specifically focused on improving SA for such studies. The first subsection focuses on SA during the initial simulation setup including the power flow, while the second is on SA during and after the simulation. Results are demonstrated on both a 10,000-bus (10K) synthetic grid [30, 48], and on the previously mentioned 110,000-bus model. While the presented techniques are generic, they are specifically demonstrated using PowerWorld Simulator version 22.

2.3.1 Situational Awareness While Setting up a Simulation

Over the years a number of techniques have been developed to help with SA including the use of onelines often at the substation level such as in Figure 2.1), tabular displays, intelligent alarming and color contouring (such as Figure 2.7; some background papers in this area include [49–51]. Our experience is that all of these techniques can be quite useful, with the most important design aspect being the ability to easily get more information on anything that seems important.

The focus here on wide-area visualization has been helped recently with more widespread availability of electric grid geographic information. This geographic information can be leveraged using geographic data views (GDVs) [41, 52] in which geographic information embedded in the electric grid model is used to draw symbols on a display in which the symbol's appearance can be dynamically modified to show model object values. GDVs can be quite useful for providing the "details on demand" mentioned earlier and will be utilized throughout the section.

Having good SA on the initial power flow solution is crucial to correctly initializing the simulations. However, before running the simulations it is important to address model instance parameter errors, a challenge common in larger-sized grid models. While errors on the individual instances do gradually get corrected, this is offset by the addition of new model types and associated instances. With the current rapid change in many electric grids worldwide with the addition of more renewable generation and storage this trend shows no signs of abating.

Two commonly used SA methods include the single machine, infinite bus (SMIB) eigenvalue analysis [53], and examining the log of any initial dynamic simulation limit violations while visualizing all of the parameters associated with a particular model type. These two methods are explained in more detail in [16].

2.3.2 Situational Awareness During and After a Simulation

The potential SA challenge during and/or after a simulation is with interpreting the large amount of data that could be generated. How much interpretation is needed depends on the application. Luckily for many simulations the SA challenge can be extremely modest. Here, the focus is on really knowing what is going on in a large grid simulation when something quite unexpected could be occurring. Examples could include debugging a new electric grid model to look for parameter errors, simulating more unusual situations, or even some of the more routine studies mentioned earlier when things don't go as planned. Hence there is a need for more sophisticated techniques.

A first step to gaining SA is knowing whether the simulation failed (and if so, when) along with noting the number and times of occurrences of the simulation generated events (SGEs). Sometimes the cause of a simulation failure is simple, such as forgetting to clear a fault, something that can be readily determined from a log. Other times it can be much more difficult to determine what occurred, requiring a much broader consideration of the results.

With this interpretation the techniques presented here are broadly divided into three groups. First, traditional time-varying graphs in which time is the x-axis parameter and the time-variation in the values of interest (i.e., the signals) is shown. Hence they show a portion of the results matrices but without much of the metadata (beyond perhaps a label). Second, ones in which the visualizations show the grid at a particular time point and animation loops are used to show the simulation response. This approach provides an opportunity for more fully showing the metadata, such as the geographic or electric location. Third, techniques that use algorithms to aggregate the overall system response with the machine learning approach of [54] and modal analysis examples.

In the first group a common technique for gaining some simulation SA is to setup a time plot of a small sample of results (signals), usually chosen from across the grid, to get a feel for the overall grid response. Examples of such plots are shown in [16]. Advantages of such figures include they are quick to draw, a key can be used to provide a label for each signal (e.g., mapping the signal to a particular bus), and they can show the complete time-variation for the signals. Disadvantages include 1) there is a lack of any spatial relationships between the signals, 2) since it is just a sample of the signals important results could be missed, and 3) it can be difficult include different types of signals such as voltage magnitudes with frequency (this could be done with multiple y-axes though with a similar limitation on the number of signals). Such graphs have a long history (e.g., [42]) and they certainly play an important role in gaining SA.

An alternative to plotting a small signal subset is to plot all of them. Given the current speed of plotting algorithms tens of thousands of signals can be quickly rendered. The advantage is that now no signals of the specified type are missed (and for this example it is clear that there is a sustained oscillation that will be considered later in the paper). Disadvantages include the loss of being able to see the individual signals, potentially longer rendering times, and because of overlap many of the signals are actually covered. An alternative for showing all of the signals is to plot the envelope of their response (i.e., the minimum and maximum at each time point).

The second general group of techniques is to show data at a particular point in time, and then utilize animation loops [55, 56] to show how the system changes with time. One commonly used technique is a geographic oneline diagram, often with an associated contour [**?**]. An example of this is shown for the 10K system in Figure 2.9 in which the contour is showing the bus voltage magnitude deviation at a simulation time of two seconds and alpha-blended is used to deemphasize the transmission grid. Hence it is showing a row of the voltage magnitude simulation results matrix with all the values shifted by the values in the first row (i.e., when simulation time is zero).

As an alternative approach, Figure 2.10 shows a GDV summary visualization [41] in which the substation and line information have been aggregated based on a geographic grid. Here the size of each rectangle (a GDV summary object) is proportional to the net real power injection for the buses in the rectangle, the color and field value give the minimum per unit voltage magnitude, and the size of the black flow arrows is proportional to the real power flow between the different regions. The contours can be combined to show other data results as well, such as contouring based on spatial variation in the bus frequency, using GDV summary visualizations showing the largest

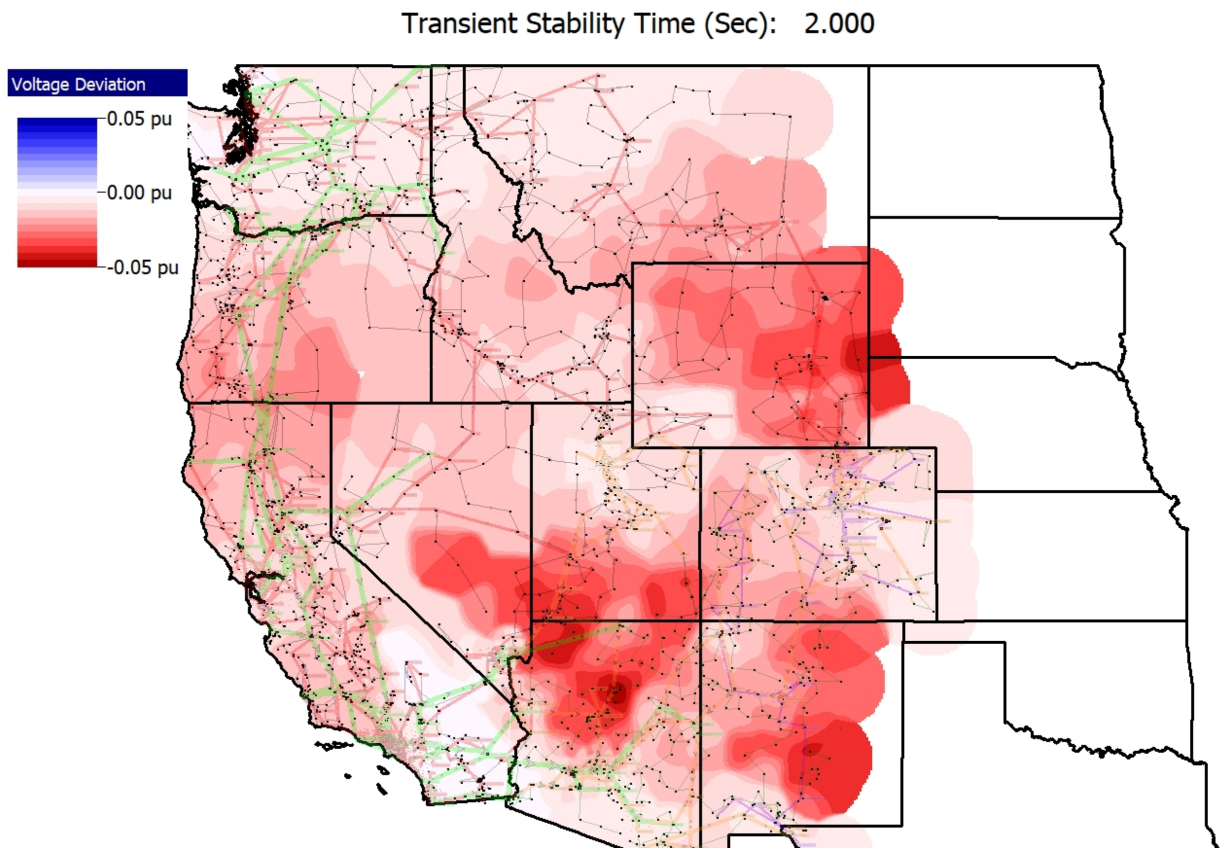


Figure 2.9: 10K Grid Voltage Deviation Contour at 2.0 Seconds

voltage deviation in the different geographic portions of the grid, and arrows showing the change in the real power flow, such as in Figure 2.10.

While by themselves such visualizations can help with SA, they are even more effective when used in animations. One animation approach that has been particularly effective in showing generator outage scenarios is develop the animations using a variable playback speed approach. For example, creating the animation to the first 10 seconds at ½ real-time, show the next 20 seconds are real-time, and show any subsequent values (i.e., the slower automatic generation control [AGC] response) at twice or more real-time.

The third general group of techniques for gaining SA is to utilize various algorithms to aggregate and summarize the overall system response. For the larger system applications here, the iterative matrix pencil (IMP) approach is particularly effective in determining the modes for large numbers of signals [57]. Modal analysis with the IMP utilizing all 10,000 voltage magnitude signals as inputs can be used to quickly determine is frequency (0.31 Hz) and the algorithm from [58] can determine its source and visualize the results. This is shown in Figure 2.11 in which the large magenta rectangle shows the source of the modal power flows in Northeast Montana, the yellow rectangles show the absorption locations (primarily in the southwest part of the grid) and the arrows show

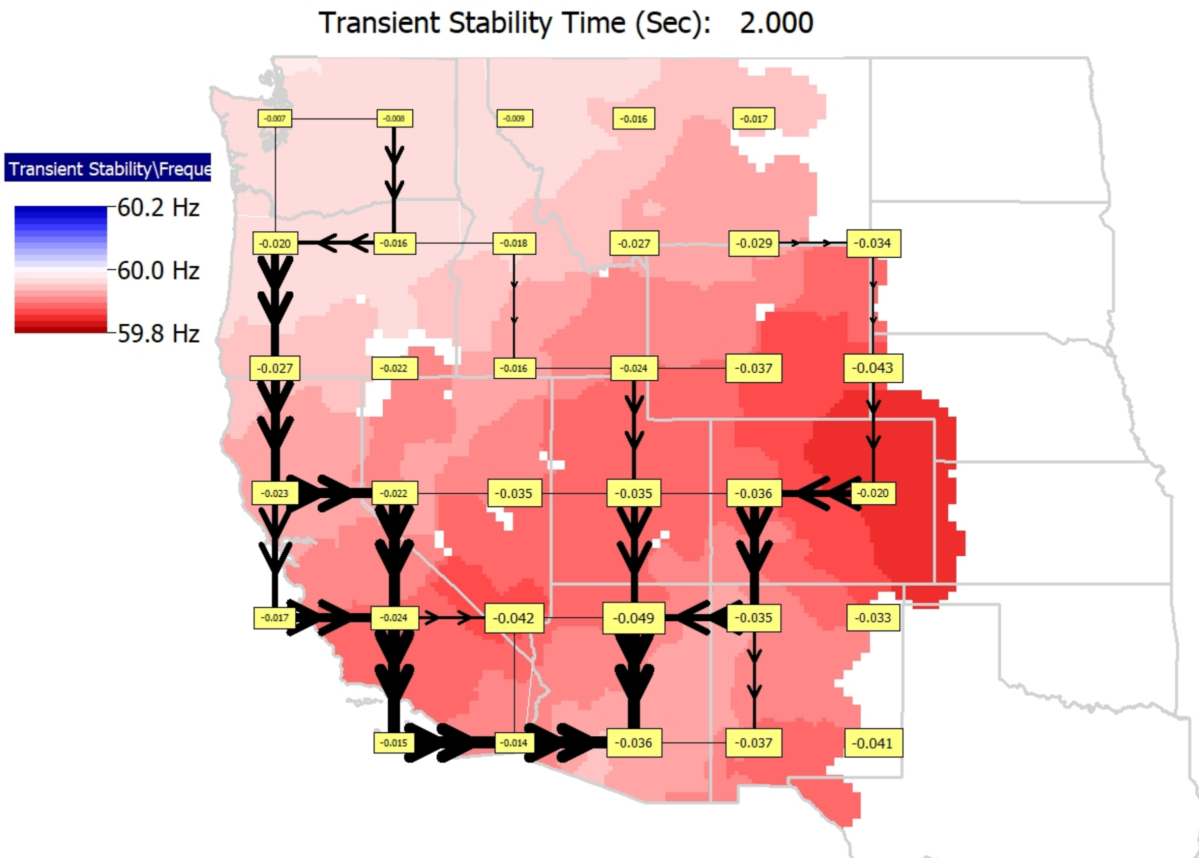


Figure 2.10: 10K Grid Voltage Deviation, Flows and Frequency at 2.0 Seconds

the modal power flows. The oscillation can be corrected by either disabling the dynamic model for the 37 MVA generator or correcting the model parameters (this error was actually deliberately induced to shown an oscillation, and could have been found with the SMIB since the generator had a positive eigenvalue).

Finally [16] provides example results from a 110K synthetic grid available at [**?**] for a generator loss contingency that includes modeling some of the AGC response.

2.4 Wide-Area Visualization of Electric Transmission Grids using the Delaunay Triangulation

Building on the situational awareness (SA) techniques in Section 2.3, the

The SA challenges associated with doing these studies depend upon a number of factors including the size of the electric grid, and, most importantly, the ultimate purpose for the user doing the study. To aid with SA, electric grid analysis software usually utilizes a number of different information presentation techniques including tabular displays and graphical visualization. These different approaches are often used synergistically to achieve overall SA. The purpose of this paper is to focus on an aspect of SA associated with the user understanding the overall state of a large-scale electric grid, with a particular focus on the graphical visualization of the overall transmission system flow patterns.

2.4.1 Power System Visualization Background

A quite common graphical visualization used with positive sequence simulations is the single-line or one-line diagram (oneline), so named because the devices associated with assumed balanced three-phase system are shown using a single line. Onelines have several different applications in electric grid operations and planning. In the context of this paper's focus on wide-area visualization of larger scale systems, the purpose of the oneline is to provide an overview of all, or a significant portion of an electric grid.

The section summarizes the work in [17] presents an algorithm to better visualize oneline information. The results are demonstrated on two example systems, a 37 bus grid from [59] and an 82,000 bus case from [30] from [**?**]. The full paper [17] presents the results on fours systems ranging from seven buses to 82,000 buses, and presents the overview onelines for these systems.

Over the years various techniques have been developed to help with wide-area visualization, particularly for oneline bus related data. One is color contouring [60], a technique that is widely used in the electric industry particularly for the display of locational marginal cost information [61, 62]. A second is the use of geographic data views (GDVs) [41, 52] in which geographic information embedded in the electric grid model is used to draw symbols on a display in which the symbol's appearance can be dynamically modified to show model object values. Our experience is that all of these techniques can be quite useful, with the most important design aspect being the ability to easily get more information on anything that seems important. To quote [63, 64] "Overview first, zoom and filter, then details on demand."

However it is much more difficult to do effective wide-area visualization of branch flows. This is partially due to the sheer number of branches, and partially due to the many overlapping branches that occur in large-scale grids with parallel branches common. This difficulty becomes more apparent when contrasting the smaller 37 bus system, with the larger 82K bus one; as systems grow it becomes more difficult to show specific details such as the individual bus voltage magnitudes or the specific outputs of generators. Some of this can be helped through zooming and panning, but zooming does come with the loss of seeing the overall system.

To address this the paper presents an algorithm to take the flows on a potentially large number of electric transmission lines, which could be either AC or DC, and visualize them using a planar graph. In general the high voltage transmission system is not planar, with line crossings common, particularly when the lines are at different voltage levels. Nevertheless the premise of this paper is a planar graph structure can be used to effectively visualize some aspects of such grids.

2.4.2 Delaunay Triangulation Visualization Algorithm

The algorithm to do this layout is straightforward and quite computationally tractable, allowing for interactive design even on large systems. As a starting point, assume that in the portion of the electric grid to be visualized there are n buses, m bus groups, and b branches. Further assume that each bus is included in a bus group and that each bus group has a unique associated spatial location, where the spatial location could be its geographic location (latitude and longitude) or an XY value associated with its position on a oneline.

The first step in the algorithm is to map each branch to the bus groups associated with its terminals. The second step is to do a Delaunay triangulation of the bus groups [65]. Computationally this can be $O(m\log m)$, and given that m is usually relatively small, this step is usually very fast. Define the set of line segments (segments) added as the result of the Delaunay algorithm as S, and of course by definition the Delaunay layout will be planar. The third step is to determine for each branch a segment path between its terminal bus groups, and then to add it to the list of branches maintained by each segment along the path. The last step is to visualize the results. This is shown in Figure 2.12. for the seven bus system utilizing dynamically size flow arrows and segments thicknesses [66,67]. More examples and a more detailed explanation of the process is given in [17]

This method can be utilized for larger systems, such as the 37-bus grid. In Figure 2.13 substations are shown using GDVs in which their size is proportional to the absolute value of the net injection for the substation (total generation minus total load) and the GDVs color is magenta if the net injection is positive and yellow if negative. Additionally, the figure includes blue reactive power flow arrows and labels on each segment to show the MW flow. In the figure the segment flows are again shown using arrows and the thickness of each line thicknesses proportional to the overall MW flow. In this figure the flows on 57 branches have been mapped into 34 segments, with at most three branches in a segment. For the branches 14 are within the bus groups (primarily the transformers) and hence not mapped to a segment, 36 are in a single segment, six in two segments and one in three segments. While the reduction from the number of branches to segments is rather modest here, it is easier to see the overall power flows. This approach could be particularly helpful in comparing different operating conditions.

As a final example the remaining figures demonstrate the approach using the 82K-bus, 76 area grid, which contains 104,125 branches spanning the contiguous US. The number of buses in each area ranges from 91 to 3229. Because of the large number of buses and branches little information about the grid's actual operation is apparent from its original oneline.

Figure 2.14 shows the initial Delaunay triangulation using the system's 76 areas as the bus groups. The size of each area GDV is proportional to the area's generation and its color based on the amount of MW exports (with red for net exports, and blue imports). The figure also shows the segments with a non-zero numbers of branches, with the number of branches in each segment ranging from one up to 161. Green arrows are used to visualize the MW flows.

2.5 Overview of Stability Considerations for a Synchronous Interconnection of the North American Eastern and Western Electric Grids

An application of the visualization techniques presented above is the analysis of connecting and synchronously operating the Eastern and Western United States Interconnections as described in [18]

This paper presents some of the dynamic considerations of an ac interconnection of the North American Eastern and Western electric grids. Currently most of the electricity used in North America (NA) is supplied by four major interconnects with each operating at 60 Hz but asynchronous with each other. These are the Eastern Interconnect (EI), the Western Electricity Coordinating Council (WECC), the Electric Reliability Council of Texas (ERCOT), and the Quebec Interconnection. All of these ac networks are internally synchronized and are linked to each other only through dc ties. For several years starting on February 7, 1967, the EI and WECC (then known as Western System Coordinating Council, with its name changed in 2002) were operated as a single electric grid, with the interconnection motivated by a desire to improve electric grid reliability as a result of the November 1965 Northeast Blackout. While this interconnection worked well initially, within months problems became apparent including oscillations on the western side and large inadvertent exchanges [68, 69]. This led to overloading of transmission facilities, major system breakups, reduced transmission capacity and a final removal of the ac interconnection in the early 1970's. Since then several back-to-back, high voltage dc (HVDC) facilities have been constructed along seam between the EI and WECC, allowing for up to 1.5 GW of east-west power transfer while the two grids operate asynchronously for each other. Over the years there have been several studies looking at a stronger connection between the EI and WECC, with some of this work focused on the economic or resource planning aspects [70, 71], and some on the use of HVDC for transmission expansion and design [72, 73]. In particular [73] focused on leveraging dc systems through upgrading the existing back-to-back (B2B) dc ties and building some new long distance HVDC lines. While this included rigorous analyses considering future capacity and carbon policies, a key area of improvement that this study did not consider is stability analysis. The feasibility of a new interconnection has been studied less frequently [74]. This is partially because as a result of the interconnection failure in the early 1970's an ac connection was viewed as, "like tying two elephants together with rubber bands; they can only go so far in opposite directions before the rubber bands snap" [75]. However, there has been significant growth in transmission and generation technologies since the 1970's along the EI WECC boundary, particularly with the now widespread application of power system stabilizers, and greatly improved electric grid monitoring and control. The need for more up-to-date assessments with improved stability models is identified in [74, 76]. Given the importance of system stability in operating these to grids synchronously, this paper [18] builds on and extends earlier results presented from a 2020 study on an ac interconnection of the EI and WECC in [46, 47]. PowerWorld Simulator Version 22 is used for all the simulation results shown due to in part to its ability to model very large systems, stability models used in both the EI and WECC grids and efficiently visualize engineering results. The paper is organized as follows. The second section presents the electric grid models used in the study. Given that some information about the actual electric grids models used in this study is designated as Critical Energy/Electricity

Infrastructure Information (CEII) [3] so it is not fully publicly available, the paper presents results for models both of the actual grid and non-CEII synthetic grids. The third section provides some overview results from the study based on the actual grid models, whereas the fourth section provides more detailed results on the study's methodology and associated visualizations using the synthetic grids. Then the fifth section summarizes the paper and discusses some future directions.

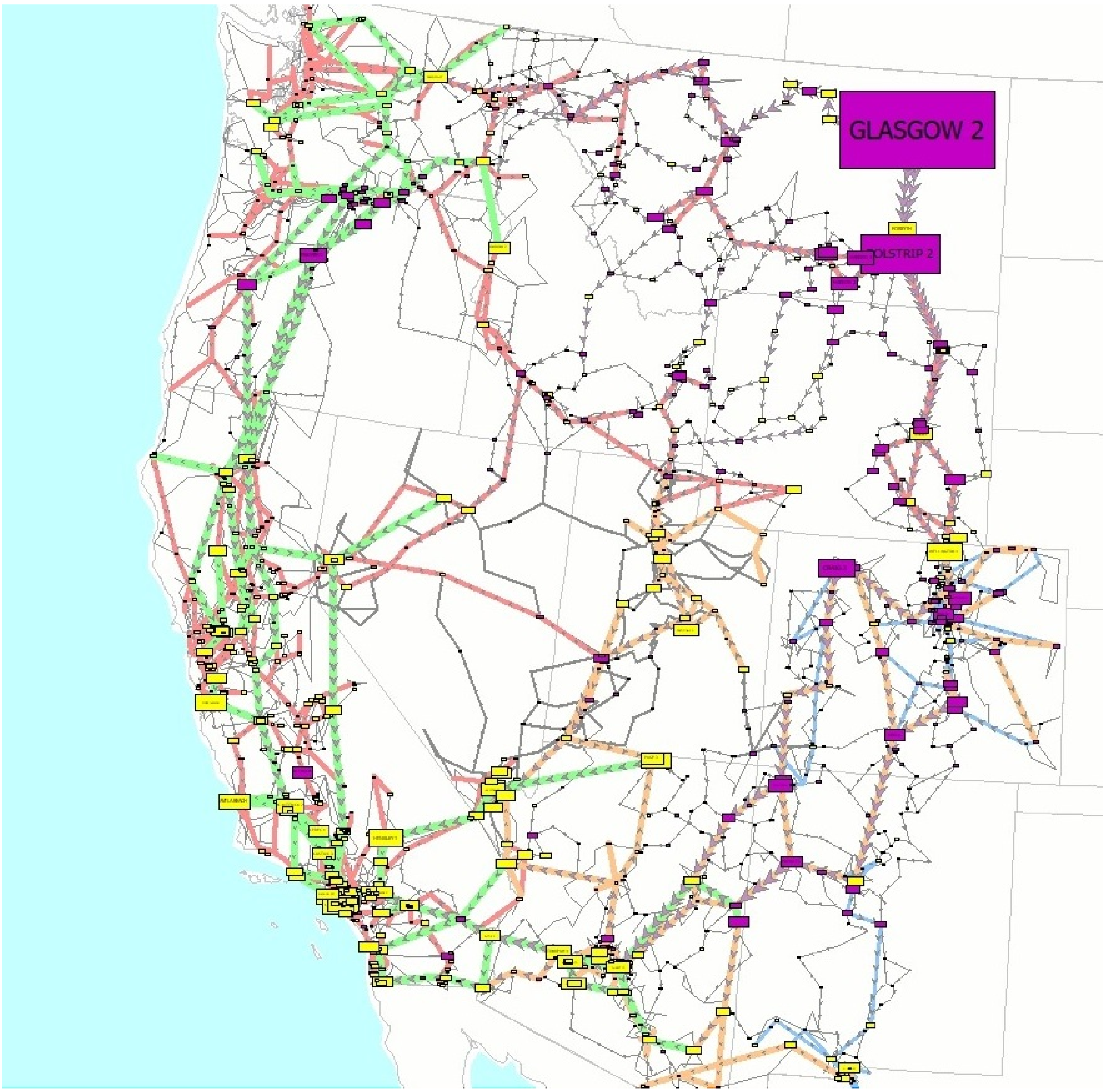


Figure 2.11: 10K Grid Visualizing the Source of the 0.31 Hz Oscillation

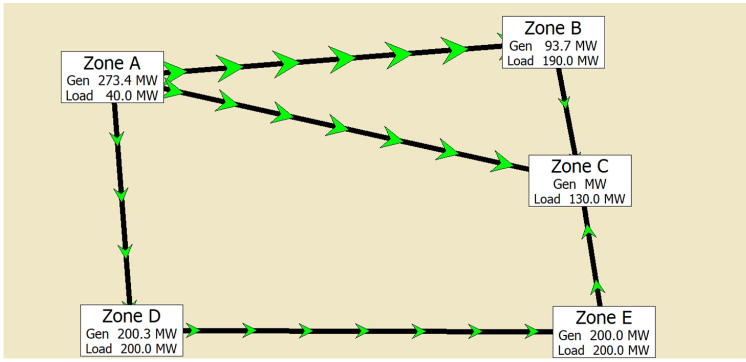


Figure 2.12: Seven Bus System Overall Transmission Flows Visualized

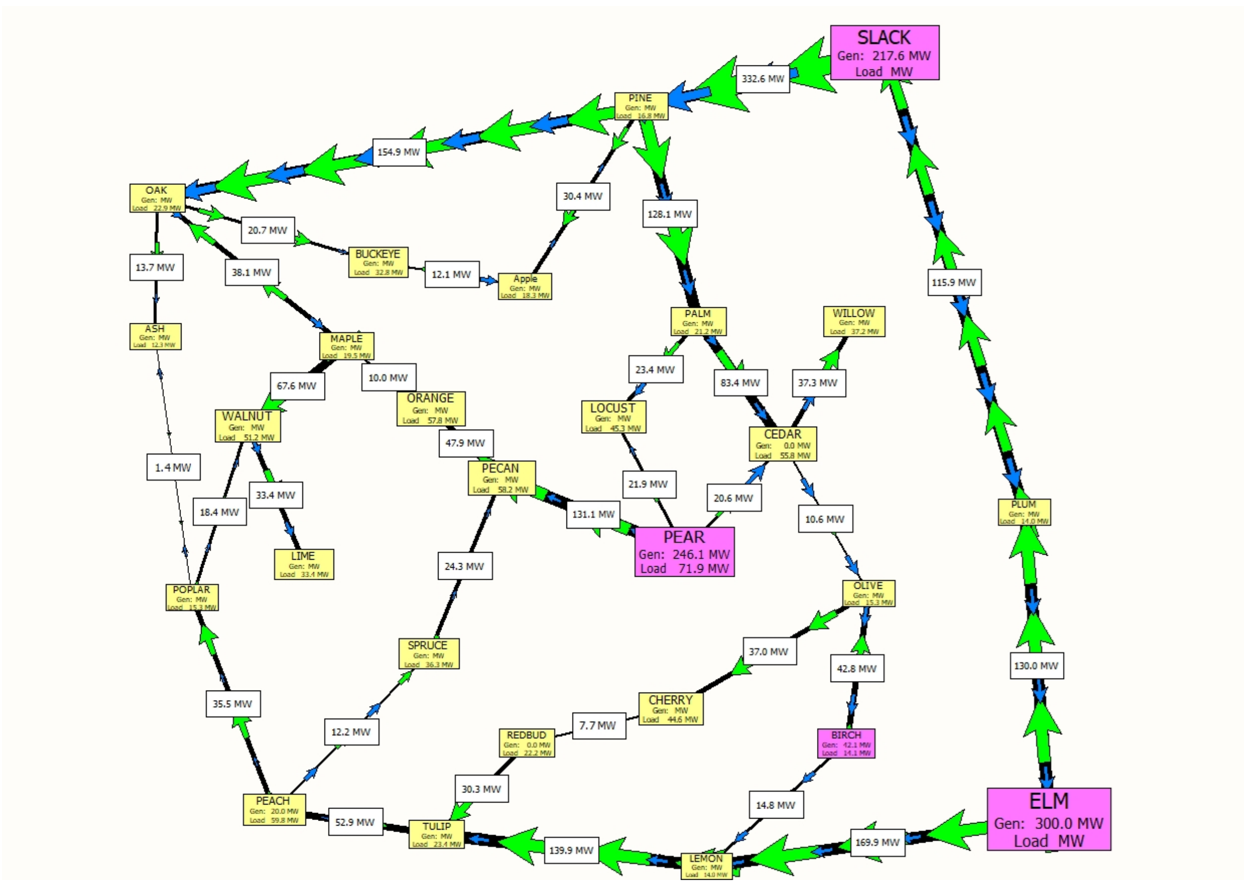


Figure 2.13: Bus System Substation-Based Visualization including Reactive Power and MW Labels

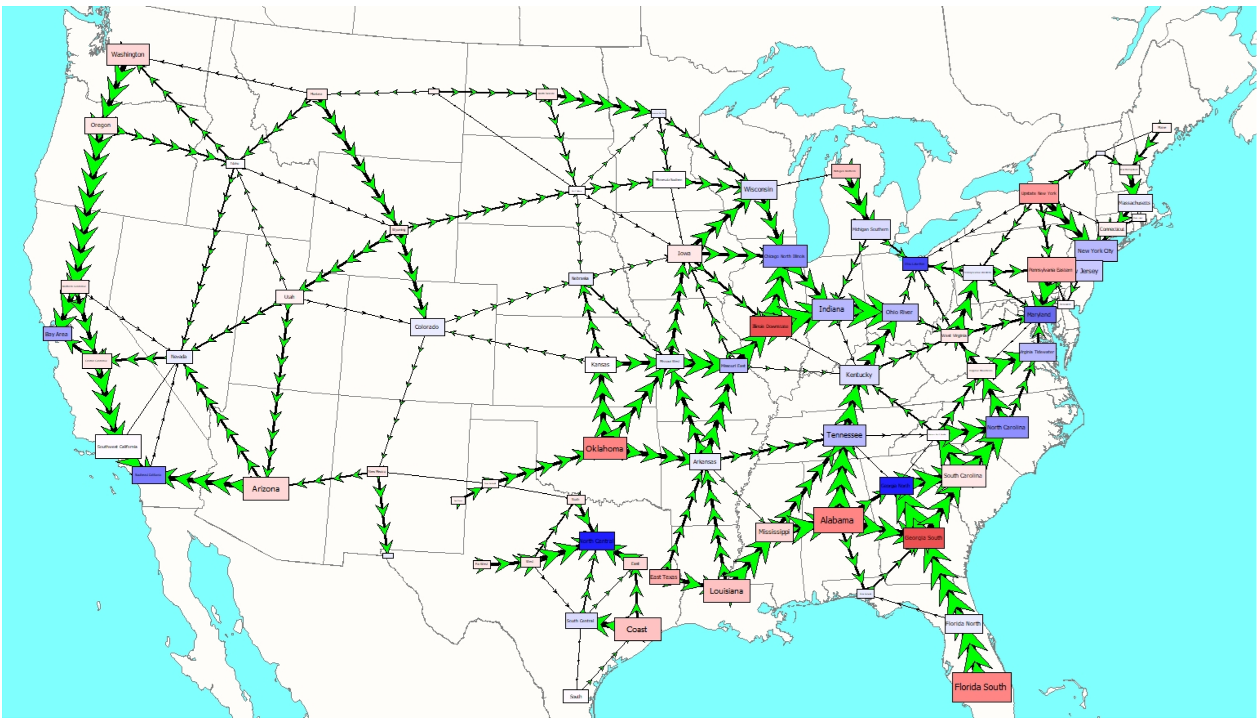


Figure 2.14: 82K System Area Full Delaunay Triangulation Visualization

3. Task 2: Developing Specific Grid Scenarios

The goal of Task 2 is to develop specific grid scenarios. We accomplished this objective we creating a geographically accurate synthetic grid modeling CAISO in California as described in Section 3.1. This new synthetic grid enables further geographically impacted research on electric power systems, such as the co-simulation of natural gas and transmission networks.

3.1 Developing a Geographically Accurate Synthetic Transmission Grid on the footprint of California

To fulfill research needs without compromising security, synthetic grids are developed with fictitious components and parameters. Typical synthetic grids include a set of electrical components, including transmission lines, transformers, buses, substations, generators, and loads with defined parameters and connections that determine how current will flow through the system. The classic IEEE test cases were among the first synthetic grid models [77], followed by recent synthetic networks created by Texas A&M University [**?**, 9], the University of Wisconsin-Madison [78, 79], the Reliability Test System-Grid Modernization Lab Consortium (RTS-GMLC) [80], and other test cases compiled into the power flow library for benchmarking AC OPF algorithms [81]. These synthetic grids provide increased security and ease of use relative to real grid models.

One shortcoming of these commonly used synthetic grids is their lack of realistic geography or geographic information. Though they are based on portions of real power system networks from the 1960's and 1970's, the standard IEEE test cases do not include any geographic information. The Texas A&M grids, the University of Wisconsin-Madison networks, and the RTS-GMLC network are all synthetic ("realistic but not real") electric power grid models that contain geographic coordinates ranging from the southwestern United States to the entire continental United States. These network models do not represent any specific actual power system, but model the characteristics and electrical behavior of real power systems [21, 82, 83]. Thus, the transmission lines, substations and buses do not correspond to any existing power systems network equipment. Also, the transmission lines are represented as straight lines connecting two nodes, rather than nonlinear paths that curve based on local topography, vegetation, and property ownership.

Realistic or accurate geography with high granularity is important for geo-referenced applications, including studies related to weather, topography, and socio-economic considerations. For example, the models in [84] balance the competing risks associated with wildfire ignition from power equipment and public safety power shutoffs. Using a geographically accurate synthetic grid to test such a model allows for synthesis with spatial data, including wildfire risk maps and demographic data. Thus, there is a need for synthetic grids without CEII-protected information that represent the existing geography of real power systems.

The main contribution of this section is a geographically accurate synthetic transmission network model that includes the service territory of the California system operator CAISO. This electric network was developed by adding approximate substation and node topologies and first principles-based synthetic network parameters to publicly available geographic data of California's electric infrastructure. The result is an open-source, non-CEII transmission network model suitable for geo-located applications. Further, the procedure described in this section can be adapted to create similar transmission and distribution models in other geographic locations.

3.2 Data Sources

The proposed geographically accurate synthetic grid was developed using several sources of publicly accessible data described in this section.

The data sources are organized in the following categories: electric infrastructure geographic information, generation data, synthetic load profiles, and transmission line parameter data.

3.2.1 California Electric Infrastructure Geographic Data

Geographic information systems (GIS) represent spatial data in the form of maps. GIS data of California's transmission lines and substations were obtained from the California Energy Commission (CEC) [1]. This accurate geographic data can be downloaded as two separate shapefiles (.shp) and opened in a number of GIS mapping software applications, such as ArcGIS Pro. These shapefiles store the location, shape, and attributes of California's actual transmission lines and substations, which are represented as line and point features, respectively.

The CEC also provides a shapefile of California's power plants. Additionally, the Energy Information Agency (EIA) releases publicly available information about California's power plants, including geographic coordinates, in the Form EIA-860 [85]. The location and capacity of the generators and plants from the CEC and EIA sources are similar, but not exactly the same. To inform generator locations in the proposed network, we selected the 2019 Form EIA-860 over the CEC power plant shapefile since it provides additional attribute fields, such as power factor, that could be useful for various applications.

3.2.2 Generation Data

In addition to geographic coordinates, the 2019 Form EIA-860 contains useful generator attribute information, such as the nameplate MW capacity, power factor, minimum MW values, fuel type and unit type. However, it does not contain details about renewable energy generation output and non-renewable generation cost curves. Since such information is needed to create the synthetic grid, we use state-wide renewable data published by CAISO for 2019 [86] and the quadratic cost curve coefficients from [87] to create the synthetic system described in this paper. To create the cost curves for steam boiler plants, [87] used typical heat rate values from [88] and multiplied them by fuel costs from the EIA website. Heat rates for natural gas (simple and combined cycle) plants were obtained from GE, and multiplied by average natural gas fuel prices. For nuclear, wind and solar, offer curve data was used. All of the generators in the Form EIA-860 were grouped into clusters, and representative generators from each cluster were manually assigned cost curves. The remaining generators were automatically assigned cost curves based on the representative generators [87].

3.2.3 Load Data

In this work, we leverage publicly available aggregate hourly load data from CAISO for 2019 [86]. This is combined with a method developed as part of the EPIGRIDS project, which provides hourly load data at census tract level granularity and captures variation in the load profiles throughout the state of California [78, 89]. In brief, The EPIGRIDS method is a a multi-stage method that disaggregates state-wide temporal load information to individual loads. To achieve this, GIS data is used to estimate the percentage of the state load consumed at each load location, as well as the distribution of residential, commercial, and industrial load types. This is combined with information on typical load profiles associated with residential, commercial, and industrial loads to create a load profile for each load bus, appropriately scaled to match the total state-wide load. In this paper, we leverage the same original methodology from EPIGRIDS, but update the load data to reflect the CAISO 2019 aggregate load.

3.2.4 Transmission Line and Transformer Parameter Data

Transmission line parameters, including resistance, reactance and susceptance, are protected data. Thus, synthetic values must be generated. We use publicly available data to generate realistic parameters for the transmission lines and transformers in the proposed synthetic system. FERC publishes historical and current annual reports of information about electric utilities in the United States. Two of these reports, the Form No. 1 “Annual Report for Major Electric Utility” [90], and the Form No. 715 “Annual Transmission Planning and Evaluation Report” [91], contain useful data for assigning line parameters. The Form No. 1 is publicly available, while the Form No. 715 is classified as CEII and no longer released to the public. Although access to the Form No. 715 can be requested, the authors opted instead to only use the Form No. 1 from 2010 and previously published average values and statistical data derived from the Form No. 715 [21, 78] to avoid concerns regarding protected data in the proposed synthetic grid.

The Form 1 was created for utility companies to report their bulk electric system assets to FERC for annual accounting. In California, three utility companies submitted the FERC Form 1 report and thus their data is utilized to create the proposed grid. The data includes voltage level, transmission line length, number of conductors per phase, conductor size, conductor type (ACSR, ACSS, XLPE, etc), transmission structure material and construction type. The granularity of the data varies, ranging from total mileage of transmission lines at each voltage level to individual transmission structure construction for individual segments of each transmission line.

The Form 1 does not contain any useful data for transformer parameters. We therefore leverage average per unit impedance values using the transformer base MVA for each pair of primary-secondary voltages from [78] and X/R ratios for all transformers with MVA values ranging from 50 MVA to 2000 MVA were obtained from [21].

While there is an abundance of publicly available data, several important aspects are needed to create a more complete power system model. This includes the topology of the system (e.g., connections between different components), line and generation parameters, and locations and parameters of transformers and reactive power compensation. We describe the methodology used to generate each of those aspects below.

3.3 Topology

The data from CEC and EIA are collections of individual components, rather than a fully connected network suitable for power system simulation and analysis. This section introduces the method for creating connectivity and then describes various data cleaning and processing steps.

The available datasets include substation locations, generator locations, and transmission line paths. The first step in achieving a connected topology is to assign substation connections to the end points of the transmission lines. To do this, we calculate the distance from a transmission line endpoint to each substation in the network, and assign the closest substation as a connection. This is repeated for each transmission line endpoint in the network. A similar process is conducted for assigning generators to substations. We calculate the distance from each generator to each substation, and assign a generator connection to the closest substation.

While this connectivity method is simple and intuitive, several data challenges cause the resulting network to be a poor representation of the CAISO network.

i) Substation locations do not represent all network nodes: Many of the transmission lines are made of several line segments, and substations do not exist at all of the end points of these line segments. This prevents transmission line segments from being connected and line termination points from being modeled correctly.

ii) Lines branch within a line segment: The transmission lines from the CEC data represent the actual line locations, but the data does not accurately model electrical circuits and collections. In particular, a multi-circuit path where a single circuit branches away at some point along the length of the line may not be correctly represented. At such points, the transmission line paths do not break and the data is missing a connection point where the branching line enters or leaves.

iii) Automated processing steps require manual cleanup: Some of the automated methods to address challenges i) and ii) cause secondary topology problems that require manual analysis and data cleaning.

iv) Substations do not include transformer connections: The data sets include transmission line voltage levels, but not the connections between voltage levels. Substation topologies must be created if multiple voltage levels connect to the same substation.

The solutions to the challenges are discussed below.

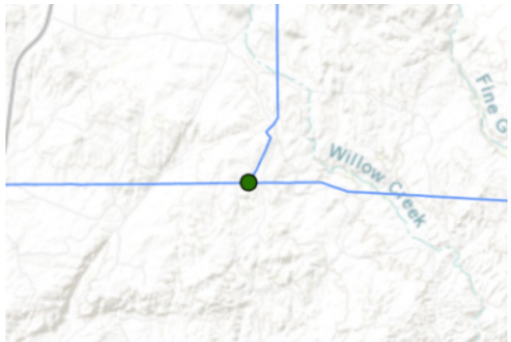


Figure 3.1: An example of a transmission line branching off a from another line. The green node was added at the vertical line's endpoint, and the horizontal line needed to be split at that point to model a connection.

3.3.1 Supplementing the network with additional nodes

The substations do not represent all necessary connections between transmission lines, and thousands of additional nodes were required to represent electrical interconnections between transmission lines. These "added nodes" are placed at every transmission line segment endpoint that does not already have a node or substation within a small search radius. A conservative search radius of just 12 meters ensures that nodes are still placed at the endpoints of very short line segments that exist within cities.

3.3.2 Breaking down transmission lines paths

Some transmission line paths are continuous line segments that must be further segmented to represent connections. An example is shown in Fig. 3.1 where the vertical line has had a node (shown as a green point) added at its endpoint. However, the horizontal line does not terminate at the node, and therefore will not be modeled as an electrical connection. This line must be segmented to add a line endpoint and model the connection to the vertical line. All lines in the network that pass through a node like this are segmented into smaller line paths. The process created a few hundred extremely short line segments, some smaller than one meter. These lines segment were manually deleted or merged with another line.

3.3.3 Resolving Topology Issues

After the above geo-processing stages are complete, the network connectivity process can be run again, but there are new challenges to solve. Certain parts of the topology, such as Midway Substation in Kilowatt, California, shown in Fig. 3.2, do not correctly connect because too many nodes have been added to nearby transmission line endpoints (due to the conservative search radius used when adding new nodes). Large substations like Midway cover a large area of land in reality but are represented as a single point, resulting in relatively large distances between the substation point and adjacent line endpoints. With these additional nodes, many lines would not correctly connect in the substation and instead link to nodes outside of the substation, preventing the topology from representing the connectivity of all of these lines. These extra nodes are removed manually, but we

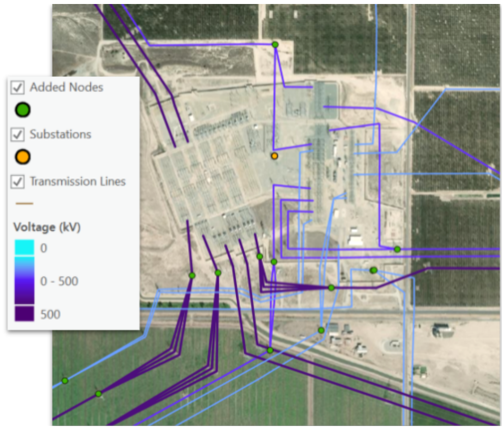


Figure 3.2: GIS representation of Midway Substation in Kilowatt, California, overlaid on top of a Google Maps satellite image.

created a tool to aid in identifying other locations where there may be connectivity challenges.

The tool calculated the graph-distance (with line length as edge weights) and geographic-distance between node pairs. If there was a large discrepancy between these values, it indicates a potential network problem where nearby nodes are not connected. This allowed easier identification of locations similar to Midway substation, where creating correct topology connections in a fully automated way is challenging. After identifying a connectivity issue, we made modifications to the geographic data to solve the problem.

One particular connectivity challenge arises when a line is assigned a connection at the same node at both ends. In these cases, an added node at an interconnection was often closer than the substation node, and therefore both line endpoints were assigned to that same added node. A list of these lines was produced and they were addressed manually by merging, extending, and deleting lines and deleting nodes in ways that allowed components to correctly connect.

3.3.4 Substation transformers

The final topology challenge is the addition of transformers. When multiple voltage levels connect at the same node, the node must be broken into multiple nodes with transformers to connect them, transformers assigned to connect nodes, and lines and generators must be reassigned a connecting node.

The process for transformer node creation is as follows.

1. Identify the number of voltage levels V connecting to a node.

2. Add $V-1$ additional nodes to the network.

3. Add new transformer branches to connect the nodes, where the highest voltage node connects to the second highest voltage node, the second highest voltage node connections to the third-highest, etc.

4. Assign each line that connected to the original node to the new node representing the correct voltage level.

5. Assign any generators to the highest voltage level in the set.

3.3.5 Final Topology

For each of the topology creation steps, we modified and cleaned some of the input data and then re-ran the automated connectivity steps. The final grid topology, after the automated and manual processing is complete, is shown in Fig. 3.3.

3.4 Assigning generator, load and renewable energy data

We next describe our methodology for assigning generation, load and renewable energy data to the grid.

3.4.1 Assigning data for conventional generators

The information pulled from the 2019 Form EIA-860 files includes static attributes, including the nameplate capacity (MW), nameplate power factor, latitude, and longitude.

In addition to the generators location and production capacity, we also need generation cost curves to model generator dispatch within the electricity market. The cost curve coefficients (see Section 3.2.2) are correlated with the 2019 Form EIA-860 generators. For 'Plant Codes' and 'Generator ID's' that directly matched between the 2010 cost curve spreadsheet and the 2019 generators, the cost curve coefficients were simply copied over. For the remaining non-renewable generators, we assign the coefficients of the closest-sized generator of the same type. For the renewable generators, we assign coefficients of zero.

3.4.2 Assigning data for renewable generation

To account for the variability of solar and wind generation, we used the state-wide generation data published by CAISO for 2019 [86]. The data consists of load and generation from different resources for every five minute interval, from which we sample points at the beginning of every hour. Before beginning, we scaled the capacities of all the solar generators to account for the lower solar capacity in the grid topology created, compared to the peak solar generation in California. The next step is to assign a portion of the total state-wide generation from renewable sources to each renewable generator in the proposed grid. The capacity of each generation was scaled according to its actual rating, as shown in equations (3.1a) and (3.1b).

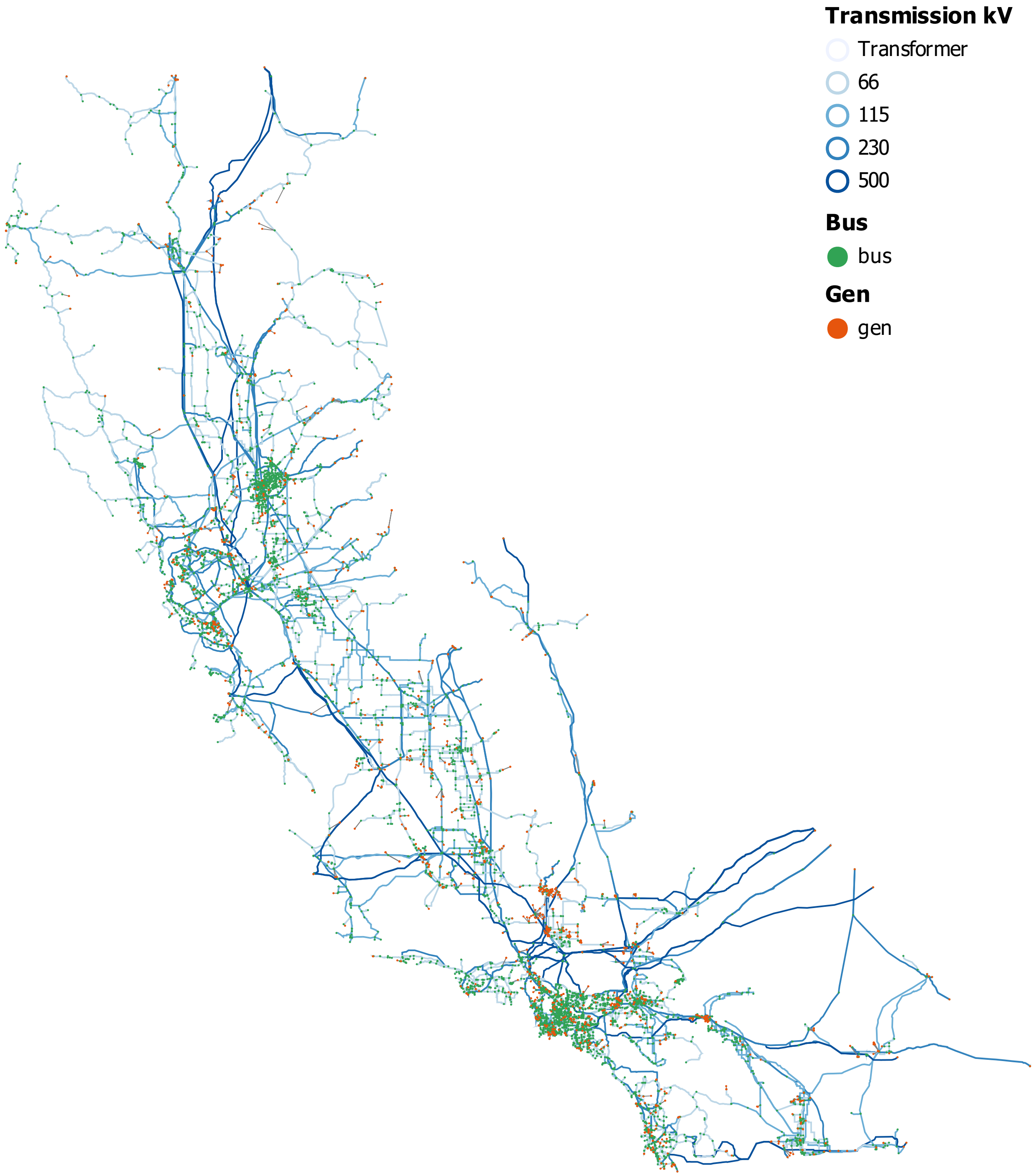


Figure 3.3: California synthetic grid topology

$$P_{PV,i}^{max,s} = \frac{P_{PV,i}^{max}}{\sum_{j\in\mathscr{PV}} P_{PV,j}^{max}} P_{PV}^{s} \quad \forall i \in \mathscr{PV}, s \in \mathscr{S} \tag{3.1a}$$

$$P_{W,i}^{max,s} = \frac{P_{W,i}^{max}}{\sum_{j\in\mathscr{W}} P_{W,j}^{max}} P_{W}^{s} \quad \forall i \in \mathscr{W}, s \in \mathscr{S} \tag{3.1b}$$

In the above equations, $\mathscr{PV}$ is the set of all solar generators and $\mathscr{W}$ is the set of all wind generators in the synthetic grid. The set $\mathscr{S}$ denotes the generation scenarios considered. $P_{g,i}^{max}$ is the actual capacity of generator i (scaled to ensure that the grid is able to produce peak solar), P_g^s represents the total amount of solar or wind generation in the system, respectively, and $P_{g,i}^{max,s}$ is scaled capacity proportional to the generation from renewable sources in scenario s.

We note that the above assignment policy keeps the ratio of renewable generation constant between generators over time. Improving the renewable energy production scenarios to account for more geographic variability is an area for future work.

Load

Hourly load data scenarios were utilized from a previous research project called EPIGRIDS [78, 89]. These loads are given per census tract, and therefore need to be assigned to specific locations in the grid. A first approximation would be to assign each load to the nearest substation from the CEC dataset. However, this assignment method would leave many substations with very high load, many with no load, and some radial substations with no load or generation. To create a better assignment of loads, we use an optimization assignment problem that minimizes the distance between the EPIGRIDS load buses and the CEC substations, with additional constraints on the load assignment shown in problem 3.2.

$$\min_x \quad \sum_{i\in\mathscr{L}} \sum_{j\in\mathscr{N}} x_{ij} c_{ij} \tag{3.2a}$$

$$\text{s.t.} \sum_{j\in\mathscr{N}} x_{ij} = 1 \qquad \forall i \in \mathscr{L} \tag{3.2b}$$

$$\sum_{i\in\mathscr{L}} x_{ij} \geq 1 \qquad \forall j \in \mathscr{N} \tag{3.2c}$$

$$\tag{3.2d}$$

Here, the load assignment variable x_{ij} is a binary variable that determines if load i is assigned to node j. The objective function minimizes the cumulative distance from the location of each load to the substations, represented by c_{ij}. The set $\mathscr{L}$ is the set of all loads from EPIGRIDS, with one load per census tract. The set $\mathscr{N}$ is the set of nodes where we require the algorithm to assign at least one load. This set $\mathscr{N}$ includes all the original substations from the CEC data set, as well as the added network nodes located at the end of radial branches that do not have generators attached. The constraints on the load assignment are that each load $i \in \mathscr{L}$ is assigned is assigned to exactly one node (3.2b), and that each node $j \in \mathscr{N}$ is assigned at least one load (3.2c). The set of loads $\mathscr{L}$ is larger than the set of nodes $\mathscr{N}$ at which we assigned the loads, permitting a feasible solution to the problem. Since this problem satisfies the conditions for total unimodularity, it produces a solution with $x_{ij} \in \{0,1\}$ without explicitly enforcing that x_{ij} is binary.

3.5 Assigning Line and Transformer Parameters

To obtain a system which gives rise to feasible and realistic power flow and optimal power flow solutions, we need to add realistic line and transformer parameters such as impedances and MVA limits. This section describes the procedure used for generating these limits. The key steps in the procedure are outlined in Fig. 3.4.

3.5.1 Input data

Choosing load and generation scenarios

To create a realistic power system network model that is feasible for a wide range of load and renewable generation scenarios, we have to consider more than one loading condition when assigning line parameters. Considering every possible load and generation scenario is not practical for computational reasons. To increase the computational efficiency of the grid creation algorithm while still designing for a large range of operating conditions, the load data was sub-sampled to create a representative set of scenarios that approximates the boundaries of the set of load and generation scenarios.

Specifically, we picked a total of 245 loading scenarios. The first 121 scenarios are chosen to be the hour with maximum load, as well as the 60 hours before and after. The next 121 scenarios are chosen to be the hour with the minimum load, as well as the 60 hours before and after. In addition, we consider 3 scenarios that represent the hour with the maximum solar generation, the hour with maximum wind generation and the hour with the overall lowest renewable generation. This ensures that we observe a range of both load and renewable generation scenarios, as well as different hours of the day and days of the week.

Generate initial generation profiles for each scenario

When creating generation profiles for the scenarios, we both want to reflect typical operating conditions (i.e., conditions that allow the lowest cost generators to run). At the same time, we should avoid over-optimizing the grid such that transmissions lines are sized only to support the lowest cost generator dispatch (economic dispatch). Also, load power demand changes throughout a day or week, and changes between months and years, this creates varying flow flow patters across the grid. If the grid creation process does not include enough variety in the generator dispatch and unit

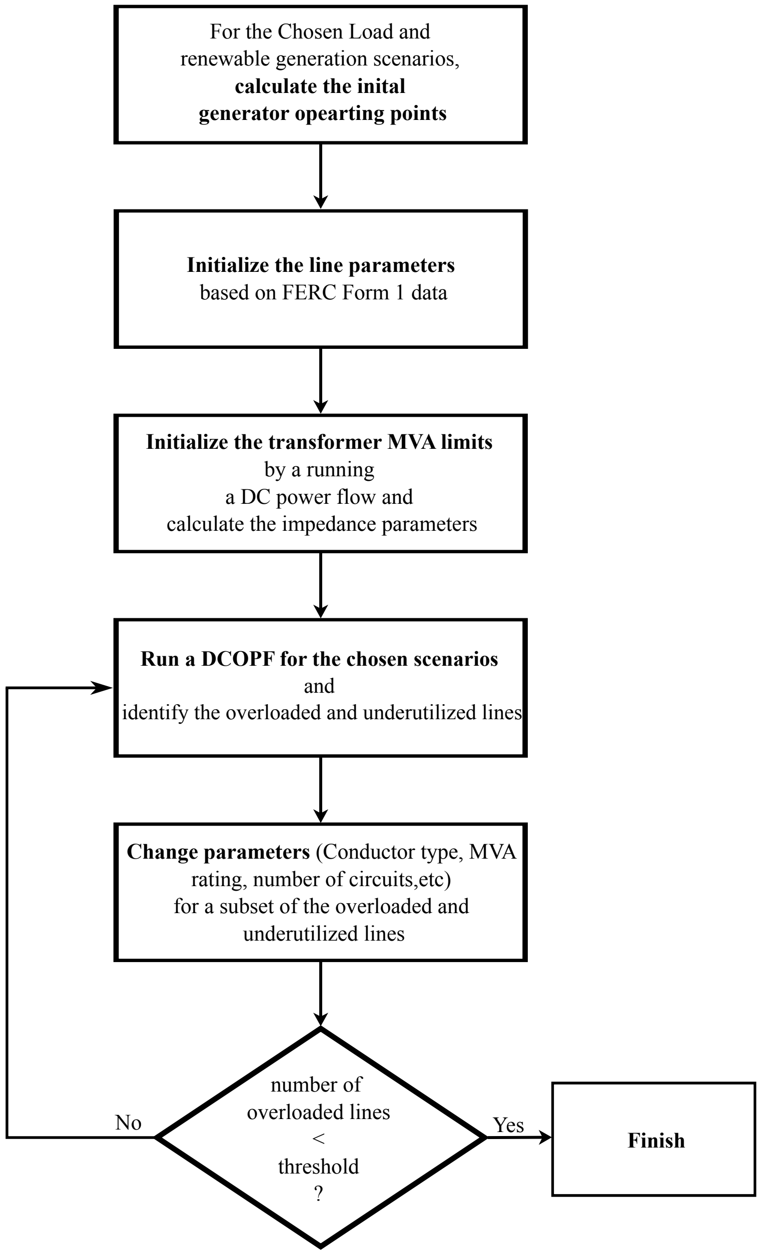


Figure 3.4: Flowchart of the grid creation process.

commitment, the grid will not be able to support varying load flow patterns and cannot transfer larger amounts of power in different directions across the grid.

Past experience has also shown that synthetic grid models are highly sensitive to the set of generators committed, and that at each generator must be dispatched at their maximum output level in at least one scenario [78]. Otherwise, the transmission lines connected to the generator points of interconnection (POI) for decommitted generators will be inadequately designed.

To obtain a varied set of power injections from the generators, we generate two sets of generation schedules for each hour using two methods we call the *economic dispatch* and *uneconomic dispatch*:

- *Economic dispatch:* For each hour, we implement a simple unit commitment algorithm. This algorithm iteratively decommits the most expensive generator until a target spinning reserve level of 10% is met. Once the unit commitment is fixed, the generators are dispatched using an economic dispatch algorithm that minimizes the generator cost subject to the total demand equaling total generation.

- *Uneconomic dispatch:* In the uneconomic dispatch, we follow a similar procedure, but instead decommit the cheapest generators to create an "uneconomic dispatch". The final injections are again computed by running an economic dispatch algorithm considering just the most expensive generators.

The economic and "uneconomic" dispatch scenarios result in all generators being dispatched at, or near, their maximum power output in at least one generator unit commitment scenario. Since we create two generation scenarios for each of our 245 load scenarios, we consider a total number of 490 power injection scenarios.

Initializing the transmission line and transformer parameters

As a final input to our method, initial transmission line and transformer impedances are assigned to each of the network branches. Instead of using a simple assignment such as assigning a uniform per-unit-length impedance to all transmission lines, we used transmission line data from the FERC Form 1 to make the initial assignment. For each transmission line at each voltage level in the proposed synthetic grid, the line in the FERC Form 1 with the closest length was identified. If the utility company listed in the CEC data matched the utility company listed in the FERC Form 1, only the lines in the Form 1 data corresponding to that utility company were examined.

We used the Form 1 data for the matched CEC transmission lines, including conductor size (in kcmil), conductor type, and number of conductors per phase, to determine ampacity limits and transmission line impedances for the corresponding transmission line in the proposed grid, following the methodology described in [78]. As part of this process, transmission line manufacturer's data sheets [92] are used to determine ampacity limits, while [93] is used to determine approximate

GMR and GMD values which are then used to compute synthetic per-unit length transmission line impedances for the lines in the Form 1. Additionally, the calculated ampacity limits are multiplied by the rated voltage of the transmission line to calculate probable MVA thermal ratings for each transmission line in the Form 1.

By examining geographic regions such as states or approximate ISO or RTO service territory, we can determine MVA ranges for each voltage level and region. This data is validated using the MVA limit ranges for transmission lines at each voltage level in [21, 78]. We combine the calculated per-unit-length impedance parameters with the MVA limit ranges to produce a table of possible conductors configurations for each transmission line. Later, we use this table to adapt the transmission line parameters as the MVA limits of lines are increased or decreased, as explained below.

The MVA values and ranges were created assuming that all conductors are ACSR [92]. However, if the conductor material type is changed to ACSS, the ampacity roughly doubles, without a substantial change to the GMR of the wire [94]. This allows the MVA rating of a transmission line to increase by up to 100% of the original value without modifying the corresponding R,X,B values.

Since the Form 1 does not contain useful transformer data, we assign an initial limit of 2000 MVA to each transmission-level transformer. The average per unit impedance values are obtained from [78] using the transformer base MVA and the primary-secondary voltages, and corresponding X/R ratios were obtained from [21]. Once the initial line and transformer parameters are calculated, we solve a DC power flow for all 490 scenarios to calculate the resulting flows through all transmission lines and transformers. We then resize the transformers to have an MVA limit equal to the maximum value calculated from the DC power flow and recompute the impedance parameters corresponding to these calculated flows.

3.5.2 Algorithm for Updating Transmission Line Parameters

The initial line parameter assignment is typically inaccurate and the generation schedules obtained with the uneconomic and economic dispatches, which do not account for network constraints, may not be feasible. In the following section, we describe our algorithm for adjusting both the generation dispatch and the line parameters to obtain a system that allows for feasible power flow and optimal power flow solutions.

Step 0: Initialization

Define the 490 power injection scenarios and initial line parameters as discussed above.

Step 1: Solve line upgrade optimization problem

For each power injection scenario, solve the optimization problem (3.3). This optimization problem attempts to minimize the size of transmission line violations while also limiting generation redispatch away from the assigned power injection schedule. To achieve this, the objective function (3.3a) is formulated with two terms: (1) a penalty on $\Delta p_{g,s}$, which measures how much generator g

is redispatched in scenario s, and (2) a penalty on $\delta_{ij,s}$, which measures the violation of the power flow limit on line ij in scenario s. The factor λ is a trade–off parameter that balances how much we penalize the generation redispatch and the amount of line limit violations in the solutions. A smaller value for λ allows for more generation redispatch and leads to fewer line updates. A larger value for λ penalizes generation redispatch more and thus forces more line upgrades. If we do not allow any redispatch at all, the procedure becomes similar to solving a power flow for each load scenario. Based on testing with several values, we set $\lambda = 0.5$ for our final grid. This seems to give a reasonable trade–off between the number of lines that need upgrades and limiting generation redispatch.

$$\min \quad \lambda \sum_{k\in\mathscr{G}} \Delta P^s_{g,k} + (1-\lambda) \sum_{(i,j)\in\mathscr{L}} \delta^s_{ij} \tag{3.3a}$$

$$\text{s.t.} \quad P_{g,k} - \sum_{(i,j)\in\mathscr{L}} \beta^k_{ij} P_{f,ij} = P_{d,k} \quad \forall k \in \mathscr{B} \tag{3.3b}$$

$$P_{f,ij} = B_i j(\theta_i - \theta_j) \quad \forall (i,j) \in \mathscr{L} \tag{3.3c}$$

$$-P^{max}_{f,ij} - \delta^s_{ij} \le P_{f,ij} \le P^{max}_{f,ij} + \delta^s_{ij} \quad \forall (i,j) \in \mathscr{L} \tag{3.3d}$$

$$P^s_{g,o} - \Delta P^s_{g,k} \le P^s_{g,k} \le P^s_{g,o} + \Delta P^s_{g,k} \quad \forall k \in \mathscr{G} \tag{3.3e}$$

$$P^{min}_{g,k} \le P^s_{g,k} \le P^{max}_{g,k} \quad \forall k \in \mathscr{G} \tag{3.3f}$$

$$\delta^s_{ij} \ge 0 \quad \forall (i,j) \in \mathscr{L} \tag{3.3g}$$

$$\Delta P^s_{g,k} \ge 0 \quad \forall k \in \mathscr{G} \tag{3.3h}$$

$\mathscr{B}$ and $\mathscr{G}$ are the set of all the buses and generators in the grid respectively. The equality constraints (3.3b) represent the DC power flow equations, while (3.3d) represent the relaxed transmission line limits and (3.3e) represent the generation limits after redispatch. The primary output of this optimization problem is the line limit violations δ^s_{ij} for each line $(i,j) \in \mathscr{L}$ in each scenario $s \in \mathscr{S}$. Note that if the DC power flow solution is feasible for the original power injection scenario s, the optimization problem would set $\delta^s_{ij} = 0$.

Step 2: Identify and upgrade overloaded lines

To identify the overloaded lines, we compute the maximum violation across all scenarios,

$$\delta_{ij} = \max_{s\in\mathscr{S}} \ \delta_{ij,s} \ . \tag{3.4}$$

We then randomly choose a subset of 541 lines to upgrade, corresponding to 5% of the total number of lines in the system[1]. If fewer than 541 lines are overloaded, we upgrade all the lines.

[1] The reason for upgrading only a subset of lines is that some overload problems may resolve by themselves in the next iteration once the other lines have been upgraded. The random choice of lines to upgrade reflects the fact that the

For each line (i, j) that is chosen for an upgrade, we use the following procedure:

(a) Using the table of possible conductor types and MVA ratings for this line, upgrade the type of conductor to the one with the closest higher MVA rating.

(b) If the conductor type has already been upgraded to the highest MVA conductor type and could not be updated using the procedure in (a), increase the number of circuits included in the line by one. Note that the maximum number of allowable circuits per line is 8.

If the line has already been upgraded to have 8 circuits, it is left in its overloaded state until the end of the algorithm. Once the algorithm terminates, we double the ratings of all the lines, which can be understood as changing the conductor type from ACSR to ACSS. By doing this, we (a) remove the remaining overloads and (b) ensure that we account for reactive power flows and losses in the lines, which were so far neglected, when assigning reactive power support.

It should be noted that during this entire process, the number of conductors in each circuit was not changed. This is because, in actual networks, the number of bundled conductors in each circuit is usually limited to two or three, with some exceptions. To get parameters that are as realistic as possible, the number of conductors was thus fixed.

Step 3: Identify and upgrade underutilized lines

A similar set of changes are performed to reduce the ratings of the lines that are under-utilized. Transmission lines with a utilization lower than a given threshold (30% in our case) are classified as underutilized. As done in the case of overloaded lines, a random subset of 541 underutilized lines are chosen. In case the number of underutilized lines is smaller than 541, all such lines are downsized.

For each of the chosen lines (i, j):

(a) Using the table of possible conductor types and MVA ratings, downsize the line by choosing the conductor type that has the closest lower MVA rating

(b) If the conductor type has already been modified to one with the lowest MVA rating, reduce the number circuits in the line by one. Since a line must have at least one circuit, do not decrease the number of circuits in a line once it equal one.

If a line has already been downsized to have just the one circuit of the smallest allowable conductor size and it is still underutilized, do not downsize it further.

power system has been evolving over a long period of time, and thus is not always built to be optimal for the present day loading conditions.

Step 4: Check termination criterion

If the number of overloaded lines is below a threshold τ, then the line resizing terminates. Otherwise, we return to solving the optimization problem in Step 1. At the beginning, the threshold τ is set to zero. However, if the algorithm fails to terminate after a certain number of iterations, the threshold is increased every iteration until the algorithm terminates. This is to prevent "cycling" where the algorithm upgrades a set of lines, causing another set of lines to become under-utilized. Correcting these under-utilized lines then causes the first set of lines to become overloaded, thus driving the algorithm into an infinite loop if the threshold τ is not increased.

3.6 Assigning Reactive Power Support

The reactive power output of the generators alone is not sufficient to maintain the voltage at each bus within its limits. Thus, to ensure that the grid gives rise to an AC power flow feasible solution where all voltage limits are satisfied, we need to add reactive power compensation elements to the network. We do this using the algorithm described below. Due to the high computational burden associated with solving AC optimal power flow for a network of this size, this algorithm considers only a single power injection scenario corresponding to the maximum load scenario with economic generation dispatch. As mentioned in Section 3.5.2, the thermal limits of the lines are temporarily doubled to ensure that the network at hand is ACOPF feasible and the reactive power flow and losses in the network are accounted for.

Step 0: Initialization

Add reactive power compensation in the form of synchronous condensers to all nodes in the network. The initial capacity of the reactive power compensation devices is set to 200 MVAr.

Step 1: Solve AC optimal power flow

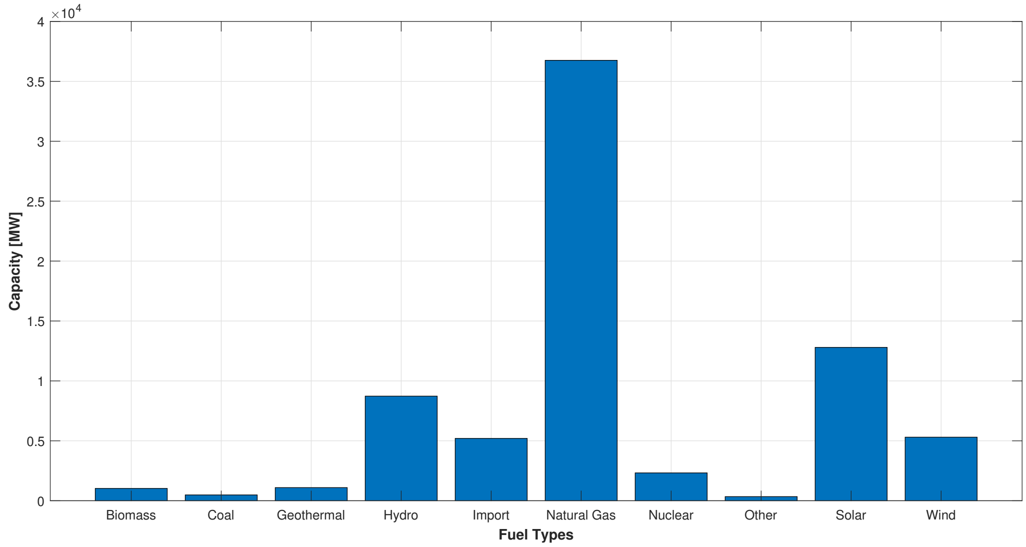
For the problem with reactive power compensation installed, we solve a standard AC optimal power flow problem which minimizes generation cost subject to AC power flow, generation, transmission and voltage magnitude constraints [95].

Step 2: Remove redundant compensation

Remove reactive power compensation from the 20% of the nodes that currently have reactive power compensation.

Step 3: Check termination criterion

If fewer than 20% of all nodes have reactive power compensation, terminate. Otherwise, go back to Step 1 and resolve the AC optimal power flow.



(a) Generation capacity (MW) of each fuel type

Step 4: Restoring the line limits

After assigning reactive power support, the thermal limits of lines with a utilization of less than 50% are brought back to their original values before they were temporarily doubled. This results in around 2.54% of lines with limits that are doubled, corresponding to an upgrade from an ACSR conductor to an ACSS conductor, as described in Section 3.5.1.

3.7 Grid Metrics and Evaluation

A brief overview of the network structure is given in Table 3.1. The created synthetic network has 8848 buses, out of which 1,483 have reactive power support in the form of synchronous condensers. There are 10,140 transmission lines and 663 transformers in the grid. The system has 2461 load buses with a peak load of 44,001.4 MW and 2,229 generators with a total capacity of 74,052 MW. Generation capacities for different fuel type is shown in Fig. 3.5a.

An important metric to evaluate synthetic networks is the node degree distribution, which captures the frequency of the node degree (number of lines connected to the substation) of each substation. Fig. 3.6 shows the node degree distribution for our synthetic grid. As expected, the node degree distribution of our network follows a downward trend and agrees closely with the trends followed by real networks [78].

Characteristics of the network branches, shown in Table 3.2, for the synthetic grid follow trends similar to those followed by the real grids presented in [78]. Any statistical comparison of the synthetic network and real grids will have shortcomings due to the fact that there are only three samples of real world networks in the USA.

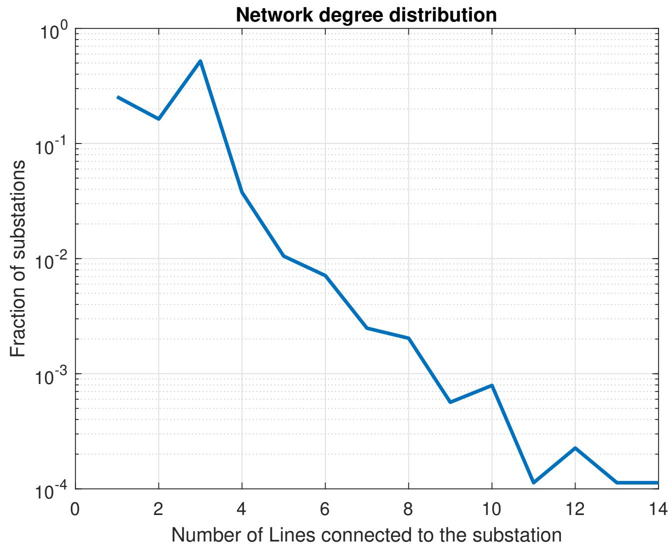


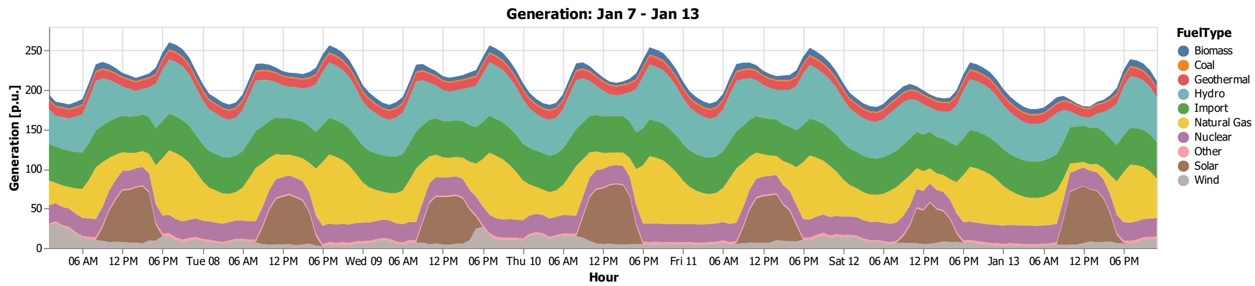
Figure 3.6: Node degree distribution for the network

Metric	Number
Buses	8848
Transformers	663
Transmission lines	10140
Generators	2229
Buses with reactive power compensation	1483
Load Buses	2461

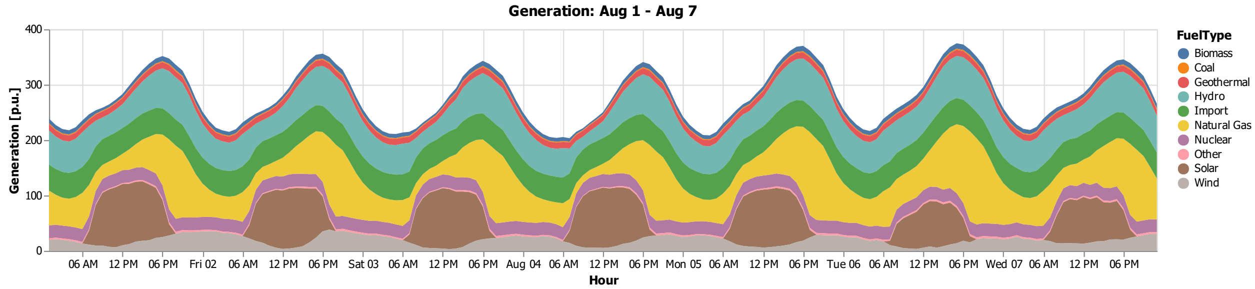
Table 3.1: Network data

Voltage level (kV)	GVA-miles	Total length (miles)	Percent of lines
66	952.40	12696	0.6485
115	1760.95	8076.1	0.2170
230	5651.86	8256.5	0.1225
500	10213.52	4637.2	0.0120

Table 3.2: Characteristics of the network branches



(a)



(b)

Figure 3.7: a) shows the generation dispatch for Jan 7 through Jan 13. b) shows the generation dispatch for Aug 1 through Aug. The winter generation profile is more constant, with less ramp in the evening, while the summer dispatch has much high load for air conditioning. Also note the shorter output duration of solar (brown) in the winter (a) compared to the summer (b).

We have evaluated the operation of the proposed grid across a year of hourly renewable generation and load scenarios. The renewable generation data comes from 2019 aggregate production data from CAISO [86]. As described in Section 3.4.2, for each hour, the capacity of each renewable generator is scaled according to its rating and the aggregate renewable production. For the load scenarios, we use a full year of the previously referenced hourly load data, as described in Section 3.2. For each hour, an DC optimal power flow is run using the `solve_opf` function with `DCPPowerModel` option from the PowerModels.jl package [95]. We then evaluate aspects of the power flow solutions, including feasibility, line loading, generation dispatch, and curtailment.

Out of the 8760 hourly time steps, there is only one time step that does not have a feasible primal point. This occurs at hour 5919, which corresponds to September 4 at 3:00pm. This indicates that this period should likely by added to the scenario set for upgrading transmission lines.

We also examined congestion on transmission lines. Out of 10,140 transmission lines in the network, there are on average 10.1 lines that are operating at their maximum capacity, and the median number of lines at capacity is 10. The highest congestion level, which occurs on August 15 at 10:00pm, has 26 lines operating at their maximum capacity. This indicates that there is a small level of congestion in this network under most scenarios, but the congestion is limited and the grid generally operates with economic efficiency.

Generation profiles for a selected two-week period are shown in Fig. 3.7. A winter profile for January 7 through 13 is shown in 3.7a, while a summer profile for August 1 through 7 is shown in 3.7b. The generation output in the summer is much higher and has a larger ramp rate, as is expected for California. This is especially pronounced for natural gas (yellow) after solar output drops after sunset. We also note that the daily production period of solar generation (brown) is shorter in the winter than the summer, which is a result of fewer hours of sunlight in the winter months.

Curtailment of wind and solar energy is limited in this grid model. The median curtailment for each hour the year is 0 MW, and on average just 26.5 MW of solar or wind is curtailed. Only 233 out of 8760 hours of the year have curtailment over 100 MW. It is important to note that this is not reflective of the much higher level of curtailment in the actual CAISO system. We suspect that the low curtailment is due in part to the fact that we scale the wind and solar capacity by the dispatched power of the CAISO market, not the total available wind or solar at a given hour. An area of future work is to improve the accuracy of the renewable generation profiles.

3.8 Using this Synthetic Grid

This synthetic grid can be used as a test case for power systems research applications. It is particularly valuable for use in geo-referenced applications, such as those relating to weather, climate change, topography, political boundaries, socio-economic considerations, and more. The synthetic grid is available in a GitHub repository[2] in MATPOWER and GIS formats.

The MATPOWER (.m) file is a standard format for storing power system component information. It contains sections for the network components, including 'branch', 'gen', 'bus', and 'load', and their attributes. Generator cost information is given under 'gencost'. Extra component data is listed under separate headings of 'branch data', 'gen data', etc. MATPOWER files are commonly used as inputs for a number of power system analysis software packages and applications, such as PowerModels.jl.

Power system models are not typically available in the GIS format. However, this format can facilitate synthesis with other geographic data. The GIS file is a shapefile (.shp) that can be opened in GIS mapping software, like ArcGIS Pro.

3.9 Summary and Conclusion

The result of the previously described methodology is a synthetic grid that represents the real locations of California's electric infrastructure with invented interconnections and parameters. To the authors' best knowledge, this is the first publicly available synthetic grid that accurately reflects the geography of California's electric transmission infrastructure.

[2]The Github repository is not currently publicly available, but may be made publicly available in the future.

There are several limitations of this synthetic grid. First, the final network is merely an approximation of California's transmission system. While the locations and paths of the components were not significantly modified, the connections and parameters are invented. This is useful for maintaining security, but it also means that any results produced using this grid do not necessarily reflect the results that would be produced from California's actual grid.

Another limitation is the connectivity of the final system. In the current version of this test case, not all of the buses are included in the largest connected network. However, we believe that this is reasonable because it is consistent with other publicly available synthetic grids, such as the Western Electricity Coordinating Council (WECC) grid model.

The authors have identified several areas for ongoing development and future work. One important area is the renewable generation time-series data. We can incorporate weather data with geographic variability to create more realistic renewable generation profiles.

4. Task 3: Exploring Decision Making with Uncertainty

In task 3, the effect of uncertainty on decision making was explored and modeled. This was accomplished by creating wildfire scenarios, quantifying their risk, and proposing an algorithm for grid upgrades to mitigate the risk of grid-caused wildfire ignitions [96] (Section 4.1). To further expand this research, section 4.2 describes dynamic co-simulation of transmission and distribution grids while considering the effect of electric vehicles on grid frequency [97]. Section 4.3 describes how scenarios can be created for research, simulations, education, and operator training [98]. Finally, Section 4.4 the creation of a model to examine the integration of PV, distribution system, transmission system, and cyber infrastructure [99].

4.1 A Framework for Risk Assessment and Optimal Line Upgrade Selection to Mitigate Wildfire Risk

In the United States, the total area burned by wildfires, wildfire frequency, and federal fire suppression costs per year have increased significantly since 1985 [100]. Wildfire prevention is an increasingly crucial effort, especially as climate change exacerbates future fire risk conditions [101]. Power line faults are one of the major sources of wildfire ignitions [102]. Downed lines, vegetation contact, conductor slap, or component failures can produce fault currents and sparks that may ignite fires under hot, dry, and windy conditions [103, 104]. The deadliest and most destructive wildfire in California's history, the 2018 Camp Fire, was ignited by an aging transmission line [105].

A particular challenge of ignitions from electric power lines is that a common factor – high wind speeds – increases both the probability of ignitions due to electric faults *and* promotes a rapid spread of the resulting fire. As a result, wildfires ignited by power lines tend to be larger than fires from other causes [106]. For example, wildfires ignited by power lines in San Diego County account for only 5% of all ignitions, but 25% of the total acres burned [102].

Several strategies may prevent ignitions from electric infrastructure [104]. During high risk conditions, utilities currently implement public safety power shutoffs [107], which de-energize the lines in high risk zones to avoid the release of fault currents and prevent wildfire ignitions. While preemptive shutoffs are effective in preventing ignitions, they can result in wide-spread power outages [108]. This consequence is particularly harmful for people that depend on electric medical devices and members of socially vulnerable communities [109], and still leaves the population exposed to non-power line ignitions. Results in [84] showed that it is possible to reduce both wildfire risk and the size of power outages by incorporating power flow modeling in shutoff decisions.

Less disruptive ignition prevention strategies include vegetation management, replacing aging components, and converting overhead power lines to underground cables [110, 111]. Undergrounding is attractive because it reduces the need for costly short- and long-term ignition prevention strate-

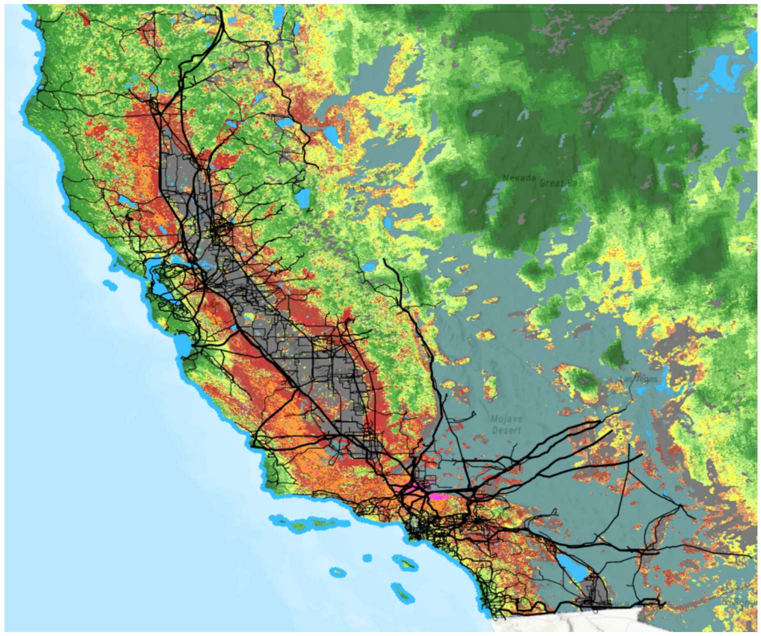


Figure 4.1: Overlay of California's transmission lines [1] with the Wildland Fire Potential Index map for August 1st, 2021 [2]. Warmer colors indicate higher wildfire potential.

gies in the future, as the ignition risk is essentially reduced to zero once the line is undergrounded. Further, underground cables are less susceptible to impacts from wildfires (e.g., flashovers due to air pollution [112] or fire damage to towers [113]), offering another argument for undergrounding in areas with high wildfire exposure. Thus, undergrounding is seen as a highly effective, though expensive, measure to reduce mitigate power line-wildfire interactions.

However, since undergrounding the entire electric grid is prohibitively expensive, we need to select lines that exhibit the highest risk. Assessing the long-term ignition risk associated with a particular power line is challenging due to the complex and time-varying nature of wildfire risk. In the United States, there are currently no publicly available databases specifically for quantifying the ignition risk of power lines and hardening strategies. Thus, there is a need among grid planners for more accessible and flexible methods of defining wildfire risk in the context of upgrade selection.

This section, and in more detail [96], addresses this gap by proposing a framework for assessing the wildfire risk associated with power lines, as well as an optimization problem to select an optimal set of lines for upgrading. The framework defines wildfire risk as a function of two components: the probability of electric faults leading to ignitions, and the potential for large wildfires and fire spread in the area around the line. The probability of electric faults leading to ignitions is hard to assess for individual lines without detailed information typically available only to utilities. However, we can incorporate known trends. For example, per mile of power line, distribution lines are three times more likely to cause ignitions compared with transmission lines [114].

In this section, we take an initial step towards filling this gap. The first contribution of our section is to examine multiple methods of defining the wildfire ignition risk associated with power lines. As a second contribution, we incorporate these metrics as inputs to an optimization model which identifies the optimal set of power lines to underground to reduce wildfire risk, while not exceeding a pre-defined budget. The model considers multiple wildfire risk scenarios to determine the line upgrades that provide the best improvements across all of the scenarios.

Finally, we demonstrate our method through a case study based on the RTS-GMLC grid [80] and the real transmission lines in California [1], in which we analyze different risk metrics and assess how sensitive the upgrade selection is to various model parameters.

4.1.1 Identifying the Optimal Set of Power Lines for Undergrounding

First, we present a simple optimization-based method to demonstrate how data on wildfire risk can be used to support decisions regarding *which power lines* should be prioritized for undergrounding in a region with significant wildfire risk. We consider a power system with N overhead power lines split into line segments represented by the set $\mathscr{L}$. Note that the number of line segments $N_L = |\mathscr{L}| \geq N$. For each line segment $l \in \mathscr{L}$, we define the associated wildfire risk as R_l and the cost of undergrounding by C_l. Further, we define a binary decision variable z_l which represents whether line segment l is undergrounded ($z_l = 1$) or not ($z_l = 0$). We represent the budget for undergrounding lines as $C^{\max}$.

With this information, we formulate a simple optimization problem to identify the optimal set of lines to underground:

$$\min_{z_l} \ \rho(R_l, z_l) \tag{4.1a}$$

$$\text{s.t.} \ \sum_{l=1}^{N_L} C_l z_l \leq C^{\max} \tag{4.1b}$$

$$z_l \in \{0,1\} \quad \forall l \in \mathscr{L}. \tag{4.1c}$$

Here, the objective function (4.1a) minimizes the total wildfire risk in the system, represented by the function $\rho(R_l, z_l)$, which depends on the wildfire risk R_l and the upgrade status z_l of each line. The constraint (4.1b) limits the upgrades to those that are possible within the defined budget, while (4.1c) requires the undergrounding variables z_l for line segment l to take on values of either 0 or 1 (i.e., partial undergrounding is not possible). Assuming that $\rho(R_l, z_l)$ can be expressed as a linear function, this simple model is a version of the classical 0-1 knapsack problem.

In this model, the risk values R_l of a transmission line l represent a single, aggregate measure of risk across the entire geographical span of the line and across multiple scenarios of daily wildfire risk. The wildfire risk function $\rho(R_l, z_l)$ further aggregates the risk values for all the individual lines into a single value. The method used to aggregate wildfire risk metrics can have a significant impact on the results. In the next section, we discuss different options for defining R_l and $\rho(R_l, z_l)$ in detail.

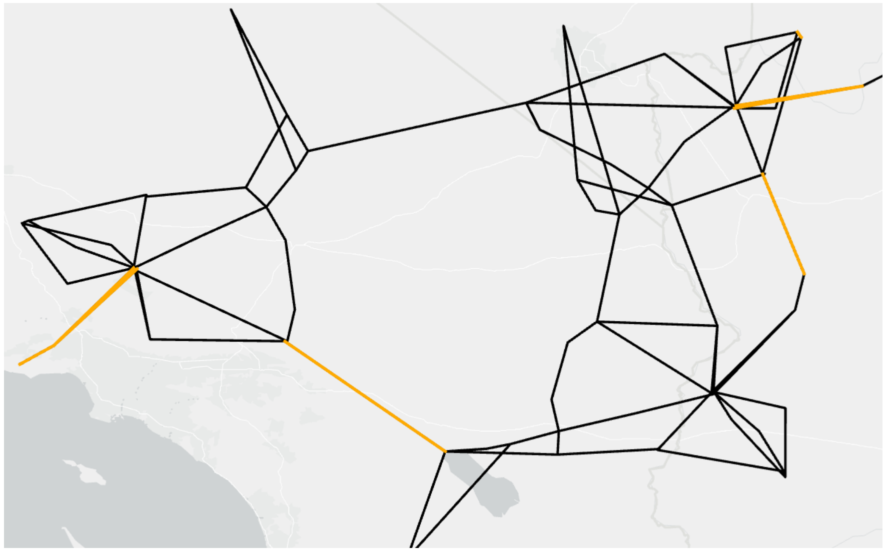


Figure 4.2: Line segments selected for upgrade (highlighted in orange) in simple example.

4.1.2 Assessing Wildfire Risk of Transmission Lines

Obtaining and synthesizing information about all of the factors that impact wildfire risk from electric grids and deriving an aggregate risk value for the entire length of the line is a challenging and time-consuming task. Furthermore, since decisions of which lines should be put underground is a long-term planning problem, we need to consider how wildfire risk varies over time. To address this challenge, we divide task of defining and computing $\rho(R_l, z_l)$ in three parts. First, we discuss how to compute risk for a small line segment at a given point in time, which is accomplished by determining the Wildfire potential $w_{i,s}$ and the probability of ignitions $\pi_{i,s}$. Then, we discuss how these risk values can be aggregated geographically (i.e., along the length of the line) and in time (i.e., across multiple scenarios) for a single line. We examine both the maximum risk or the cumulative risk metric $r_{l,s}$ for a given transmission line. Finally, we discuss how to aggregate risk across all lines in the network for both risk metrics to define the total system risk $\rho(R_l, z_l)$. This process is explained in more detail in [96].

4.1.3 Test Case Results

The proposed risk assignment and upgrade selection methods are demonstrated through several test cases as presented in [96]. First, we solved the cumulative problem variant for the RTS-GMLC system with non-segmented lines. In this example, the cumulative model chooses 8 lines to be upgraded out of the total 104, as shown in Fig. 4.2. This corresponds to 300 miles, or 9% of the total line length.

Additionally, we considered splitting the lines into pieces with a length of at most 10km or 1km. Using shorter line segments allows us to develop a more detailed plan and more carefully target undergrounding efforts in the highest risk zones. However, shorter line segments increase the computational requirements of our model, both in data processing (i.e., obtaining risk values R_l^{cum}

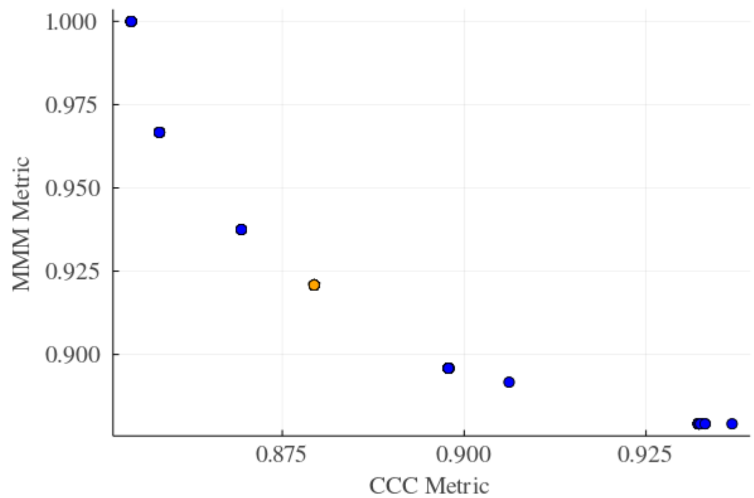


Figure 4.3: Pareto curve of normalized system-wide risk values that result from different values of the trade-off parameter. The orange point corresponds to $\alpha = 0.4$, which is the case used for comparison with the other risk metrics.

and $R_l^{\max}$ values for a large number of line segments l) and in the optimization problem (i.e. each line segment requires the introduction of a binary decision variable z_l to represent whether or not that segment should be undergrounded).

To incorporate both the cumulative and maximum risk metrics, we added a trade-off parameter α that varied between 0 and 1. Fig. 4.3 shows a Pareto plot the solutions that result from varying α. Each point on the curve represents a solution where lines are selected for upgrade such that both the cumulative and maximum risk are minimized, however, the relative importance of each metric is changing as α changes, leading to different solutions.

To examine the solutions in more detail, we solve the problem once for each variant assuming a constant budget of 300 million USD. For the trade-off formulation, the trade-off parameter of $\alpha = 0.4$, which corresponds to the orange point in Fig 4.3. It was chosen for as a solution that shows significant (although not exactly equal) reductions in both the cumulative- and maximum-based risk. In each case, the reductions of maximum risk and cumulative risk are computed based on the selected upgrades and are summarized in Table 4.1. We observe that each method upgrades a similar number of line segments, which is as expected because the budget is the same across all problems. However, the percentage reduction in risk for the maximum and cumulative metrics varies between the different problem variants.

To further analyze the solutions, Fig. 4.4 shows the 10-km line segments selected for undergrounding for the three problem variants. The sets of lines selected by the cumulative and maximum formulations do not intersect, showing that the choice of risk definition has a significant impact on undergrounding decisions. Further, we observe that the trade-off variant includes a combination of line segments from each of the other solutions. This highlights that a trade-off between cumulative

Table 4.1: Comparing Risk Metrics.

	Risk Metric Minimized		
	maximum	cumulative	Trade-off
Segments Upgraded	25	25	26
maximum risk [% reduction]	8.76	8.99	9.43
cumulative risk [% reduction]	6.65	14.58	12.06

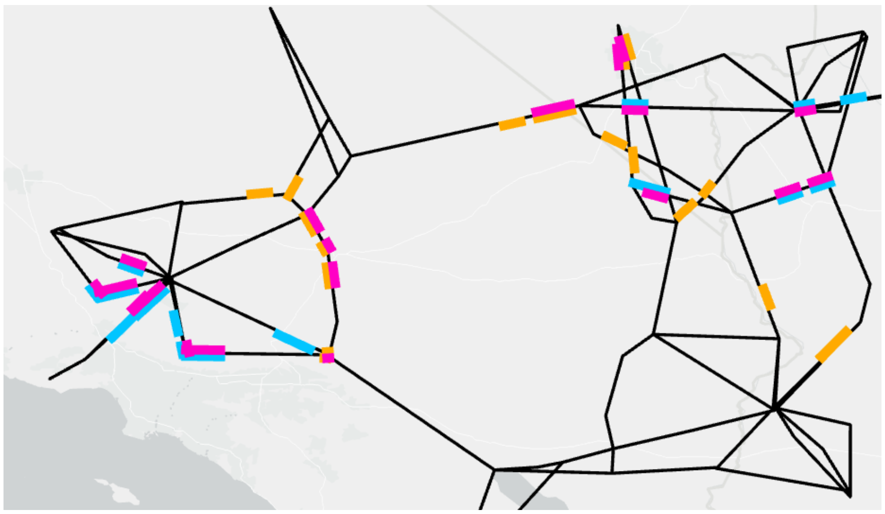
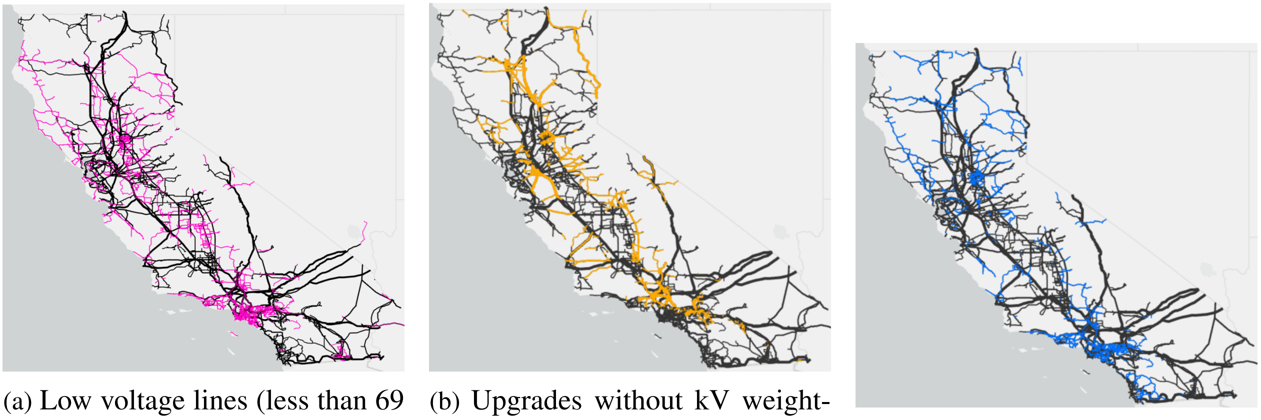


Figure 4.4: Plot of RTS-GMLC 10-km line segments selected for upgrade for the cumulative (blue), maximum (orange), and trade-off (magenta) problem variants. Segments are offset for visibility.

and maximum may be useful to reduce both high point-wise and high cumulative risk.

Finally, we included the impact of transmission voltages on the wildfire risk analysis. To incorporate the fact that the relative probability wildfire ignitions is three times higher for distribution grid lines, we assume that $\pi_{i,s} = 1 \cdot \pi$ for transmission lines and $\pi_{i,s} = 3 \cdot \pi$ for distribution lines. We use $V_{\text{dist}^{\text{max}}} = 69$ kV as a cutoff value to distinguish between transmission and distribution lines. The optimization problem is solved twice for the CEC transmission system, once without any voltage-based weighting, and once with kV-weighting.

We tested this improvement on a geographically accurate California grid (described below in Section 3.1) which exhibits a range of line voltages from 33 to 500 kV. The system serves as a large-scale example with roughly 6800 transmission and distribution lines. The same budget of 15 billion USD is used for both cases. The lines selected for upgrade by these two solutions are shown in Fig. 4.5b and Fig. 4.5c. Totals of 2536 and 3748 lines are selected for upgrade for the unweighted and kV-weighted solutions, respectively. There are 1373 lines that are exact matches, amounting to approximately 54% of the unweighted and 37% of the kV-weighted upgraded lines.



(a) Low voltage lines (less than 69 kV)

(b) Upgrades without kV weighting

(c) Upgrades with kV weighting

Figure 4.5: Plot of California Energy Commission lines showing (a) distribution lines (less than 69 kV) in pink (b) upgrades selected without kV weighting in yellow and (c) upgrades selected with kV weighting in blue.

4.2 Transmission-Distribution Dynamic Co-simulation of Electric Vehicles Providing Grid Frequency Response

This section explores decision making with uncertainty, specifically regarding EV integration and dynamic grid simulations. More details can be found in [97]

Many countries have set goals toward or are planning to reach a carbon emissions-free power sector and to reduce carbon emissions of the transportation sector during the next two decades. As a result, an increasing number of electric vehicles (EVs) and charging infrastructure will be deployed in the transmission and distribution networks. Because inverter-based resources—such as EVs, distributed photovoltaics (DPV), and energy storage—are connected to the grid through power electronic devices, the total inertia of the system is decreasing and making the system more vulnerable to frequency fluctuations [115]. Different control strategies for the generation units and storage can be adopted to restore the frequency response by providing real power support [116, 117]. These frequency regulation services, including both primary frequency response (PFR) and secondary frequency response (SFR) [118], can balance the system total load and generation.

EVs, equipped with a battery, have the capability and flexibility to provide fast frequency response, including PFR and SFR, to help mitigate system frequency fluctuations and to enhance system frequency stability; however, this vehicle-to-grid (V2G) frequency regulation provision may impact both the bulk power system frequency response and the local distribution network voltage profiles. Because the charging infrastructure is usually designed to sustain charging of EVs at the rated power, the V2G discharging for frequency regulation could increase local voltage and lead to overvoltage violations.

To synthetically study the impacts of the EVs' frequency regulation on both the bulk power system and a distribution network, this paper [97] leverages a transmission-and-distribution (T&D) dynamic co-simulation model that can simultaneously perform the bulk system dynamic simulation and distribution power flow analysis. The coordination between EVs and other DERs, such as DPV, for frequency regulation is studied. Multiple participation strategies for the frequency response from EVs are investigated. The main contributions of this paper can be summarized as follows:

- An innovative EV dynamic model considering EV owners' participation willingness has been developed and added to the T&D dynamic co-simulation model [118] to enable the analysis of the frequency response from EVs.

- The coordination between the EV and DPV frequency regulation is studied, which provides guidance for future coordination optimization of DER grid services.

- The impacts of EV frequency regulation are analyzed, including PFR and SFR on bulk system frequency response and distribution voltage. Multiple participation strategies of the frequency response from EVs are investigated.

4.2.1 EV Model and Simulation Framework

EV Model

Plug-in EVs have promising capabilities to provide several T&D grid services [119]. Because EVs are essentially inverter-based resources, we developed an EV dynamic model based on the Western Electricity Coordinating Council PVD1 model [120]. Here, we added 1) a parameter *Pcap* that models the participation strategies of EV; 2) the state-of-charge (SOC) related blocks that decide the current flowing in and out of the battery, as shown in Fig. 4.6. Note that a generic model of PFR is also included in Fig. 4.6. The overall dynamic model can represent general EV battery behaviors, which is added to ANDES [121], a grid electro-mechanical dynamics tool.

More specifically, *pcap* added in this model limits the participation of an EV to provide frequency regulation. *pcap* is in range [-1, 1], and the meanings of representative values are explained here. When *pcap* =-1, the EV's maximum power is 100% charging, which means that the EV cannot provide PFR and SFR. *pcap* =1 means that the EV's maximum power is 100% discharging, and the EV can change its status from charging to discharging to provide PFR and SFR. Similarly, *pcap* =-0.5 and 0.5 mean that the EV's maximum power is 50% charging and 50% discharging, respectively. *pcap* =0 represents that the EV's maximum power is 0, which means that the EV is not charging or discharging.

PFR

PFR uses droop control, i.e., when the frequency deviation is larger than a PFR deadband, the EV changes its active power output accordingly. An additional power output, P_{drp}, is added to the

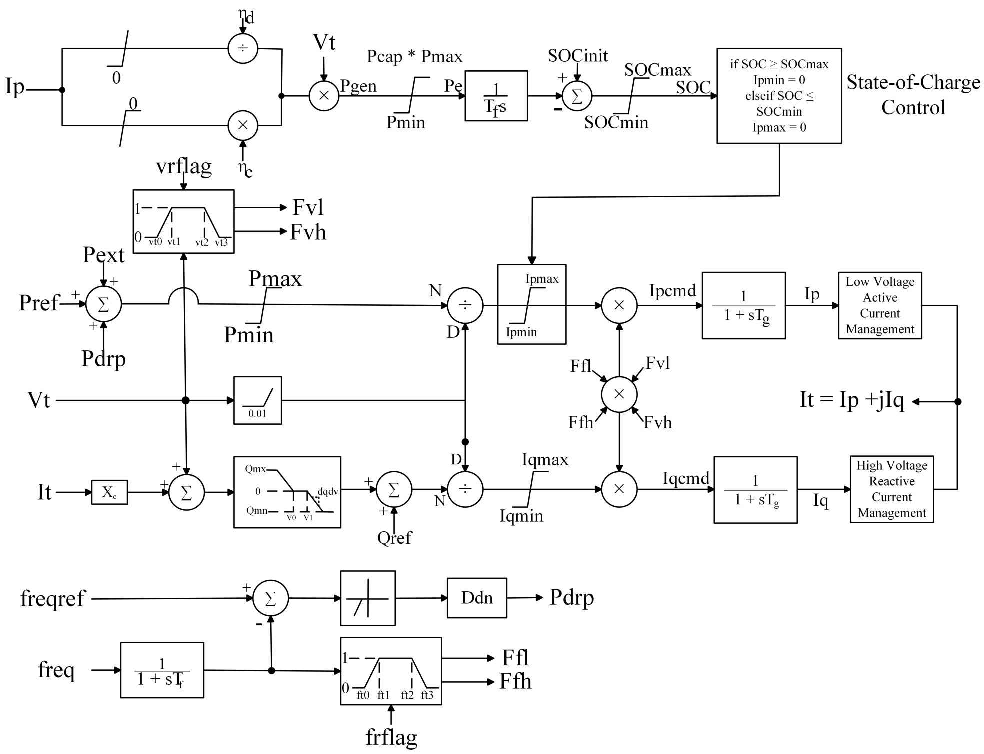


Figure 4.6: Block diagram for the EV dynamic model including PFR.

generation output:

$$P_{drp} = \begin{cases} \frac{(60-db_{UF})-f}{60} D_{dn} & \text{if } f < 60 \\ \frac{f-(60+db_{OF})}{60} D_{dn} & \text{if } f > 60 \end{cases} \tag{4.2}$$

where db_{UF} and db_{OF} are the underfrequency and overfrequency deadband, respectively; and D_{dn} is the per-unit power output change to 1-p.u. frequency change (frequency droop gain).

SFR

SFR [122] is enabled by an automatic generation control (AGC) model that includes two components: an area-level (assuming one area in this paper) estimation of the area control error (ACE) and a plant-level control that receives the SFR reference power, P_{ext}, for each plant. ACE represents the system generation and load imbalance. ACE is calculated as:

$$ACE_{tt} = 10B(f_{reqm,tt} - f_0) \tag{4.3}$$

where tt is the AGC time interval index; ACE_{tt} is the ACE at the AGC interval tt; $f_{reqm,tt}$ is the measured system frequency at the AGC interval tt; f_0 is the system reference frequency; and B is the frequency bias in $MW/0.1Hz$. After a frequency error tolerance deadband, f_{db}, a proportional-integral (PI) control is applied on the ACE signal to calculate the control variable, $u(t)$ (i.e., AGC

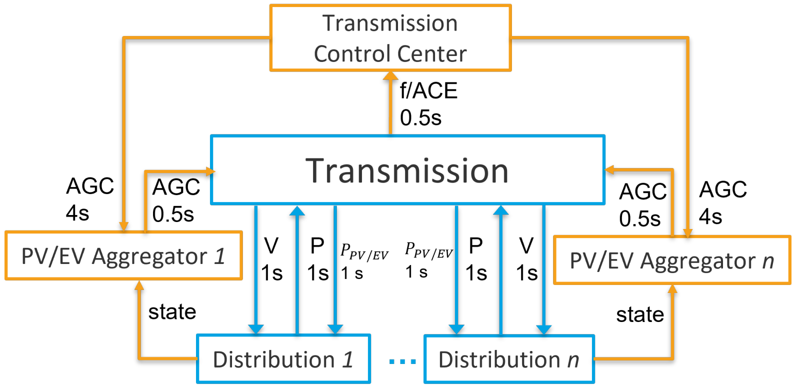


Figure 4.7: Simulation components with information exchange

signal); K_P and K_I are the coefficients of the AGC PI controller:

$$u(t) = -K_P ACE - K_I \int ACE. \tag{4.4}$$

The AGC signals are normally updated every 4 s in the field. The output from the PI controller is allocated to each AGC generator considering the unit's participation factor, resulting in the final AGC control reference for each unit. Note that the participation factor of each unit is decided by a real-time economic dispatch that is normally updated every 5 minutes. Each EV's participation factor can be updated by the corresponding EV aggregator and/or under a different time interval based on the local aggregator's optimization.

T&D Dynamic Co-simulation Platform

This section introduces the T&D dynamic co-simulation framework for studying effect of EVs on frequency response. The backbone of this framework is developed in [118]. The co-simulation framework is based on the HELICS platform and the open-source power system simulator ANDES and OpenDSS [123]. HELICS is an open-source, cyber-physical co-simulation framework for energy systems. Following are a few key concepts of HELICS that are relevant here: federates, brokers, simulators, and messages; for more details, see [124].

The developed EV component enables EV frequency response studies. Assume that the overall system comprises a transmission system; a control center; and an EV aggregator and a photovoltaic (PV) aggregator for each load bus, as shown in Fig. 4.7. The transmission system sends the system frequency and the ACE signals to the transmission control center every 0.5 second, where the AGC signals are calculated with the PI controller and sent to the EV and PV aggregators every 4 seconds. This setup is modeled in HELICS, where the transmission simulation federate uses ANDES, and the distribution quasi-static time-series power flow uses OpenDSS.

More information, including the case studies and results, can be found in [97].

4.3 Design Considerations for Operational Power System Simulation Scenarios

This section describes the scenario generation work presented in [98]

In addition to grid planning and operations, power system simulations can play an important role in formal engineering education, on-the-job training, and power system research. These simulations are well established, continuously evolving, and can impact the future of smart grid development [125]. Industry has long used operator training simulators (OTSs) and dispatcher training simulators (DTSs) to train their personnel [126]. Numerous tools exist for demonstrating the operation of the electric grid [127–133]. These simulators can emulate real-world systems and historical data can be used to design them to optimally train employees, on the system they will be operating. However, due to the sensitive nature of our power system infrastructure, students and researchers often never get the opportunity to experience one of these simulations.

Recent developments in phasor measurement unit (PMU) time-frame interactive simulation environments [133, 134], such as the Dynamic Simulator (DS), as well as the accompanying creation of large-scale, realistic synthetic grid systems [9, 48, 135, 136] have made these simulations accessible to students and researchers alike. Given these sophisticated tools and models, the challenge becomes how to use them effectively for purposes of training and education, for a variety of audiences. Short duration, steady-state or single contingency scenarios can be a good place to start. Examples include the textbook type of exercises in [137]. These help teach concepts such as contingency and sensitivity analysis, and other basic power system principles. This paper [98], however, focuses on the development of longer, more complicated, real-time, interactive simulation scenarios which are meant to mimic the role an operator would play in a control center.

Scenario design has been a key part of power systems operations training, associated with DTS's [138] and OTS's [139]. A scenario can be described in simple words as "the running of an event group with a base case" [138]. In [140], the instructor is responsible for applying and changing the scenarios at certain intervals, which may include load changes, faults, change in generator voltage setpoints, etc. In [139], a heuristic method was developed to automate the creation of scenarios to match the training goals along with trainee experience, adopting methods from artificial intelligence. However, this is computationally intensive and relies on the collection of a large amount of actual, power system operational data and is more suited for industry applications, where the data is usually proprietary. This paper aims to address some of these gaps by creating these scenarios based on publicly available grid models, so that the scenarios are not proprietary or protected. In addition, not only are all these scenarios pre-programmed but the DS simulating them runs a full transient stability simulation in real-time, as opposed to the above examples.

Longer, more "realistic" simulation scenarios could be of immense value to both students and researchers. Students can gain a feel for what it is like to operate a power system during both normal and emergency situations, in real-time. Researchers can use these simulations to evaluate the effectiveness of new visualizations, interfaces, operator tools, or training techniques. They can measure the impact of human factors on different aspects of power grid operations.

While the value of these operational scenarios is clear, the design possibilities are virtually endless. This paper [98] describes the design of real-time, interactive, operational scenarios of realistic grid system, with three scenarios detailed. The first is a single-user voltage control simulation of a large system, with a dynamic load profile. The second is a multi-user simulation, designed to mimic a typical control room, in which users are controlling a medium sized sub-set of a larger system, also with a dynamic load profile. The final scenario is a single-user simulation of a small system during a geomagnetic disturbance event (GMD). The full paper, [98] describes the development and features of all the scenarios.

4.4 A study of Cyber-Attack Resilience in a DER-Integrated Synthetic Grid Based on Industry Standards and Practices

The objective of this section (and detailed more in [99]) is to develop the required simulation infrastructure and methodology to analyze the stability issues and provide holistic study of impacts of malicious cyber-attack on power systems with a high penetration of DERs. Furthermore, the goal is for the outcomes of this section to be used to help propose an attack-resilient framework for critical infrastructure and provide a quantifiable resilience rubric for secure integration of DERs. Specially, the section aims to present and simulate the architecture involved with DER integration, the cybersecurity challenges introduced due to the integration, and steps required to mitigate the challenges and increase resilience to the infrastructure. Hence, this work has developed a Simulink model of Solar PV [141]. This has been coupled with a transmission grid as well as a communication network to create a coupled infrastructure. This work described can open the gateway to different analyses and red team/blue team assessment of DER systems in real time in an emulation platform.

The case used to design this simulation is from a dataset called syn-austin-TDgrid-v03, which is a highly detailed synthetic electric grid data set for combined transmission and distribution systems [142]. This case represents a synthetic grid version of the Travis County of central Texas with 140 substation and 448 feeders [142]. This case was selected, because this detailed T&D system facilitates analysis of coupled infrastructure as per the goal of this section.

In designing the simulation, the development included a Simulink model file with .slx extension which is block diagram model of 30KW of PV array based on a specific manufacturer design, SunPower SPR-305E-WHT-D. In attempt to model the case as realistic as possible, the model was designed to follow the rooftop solar generation capacity and geographical location data made available by Austin Shines Project and Pecan Street Inc. for Austin, Texas. The generation pattern of the DER in this case has been modelled according to the data of participating households in Austin Shines project [143].

To provide the cyber aspect to the system, a directional network was created, connecting all the DERs in the system through aggregators. All three files combined provide an infrastructure for studying voltage stability issues caused by cyber threat to DERs in the grid.

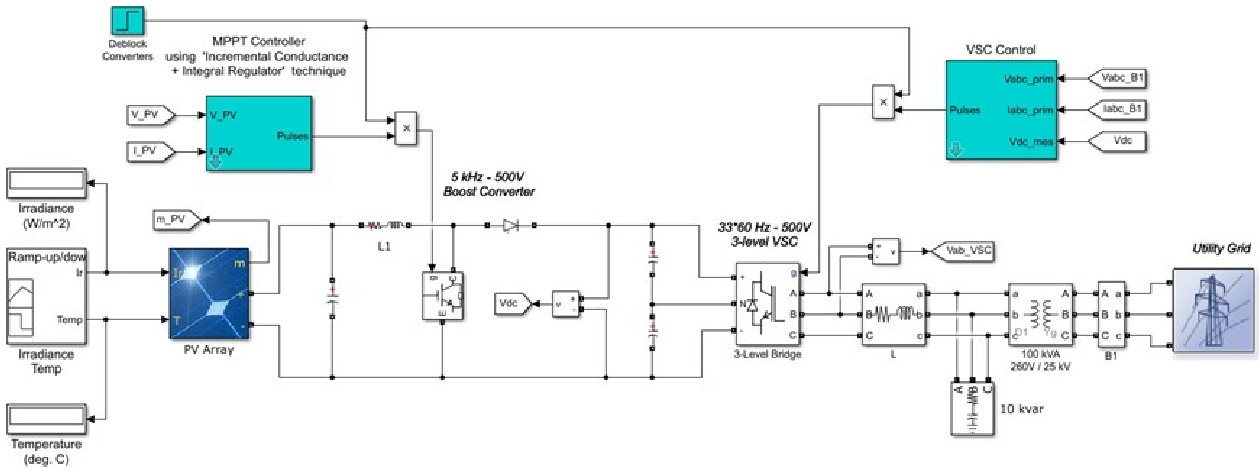


Figure 4.8: Model of single PV array connected to grid.

The work in this section uses PowerWorld to perform the modeling and analysis of the transmission system, OpenDSS for the location of substations and connection nodes available in provided distribution system data, and MATLAB and Simulink for the modelling of PV system. These software packages were then tied together through scripts of Python 3.

4.4.1 Grid Data

The grid data used in this section is based on research on synthetic grids. Synthetic grids are realistic and fictional power network models. They include detailed representation of generators, loads, transmission lines, and transformers [9]. The synthetic grid is based on geographically sited, publicly available data and statistics about the physical grid. This allows co-simulation and coupled infrastructure studies. Test Case

Test Case

In the section, we have utilized the highly detailed Synthetic Electric Grid Data Set for combined transmission and distribution systems [142] of Travis County, TX as the test case to demonstrate the coupling of power and DER networks. It includes the city of Austin and surrounding areas in central Texas. This data set serves 307,236 customers loads with total system peak of 3,254 MW. There are in total 140 substations in the system, with 69 kV and 230 kV nominal voltage level. This data set includes a mix of 448 rural, suburban, and urban feeders, and 132,406 distributed transformers. There are in average 5.3 consumers per distribution transformer; the distribution transformer capacities are in the range 10-1500 kVA, and ANSI ratings are used for the maximum allowed voltage range. The synthetic distribution network models diversity in the following terms: 1) There are urban, suburban, and rural circuits in the data sets adapted to the different characteristics and dispersion of consumers. In particular, the urban/suburban and rural circuits have different design

targets, for example, network length and reliability. 2) Several distribution nominal voltage levels

Table 4.2: The key statistics of the synthetic system Travis160

Statistics	Quantity
Customer loads	307,236
Generator units	39
Feeders	448
69 kV transmission lines	229
230 kV transmission lines	34
Transmission buses	160
Distribution electric nodes	1,654,691

	A	B	C	D	E	F
1	Area	TransmissionSub	DistributionSub	NodeName	lon	lat
2	p39u	p39uhs4_1247_69	p39uhs4_1247	p39udm7	-89.381	29.76638
3	p39u	p39uhs4_1247_69	p39uhs4_1247	p39udm8	-89.3809	29.76628
4	p39u	p39uhs4_1247_69	p39uhs4_1247	p39udm9	-89.3816	29.76661
5	p39u	p39uhs4_1247_69	p39uhs4_1247	p39udm10	-89.3758	29.77101
6	p39u	p39uhs4_1247_69	p39uhs4_1247	p39udm11	-89.3756	29.7711
7	p39u	p39uhs4_1247_69	p39uhs4_1247	p39udm13	-89.3754	29.77101
8	p39u	p39uhs4_1247_69	p39uhs4_1247	p39udm14	-89.3751	29.77111
9	p39u	p39uhs4_1247_69	p39uhs4_1247	p39udm18	-89.3753	29.77076

Figure 4.9: Connection of the transmission vs the distribution node based on the Voronoi polygons method.

are considered, specifically 4kV, 12.47kV, and 25kV. 3) Several approaches for voltage management are considered: voltage regulators and/or capacitor banks. 4) The loading of the network components depends on the discrete network components available in the input catalog. It is important to note that while the load is realistically modeled, the electric network that supplies the load in this synthetic test case is intentionally designed to be different from the actual system on the same geographic footprint. This prevents the synthetic data set from revealing critical energy infrastructure information, but still provides the users realistic test cases to develop techniques that can be applied to the real system. This test case is publicly available for download at [**?**].

Substation Service Area

Substation service areas are defined to simplify the mapping of DER generations from a distribution node to the transmission-level substation and to provide an understanding of the geographic service of the system. Establishing the service territory of each transmission substation leverages the geographic data on the synthetic system as well as the topology of the distribution system in the Travis160 synthetic case and uses Voronoi polygons to establish tessellating service territories with the electric model's nodes central to each region.

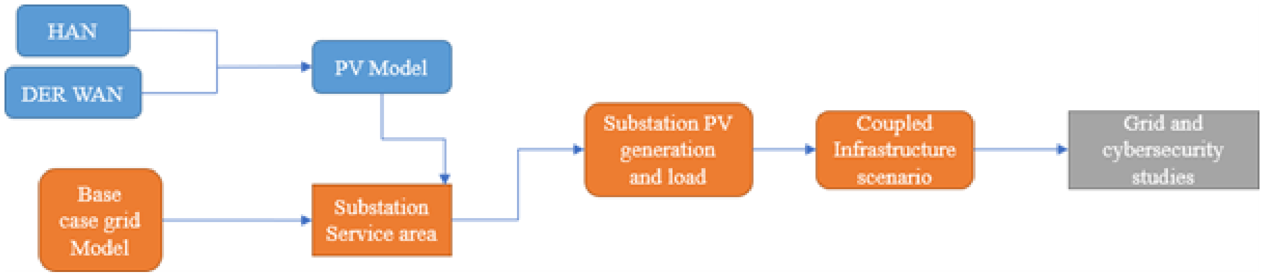


Figure 4.10: System architecture diagram for the study.

The service area mapping procedure is summarized below: 1) Select a transmission-level substation, 2) Identify which distribution feeders correspond to the selected substation, 3) Obtain geographic coordinates of identified distribution feeder nodes, 4) Create Voronoi polygons to represent the reach of each distribution node, 5) Aggregate Voronoi polygons to represent the selected transmission-level substation's service area, 6) Repeat steps 1 through 5, iterating through transmission level substations.

If the distribution system topology is not made available, service areas can be approximated by creating Voronoi polygons [144] for each of the transmission-level substations. The parsed data for transmission and distribution is available in Figure 4.9. Where each row provides info for corresponding transmission substation, distribution substation, geographical location of each DER node calculated on Voronoi polygon algorithm [144]. Thus, this data can be used to correctly map the DERs to transmission node in synthetic grid based on the corresponding distribution node data generated from Voronoi algorithm.

4.4.2 Case Studies and Results

This section provides the insight into the steps employed for the study. First, it describes parameter correction and simulation of DERs based on data from Austin SHINES. The second step involves the mapping of DER network to transmission system. The third step is to generate the communication network for the DER network in system. Finally, studies are done executing cyber threat scenarios on the DERs to study the impact on grid. The system architecture is displayed in Figure 4.10.

The city of Austin gets fixed tilt sunlight hours (the amount of hours of sunglight a fixed tilted non-tracking solar panel receives) of 5.3 hours per day and averages 4.0- 4.5$\frac{\text{kWh}}{\text{m}^2}$ of solar Irradiance [145]. Hence for the simulation, solar irradiance was assumed as displayed in Figure 4.11. The aggregate Irradiance for the day comes out to be 4.3$\frac{\text{kWh}}{\text{m}^2}$ to make it as realistic as possible.

The other constant required for simulation is hourly temperature. And the temperature used in the study is between 77°F to 90°F to simulate an average in the location. Figure 4.12 displays the fluctuation of temperature used in the simulation. The data set generated here is of 30000-time steps, spread over 24 hours.

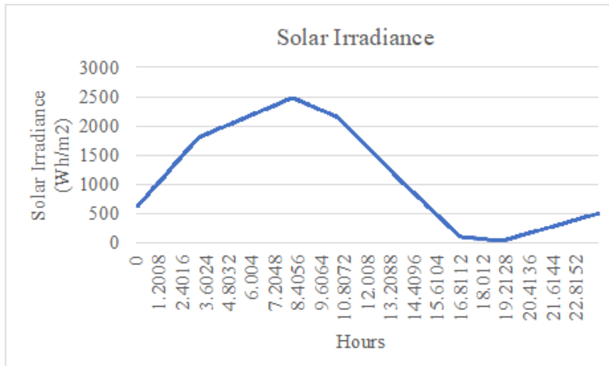


Figure 4.11: Solar Irradiance used for simulation of DER.

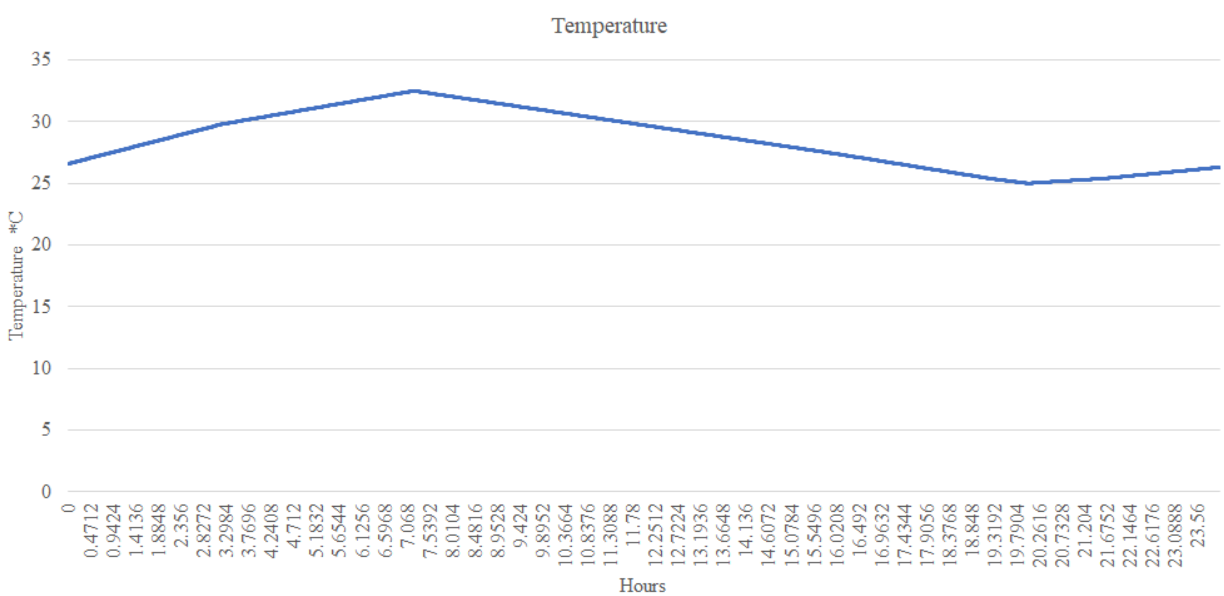


Figure 4.12: Temperature used for simulation of DER.

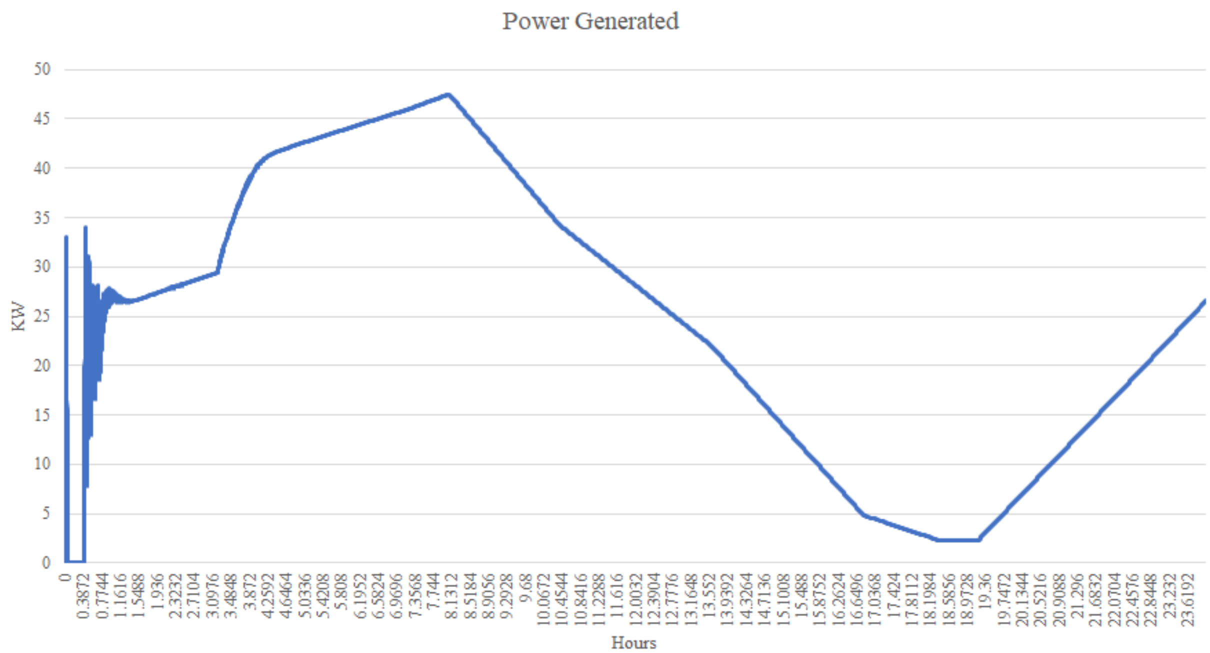


Figure 4.13: P_{mean} generated by single PV array during 24-hour simulation.

The DER nodes are mapped to the transmission system according to the substation service areas established in 4.4.1. If the DER node falls within the geographic footprint of a substation, it indicates that its most proximate distribution point of interconnection would aggregate to the specified transmission-level substation and thus, its load is best represented as an addition to the identified transmission-level substation. The DERs are mapped as generation, where the dataset ensures that generation is variable on a schedule.

Figure 4.13 displays the variable generation of PV based on parameters provided to the model. The maximum generation is 46.8KW which occurs during afternoon duration of the simulation. For 30% penetration, a total of 182.52MW of generation is scheduled through PV arrays. Assuming each DER cluster consists of 100 rooftop PV arrays. Thus, the total open generation available for the case is 3517.78MW through traditional sources and 182.52MW from attached DERs, making them a critical generation unit for the system. The schedules and schedule subscription feature available in PowerWorld are used here to update the load and generation of mapped DERs to the system.

Network for the DERs

To simulate a real-time network associated with DERs in the Austin power system, a primitive network graph methodology using vertices and edges, that represents a cyber-physical network of the county is used. The vertices denote network nodes or physical devices, and the edges denote a communication link between the nodes. Vertices and edges have attributes. A vertex that is a DER

Figure 4.14: Visualization of generated network graph for mapping of DER in the communication network

with router connection could have attributes such as the IP address, configuration, and generation information. Attributes of an edge could include, for instance, the protocol used in that link, such as DNP3. The NetworkX Model employed here has also parameters specifying the connection nodes, data, and location of particular DERs. This model provides PV generation data at geographically represented nodes each hour of the day. This information provides essential details for the coupled infrastructure: the connectivity to entire communication network. Hence, it allows the study of execution and defense of assumed threats to the system.

Due to limitation of NetworkX to execute any threat commands to power system simulation software, the connection between two softwares is an assumption in this study. The network graph generated by NetworkX is visualized in Figure 4.14.

The graph model has each of 39 DER cluster as a nodewith attribute of maximum MW generated among other data. They are connected in a directed graph fashion towards their respective aggregators. The DiGraph class provides additional methods and properties specific to directed edges. The aggregators are connected in bidirectional fashion and have ability to make multiple edges between two aggregators using MultiDiGraph class available in NetworX library [145].

A separate script is employed to simulate the network threat to aggregators in an ordered fashion. In the first step, two aggregators are taken out to simulate 25% blackout of DERs in the area. Then, all four aggregators are taken out of service to simulate 100% DER blackout in the area. The summary of the steps employed in this setup is displayed in figure 4.15.

A key question to address while considering increasing PV in grid how the grid changes along with it. The simulation environment allows the study of this impact on grid due to increased PV.

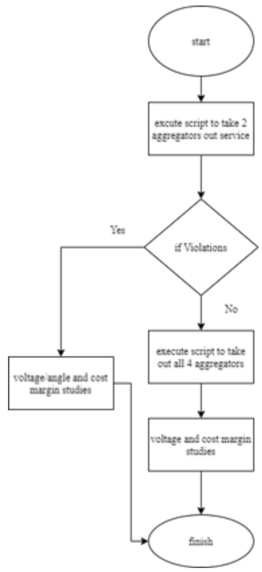


Figure 4.15: Steps involved in the study of voltage loss due to loss of DERs

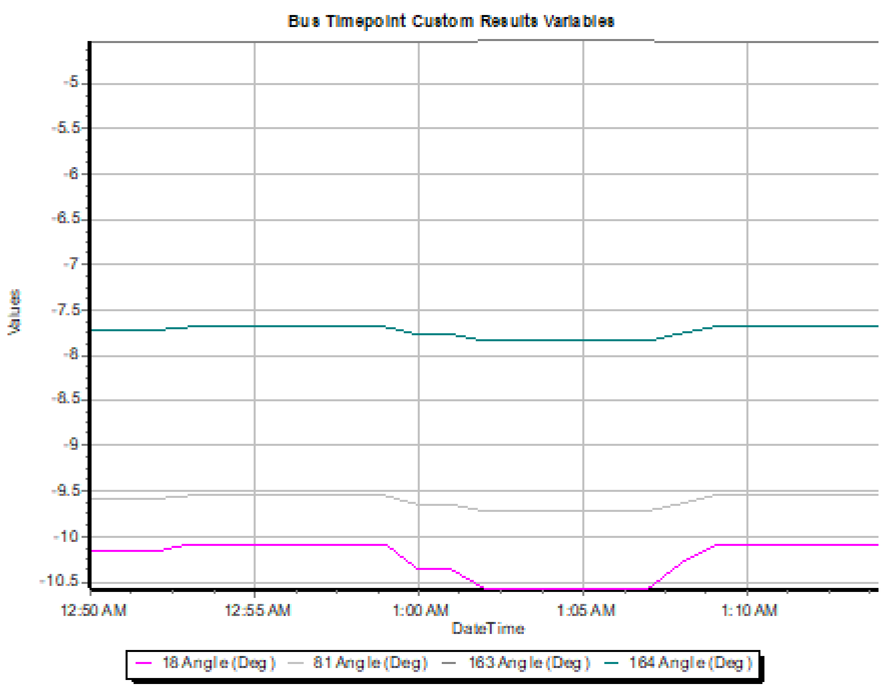


Figure 4.16: Voltage/angle during 25% DER blackout threat.

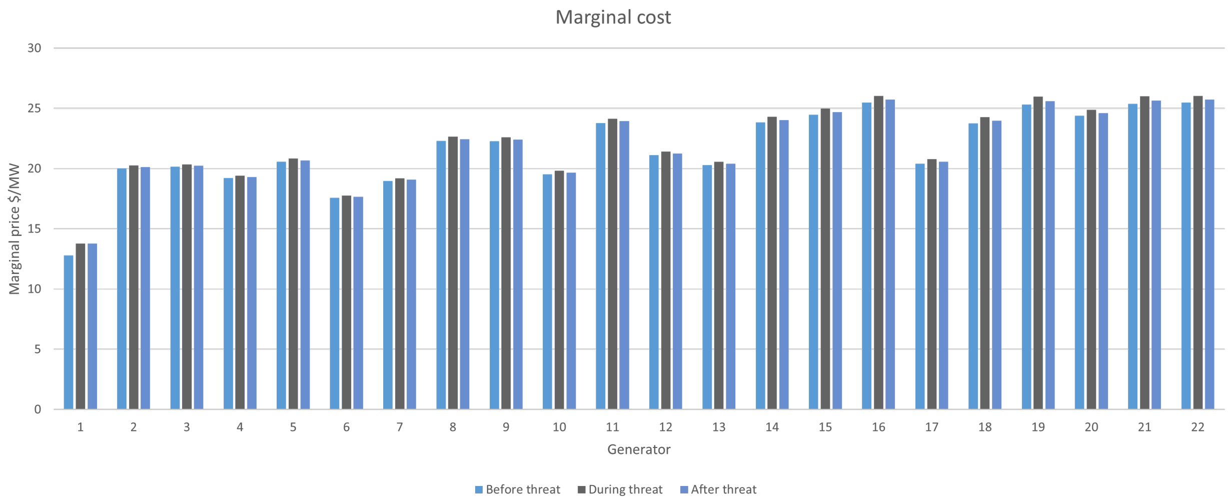


Figure 4.17: Marginal cost before, during and after 25% threat.

Results of the simulation as present below, proved that while 30% PV integration did not cause any violations, the most noticeable effects were change in energy cost, generation cost ($/hr) and marginal cost of production was decreased.

System loading studies were also performed to indicate any vulnerable points in case of extreme load scenarios. During the 24-hour period of the simulation, no bus voltage limits are violated.

Figure 4.16 shows the marginal drop in bus voltage angle due to 25% drop in generation by DERs in the system which accounts for 45.63MW only in a system of 3517.78MW. Hence there is nominal change in Marginal cost of generation for most of the traditional generators as displayed in figure 4.17.

While in case of 100% DER blackouts, there is significant voltage drop (p.u) with most of the buses attached to DERs generation units. The maximum is an almost 40% drop in voltage for bus 156.

Similarly, marginal cost analysis shows significant increase in cost/MW during the event of threat. Figure 4.18 displays the changes in marginal cost during and after the event for few generators mostly effected by the event, while Figure 4.19 provides insight into the per unit voltage drop experienced at each bus effected during the event.

Along with above mentioned results and analysis methods, there were different analyses performed on setup at different threat levels. And the data presented above is validated through the data provided by ERCOT in their study of renewables in the system [146]. The ERCOT DERAU1 model's ride through response abnormal voltage was modeled according to IEEE 1547-2018. ERCOT's study even through only considered 5% and 10% penetration for net load of 3150MW.

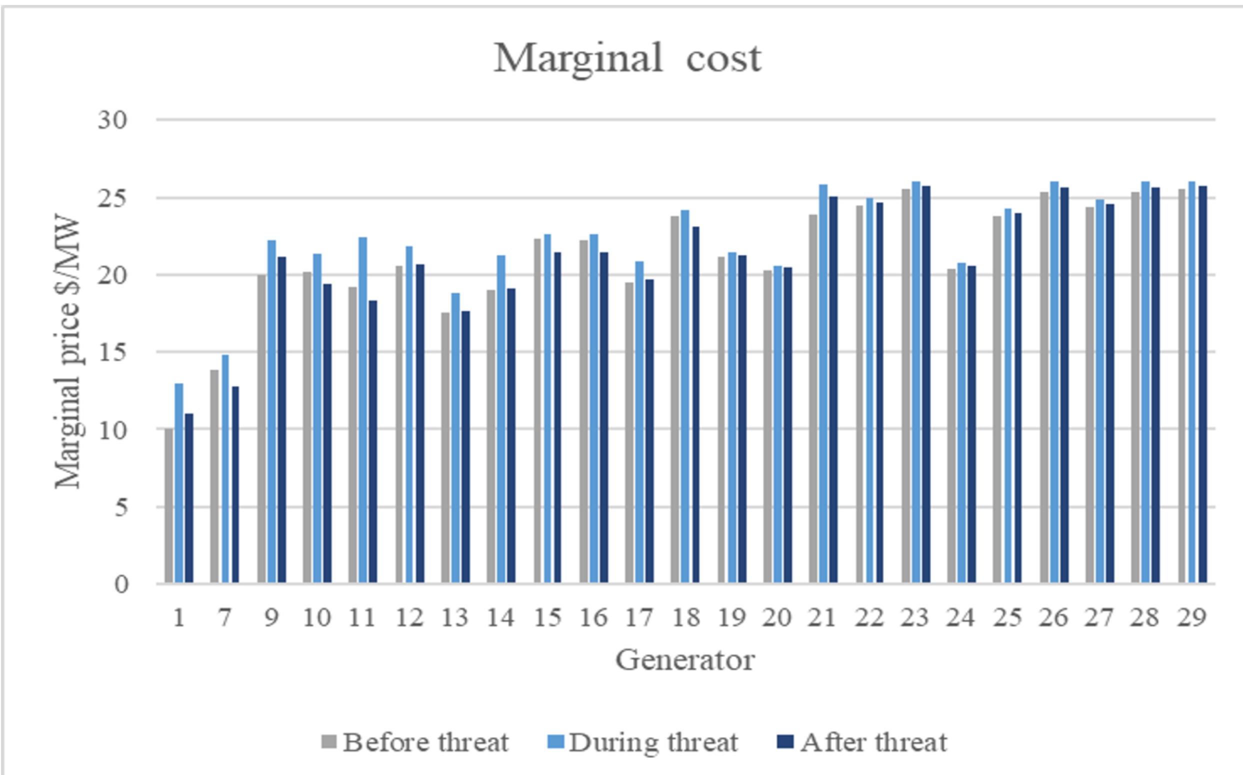


Figure 4.18: Marginal cost for generators before, during and after 100% threat event.

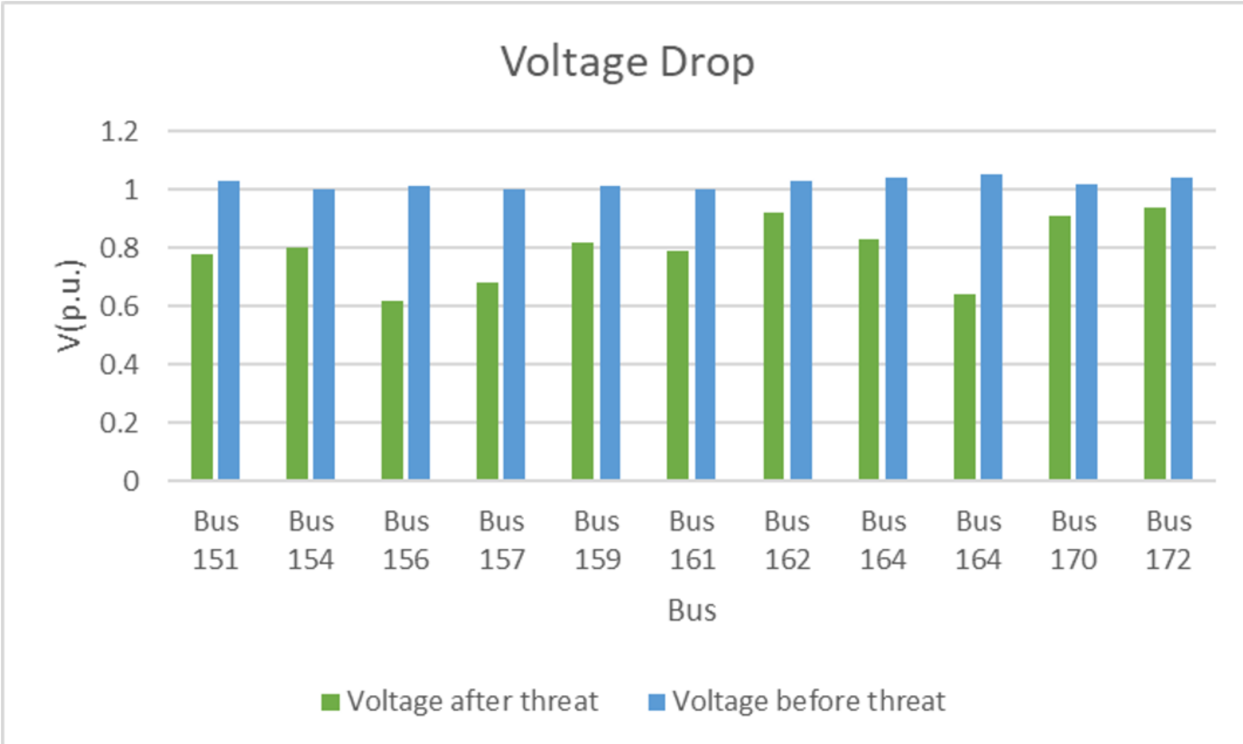


Figure 4.19: Bus voltage drop before, after 100% threat event.

summary and conclusion

This thesis introduced the interconnection of DERs to the power system, coupled with cyber infrastructure. The introduced DERs are then disconnected from the communication network due to a assumed cyber threat. This thesis also provides insight into design and deployment of photovoltaic arrays to the grid. It also discusses the most updated interconnection standards i.e IEEE 1547-2018, essential communication requirements for DERs like IEC 61850, standardization of cybersecurity for user and device authentication etc, being employed in industry across USA. The second part of thesis provide the vulnerability analysis of the various parameters involved the system and highlights the vulnerability of system to operational impacts introduced through loss of service of DERs in the system. This grid impact analysis is validated by ERCOT study [146] which also concluded that DER can negatively impact the net load serving capability of the grid (even at relatively low penetration levels) and needs to be explicitly modeled to tackle the potential reliability issues. The study also validates the interconnection discussion on ride through voltage, clearing time and other benefits of dynamic voltage support introduced by IEEE 1547-2003. In summary, the contribution of this thesis to existing literature is to:

1. Extend and fine-tune the design of photovoltaic array models for interconnection to existing grid systems

2. Develop infrastructure of DER interconnected to transmission system using Voronoi polygon

technique.

3. Provide insight into grid impacts such as voltage drop caused due to loss of DER in the system. By presenting a modeling pipeline from the PV modules to the cyber and physical transmission and distribution power system infrastructure, this thesis provides an approach for the realistic study the vulnerability of system towards newly introduced DERs.

4. Provide insight into the updated communication standards to monitor and control the DERs in real time and requirements for emergency command, threat assessment and utility-DER interactions employed in industry that bolster resilience in the interconnection.

5. Task 4: Expanding the Scope of Synthetic Grids for Coupling with Other Infrastructures

In the final task of this report, the researchers examined the impact of other infrastructure models on the eclectic system. Specifically, Section 5.1 examined the coupled electric grid and transportation networks and studied the impact of electric vehicle charging on the grid's performance [147]. Section 5.2 further studied the emission impacts of electric vehicle charging [148].

5.1 The Economic and Technical Impacts of Houston's Electric Vehicles on the Texas Transmission System: A Case Study

Due to maturing technology, declining costs, and increased support for clean transportation, electric vehicles (EVs) are on the rise. As of 2020, the transportation sector represented only around 2% of global electricity demand. However, recent studies show that by 2050, transportation is expected to account for 10% of total global electricity demand [149]. This trend toward increased electrification of the transportation sector requires extensive planning to prepare the electric grid for a variety of possible adoption scenarios. The impacts of transportation electrification varies depending on aspects of the adoption scenarios such as the penetration of EV integration and the charging models used. Therefore, there exists a need to model EV integration scenarios so that researchers can identify possible problems to various aspects of power grid planning and operations.

One area of particular interest pertains to identifying the infrastructure changes that are necessary to support an increased EV integration. Particularly, one identified impact of EV charging is a change to the peak demand under certain charging models [150]. These sudden changes in peak demand can lead to the line overloading as current increases to maintain power supplies. The line overloading and congestion results in huge changes in the locational marginal price (LMP) of electricity [151] and can lead to an accelerated component aging, increased resistive losses, and fire safety issues from overheating lines or transformers that impact the reliability of the components of the electric grid. However, if EV charging load schedule is encouraged during off-peak hours, LMPs may even decrease as a result of congestion prevention in peak hours [152].

Because many factors must be considered with the increasing integration of EV charging into the power grid, it is imperative that studies integrate realistic models of both the transportation and electric systems. This work relies on an established coupled infrastructure approach using detailed models of both a realistic electric grid and actual transportation network to analyze the impact of EV integration on line loading and LMPs of the transmission system for multiple levels of EV penetration. Publicly available data of transportation system is used for estimating EV charging patterns based on their type, location and schedules and the charging demand is integrated to a realistic but not real synthetic power grid that is created over the footprint of Texas, United States. This research provides a fundamental insight into the impact of incorporating electric vehicles into

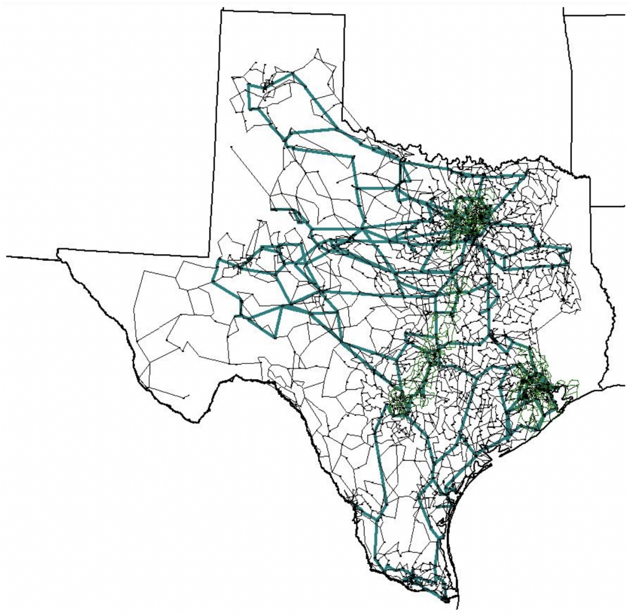


Figure 5.1: TX7k transmission system

a realistic large-scale electric system with more than 7,000 buses considering reliability and system costs.

5.1.1 Transmission System Modeling

In North America, electric grid models are considered critical energy infrastructure information (CEII) and access to those are restricted and detailed results often cannot be published. As such, this study leverages a synthetic grid that is created over the Texas footprint that is realistic enough to mimic actual grids. Synthetic grid models are publicly available at [153] and have been validated to be functionally similar to the built grids in North America [21] without compromising CEII. The development methodology of these grids is documented in [8, 9, 14, 30], and [154] details the inclusion of generator cost curve information, a feature of the synthetic grids which is essential for the performance of economic studies in this paper. The associated load time series are based on an estimated composition ratio of residential, commercial, and industrial customer load. Publicly available prototypical residential/commercial building, and industrial facility load time series are then aggregated to the buses through a heuristic optimization process [155], [156].

The transmission system used in this study is the TX7k network, a system comprised of nearly 7,000 buses geographically sited on the Electric Reliability Council of Texas (ERCOT) footprint using the same voltage levels as the built ERCOT system, shown in Figure 5.1. An overview of case information is provided in Table 5.1. This grid has a corresponding synthetic distribution system [142], the topology of which is leveraged in mapping EV loads to the transmission-level grid.

Table 5.1: System Information for TX7k Grid

Attribute	Value
Buses	6717
Transmission Voltage Levels	69 kV, 138 kV, 345 kV
Peak Load	80 GW
Generation Capacity	100 GW

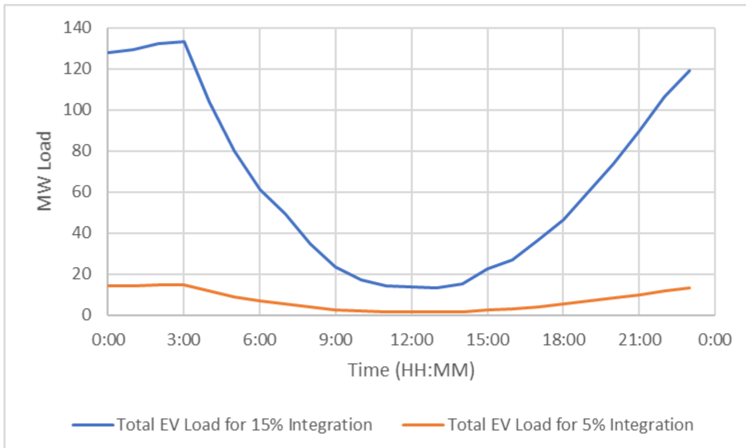


Figure 5.2: Charging profiles for 5% and 15% of EV integration

5.1.2 EV Load Modeling

Due to a presently low market penetration of EVs (in Texas in 2021, only 0.24% of vehicles registered in Texas were EVs [157]), there exists a lack of availability of widespread EV charging data. Thus, simulations are useful for generating EV charging data. The modeling of EV loads relies on an underlying transportation network model and traffic flow simulations coupled with charging behavior models. This modeling process was demonstrated in [144] and is applied in this paper for the greater Houston region.

A dynamic traffic assignment (DTA) model provides a mesoscopic analysis of traffic flow over a spatio-temporal resolution. The DTA model uses the transportation network and travel demand models to generate a trip trajectory and calculates on-road energy consumption of EVs. For a defined market penetration of EVs, trips are randomly assigned to be EV or non-EV trips. The vehicle range of those designated to be EVs is assigned based on the proportion of 100-mile, 200-mile, and 300-mile ranges from EV sales data [158].

Charging behavior was modeled with the goal of creating a realistic EV load profile. This was accomplished using a microscopic charging behavior model that accounts for characteristics of daily travel as well as various levels of anxiety of drivers. Thus, the resulting charging load incorporates variation based on time-of-day, remaining battery range, and trip characteristics. The outcome of

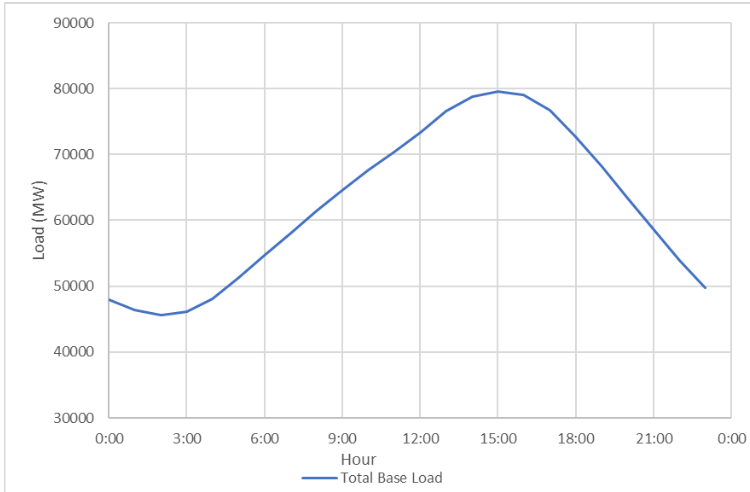


Figure 5.3: Load profile for the base case

using this behavior model is a charging load that is higher overnight, reflecting people charging their vehicles towards the end of the day when their batteries are more depleted after their daily travel.

The outcome of the EV modeling is a charging load time series at various locations in the synthetic system. These loads are incorporated to the electric grid model by the procedure developed in [144]. In summary, substation service areas are created using Voronoi polygons around the geographic location of the substation and the EV charging loads are mapped to the substations serving their respective locations. The EV charging loads are represented by loads added to buses within the substations in the electric grid model.

The case studies and results are presented in [147].

5.2 Generation Dispatch and Power Grid Emission Impacts of Transportation Electrification

Finally, research was conducted on the impact of electric vehicles (EVs) on generation dispatch and grid emissions, and is presented in [148]

In recent years, there has been rapid growth in the development and adoption of electric vehicle (EV) technologies, right from the vehicles themselves to charging infrastructures. A major driver behind this is the growing push for clean energy, which is offered by EVs with their zero tail-pipe emissions. However, there may be other sources of emissions attributed to the growing number of EVs. Specifically, the concern is with the emissions from generators in the bulk power grid that now to have supply the additional EV load. Hence, an environmental analysis of the benefits of EVs over internal combustion engine (ICE) vehicles should account for the increase in generator emissions for charging the EVs compared to the ICE tail-pipe emissions.

In regards to EV benefits, reference [1] provides a comprehensive review of existing literature on the economic benefits of EV integration to different energy market players, namely power generation companies, distribution system operators, EV aggregators, and end users. While economic benefits are important, environmental benefits are a primary function of EVs and should be evaluated. Though not explicitly discussed in [1], generating unit emissions are often considered as one of the generator "costs" that problems such as OPF or SCOPF seek to minimize. These emissions could be minimized with strategic charging strategies such as avoiding charging during peak times, and taking advantage of high renewable generation output periods.

With this in view, this paper [148] describes the impacts of EVs on generator emissions, considering different scenarios of EV penetration, charging strategies, generation mix, and wind curtailment. A synthetic grid representing the footprint of Texas is used as the case study. Hourly EV charging load for multiple cities in this footprint is considered in hourly SCOPF simulations. This charging load from on-road EV operation is developed based on a regional-level transportation simulation and charging behavior simulation, considering different EV penetration levels, congestion levels, and charging strategies. This EV load is then mapped to the appropriate grid substations leveraging the geo-mapping method developed in our prior work [2] to map nodes between transportation and grid networks. The previous paper also provided some preliminary results on grid impacts such as transformer loading and change in generation dispatch by fuel type due to the inclusion of EV charging load. The focus there was on a much smaller footprint, i.e. Travis Country, TX and the grid model used consisted of around 160 buses.

Hence, building on the work of [2], this paper [148] has the the following new contributions:

- Regional, statewide analysis for a comprehensive system study, i.e. modeling the entire transmission grid to account for realistic generation profiles and flows
- Multi-city EV load analysis
- Wind curtailment modeling in the dispatch problem
- Geographic visualizations of generator emission changes

More information can be found in [148], including the test system used for the analysis, the methodology right from calculating the EV load to mapping it to transmission substations, the dispatch and emission studies, and the results of this process for two different test systems.

References

[1] "California electric transmission lines," California Energy Commission, September 2020. [Online]. Available: https://gis.data.ca.gov/datasets/CAEnergy::california-electric-transmission-lines/about

[2] "Wildland Fire Potential Index," US Geological Survey. [Online]. Available: https://www.usgs.gov/fire-danger-forecast/wildland-fire-potential-index-wfpi

[3] Federal Energy Regulatory Commission (FERC), "Critical energy/electric infrastructure information (ceii) regulations." [Online]. Available: https://www.ferc.gov/enforcement-legal/legal/major-orders-regulations/critical-energyelectric-infrastructure-information

[4] R. Christie, "Power systems test case archive: 118 bus power flow test case," Tech. Rep., 1993. [Online]. Available: https://labs.ece.uw.edu/pstca/pf118/pg_tca118bus.htm

[5] D. Krishnamurthy, W. Li, and L. Tesfatsion, "An 8-zone test system based on iso new england data: Development and application," *IEEE Transactions on Power Systems*, vol. 31, no. 1, pp. 234–246, 2016.

[6] Q. Zhou and J. W. Bialek, "Approximate model of european interconnected system as a benchmark system to study effects of cross-border trades," *IEEE Transactions on Power Systems*, vol. 20, no. 2, pp. 782–788, 2005.

[7] T. Xu, A. B. Birchfield, K. M. Gegner, K. S. Shetye, and T. J. Overbye, "Application of large-scale synthetic power system models for energy economic studies," in *Proceedings of the 50th Hawaii International Conference on System Sciences*, 2017, Conference Proceedings.

[8] K. M. Gegner, A. B. Birchfield, T. Xu, K. S. Shetye, and T. J. Overbye, "A methodology for the creation of geographically realistic synthetic power flow models," in *2016 IEEE Power and Energy Conference at Illinois (PECI)*. IEEE, 2016, Conference Proceedings, pp. 1–6.

[9] A. B. Birchfield, T. Xu, K. M. Gegner, K. S. Shetye, and T. J. Overbye, "Grid structural characteristics as validation criteria for synthetic networks," *IEEE Transactions on Power Systems*, vol. 32, no. 4, pp. 3258–3265, 2017.

[10] J. M. Snodgrass, "Tractable algorithms for constructing electric power network models," Ph.D., 2021.

[11] U. S. C. Bureau, "Zip code tabulation areas," Tech. Rep., 2019. [Online]. Available: https://www.census.gov/geographies/reference-files/time-series/geo/gazetteer-files.html

[12] ——, "Census tracts," Tech. Rep., 2019. [Online]. Available: https://www.census.gov/geographies/reference-files/time-series/geo/gazetteer-files.html

[13] U. E. I. A. (EIA), "Form eia-860," Tech. Rep., 2019. [Online]. Available: https://www.eia.gov/electricity/data/eia860/

[14] F. Safdarian, A. B. Birchfield, K. S. Shetye, and T. J. Overbye, "Additional insights in creating large-scale, high quality synthetic grids: A case study," in *2021 IEEE Kansas Power and Energy Conference (KPEC)*, 2021, pp. 1–6.

[15] J. Snodgrass, S. Kunkolienkar, U. Habiba, Y. Liu, M. Stevens, F. Safdarian, T. Overbye, and R. Korab, "Case study of enhancing the matpower polish electric grid," in *Texas Power and Energy Conference*, 2022, Conference Proceedings.
[16] T. J. Overbye, K. S. Shetye, J. Wert, W. Trinh, A. Birchfield, T. Rolstad, and J. D. Weber, "Techniques for maintaining situational awareness during large-scale electric grid simulations," in *IEEE Power and Energy Conference at Illinois (PECI)*. IEEE, 2021, Conference Proceedings.
[17] T. J. Overbye, J. Wert, K. S. Shetye, F. Safdarian, and A. B. Birchfield, "Delaunay triangulation based wide-area visualization of electric transmission grids," in *IEEE Kansas Power and Energy Conference (KPEC)*. IEEE, 2021, Conference Proceedings.
[18] T. Overbye, K. Shetye, J. Wert, H. Li, C. Cathey, and H. Scribner, "Stability considerations for a synchronous interconnection of the north american eastern and western electric grids," in *Proceedings of the 55th Hawaii International Conference on System Sciences*, 2022, Conference Proceedings.
[19] "Texas A&M University Electric Grid Test Case Repository." [Online]. Available: https://electricgrids.engr.tamu.edu/electric-grid-test-cases/datasets-for-arpa-e-perform-program/
[20] V. Krishnan, B. Bugbee, T. Elgindy, C. Mateo, P. Duenas, F. Postigo, J.-S. Lacroix, T. G. San Roman, and B. Palmintier, "Validation of synthetic us electric power distribution system data sets," *IEEE Transactions on Smart Grid*, vol. 11, no. 5, pp. 4477–4489, 2020.
[21] A. Birchfield, E. Schweitzer, M. Athari, T. Xu, T. Overbye, A. Scaglione, and Z. Wang, "A metric-based validation process to assess the realism of synthetic power grids," *Energies*, vol. 10, no. 8, p. 1233, 2017. [Online]. Available: https://www.mdpi.com/1996-1073/10/8/1233/htm
[22] "Powerworld simulator." [Online]. Available: https://www.powerworld.com/
[23] FEMA, "United states transmission grid map," Tech. Rep. [Online]. Available: https://upload.wikimedia.org/wikipedia/commons/d/d4/UnitedStatesPowerGrid.jpg
[24] K. S. Shetye, T. J. Overbye, A. B. Birchfield, J. D. Weber, and T. L. Rolstad, "Computationally efficient identification of power flow alternative solutions with application to geomagnetic disturbance analysis," in *2020 IEEE Texas Power and Energy Conference (TPEC)*. IEEE, 2020, Conference Proceedings, pp. 1–6.
[25] T. J. Overbye and R. P. Klump, "Effective calculation of power system low-voltage solutions," *IEEE Transactions on Power Systems*, vol. 11, no. 1, pp. 75–82, 1996.
[26] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, "Matpower: Steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Transactions on power systems*, vol. 26, no. 1, pp. 12–19, 2010.
[27] R. Korab, *Description of case2746wop - Matpower*. [Online]. Available: https://matpower.org/docs/ref/matpower5.0/case2746wop.html
[28] NERC, "Standard tpl-001-4 - transmission system planning performance requirements," Tech. Rep., 2015. [Online]. Available: https://www.nerc.com/pa/Stand/Reliability%20Standards/TPL-001-4.pdf
[29] ——, "Tpl-007-4 - transmission system planned performance for geomagnetic disturbance events," Tech. Rep., 2020. [Online]. Available: https://www.nerc.com/pa/Stand/Reliability%20Standards/tpl-007-4.PDF

[30] A. B. Birchfield, K. M. Gegner, T. Xu, K. S. Shetye, and T. J. Overbye, "Statistical considerations in the creation of realistic synthetic power grids for geomagnetic disturbance studies," *IEEE Transactions on Power Systems*, vol. 32, no. 2, pp. 1502–1510, 2016.

[31] D. Elliott, M. Schwartz, S. Haymes, D. Heimiller, G. Scott, M. Brower, E. Hale, and B. Phelps, "New wind energy resource potential estimates for the united states (presentation)," National Renewable Energy Lab (NREL), Tech. Rep., 2011.

[32] N. Rhodes, L. Ntaimo, and L. Roald, "Balancing wildfire risk and power outages through optimized power shut-offs," *IEEE Transactions on Power Systems*, 2020.

[33] M. Ghamsari-Yazdel, M. Esmaili, F. Aminifar, P. Gupta, A. Pal, and H. Shayanfar, "Incorporation of controlled islanding scenarios and complex substations in optimal wams design," *IEEE Transactions on Power Systems*, vol. 34, no. 5, pp. 3408–3416, 2019.

[34] Z. Ma, H. Zhong, Q. Xia, C. Kang, Q. Wang, and X. Cao, "A unit commitment algorithm with relaxation-based neighborhood search and improved relaxation inducement," *IEEE Transactions on Power Systems*, vol. 35, no. 5, pp. 3800–3809, 2020.

[35] E. Cotilla-Sanchez, "Attack resilient data analytics for cascading failure mitigation," http://site.ieee.org/pes-cascading/files/2018/08/6-Cotilla-Sanchez.pdf, 2018, accessed 18 December 2021.

[36] M. Ensley, "Toward a theory of situation awareness in dynamic systems," *Human factors*, vol. 37, pp. 85–104, 1995.

[37] C. D. Wickens, "Situation awareness: Review of mica endsley's 1995 articles on situation awareness theory and measurement," *Human factors*, vol. 50, no. 3, pp. 397–403, 2008.

[38] C. Basu, M. Padmanaban, S. Guillon, L. Cauchon, M. De Montigny, and I. Kamwa, "Situational awareness for the electrical power grid," *IBM Journal of Research and Development*, vol. 60, no. 1, pp. 10:1–10:11, 2016.

[39] J. Wu, K. Ota, M. Dong, J. Li, and H. Wang, "Big data analysis-based security situational awareness for smart grid," *IEEE Transactions on Big Data*, vol. 4, no. 3, pp. 408–417, 2018.

[40] Y. Liu, W. Yao, D. Zhou, L. Wu, S. You, H. Liu, L. Zhan, J. Zhao, H. Lu, and W. Gao, "Recent developments of fnet/grideye—a situational awareness tool for smart grid," *CSEE Journal of Power and Energy Systems*, vol. 2, no. 3, pp. 19–27, 2016.

[41] T. J. Overbye, J. L. Wert, K. S. Shetye, F. Safdarian, and A. B. Birchfield, "The use of geographic data views to help with wide-area electric grid situational awareness," in *IEEE Texas Power and Energy Conference (TPEC)*. IEEE, 2021, Conference Proceedings.

[42] M. S. Dyrkacz, C. C. Young, and F. J. Maginniss, "A digital transient stability program including the effects or regulator, exciter, and governor response," *Transactions of the American Institute of Electrical Engineers. Part III: Power Apparatus and Systems*, vol. 79, no. 3, pp. 1245–1254, 1960.

[43] WECC, "Tpl-001-wecc-crt-3.2," 2019. [Online]. Available: https://www.wecc.org/Reliability/TPL-001-WECC-CRT-3.2.pdf

[44] NERC, "Glossary of terms used in nerc reliability standards," Tech. Rep., 2020. [Online]. Available: www.nerc.com/files/glossary_of_terms.pdf

[45] R. H. W. Lee, K. S. Shetye, A. B. Birchfield, and T. J. Overbye, "Using detailed ground modeling to evaluate electric grid impacts of late-time high-altitude electromagnetic pulses (e3 hemp)," *IEEE Transactions on Power Systems*, vol. 34, no. 2, pp. 1549–1557, 2019.

[46] K. S. Shetye, T. J. Overbye, H. Li, J. Thekkemathiote, and H. Scribner, "Considerations for interconnection of large power grid networks," in *IEEE Power and Energy Conference at Illinois (PECI)*. IEEE, 2021, Conference Proceedings.
[47] T. Overbye, K. Shetye, H. Li, W. Trinh, and J. Wert, "Feasibility assessment of synchronous operations of the north american eastern and western interconnections," *PSERC Report S-92G*, 2021.
[48] T. Xu, A. B. Birchfield, and T. J. Overbye, "Modeling, tuning, and validating system dynamics in synthetic electric grids," *IEEE Transactions on Power Systems*, vol. 33, no. 6, pp. 6501–6509, 2018.
[49] T. Overbye, D. Wiegmann, and R. Thomas, "Visualization of power systems," *PSERC Publication 02-36*, 2022.
[50] P. C. Wong, K. Schneider, P. Mackey, H. Foote, G. Chin Jr, R. Guttromson, and J. Thomas, "A novel visualization technique for electric power grid analytics," *IEEE Transactions on Visualization and Computer Graphics*, vol. 15, no. 3, pp. 410–423, 2009.
[51] T. J. Overbye and J. Weber, "Smart grid wide-area transmission system visualization," *Engineering*, vol. 1, no. 4, pp. 466–474, 2015.
[52] T. J. Overbye, E. M. Rantanen, and S. Judd, "Electric power control center visualization using geographic data views," in *iREP Symposium - Bulk Power System Dynamics and Control - VII. Revitalizing Operational Reliability*. IEEE, 2007, Conference Proceedings.
[53] F. P. Demello and C. Concordia, "Concepts of synchronous machine stability as affected by excitation control," *IEEE Transactions on power apparatus and systems*, vol. 88, no. 4, pp. 316–329, 1969.
[54] S. Dutta and T. J. Overbye, "Feature extraction and visualization of power system transient stability results," *IEEE Transactions on Power Systems*, vol. 29, no. 2, pp. 966–973, 2013.
[55] Y. Liu, "A us-wide power systems frequency monitoring network," in *IEEE Power Engineering Society General Meeting*. IEEE, 2006, Conference Proceedings.
[56] Y. Zhang, L. Chen, Y. Ye, P. Markham, J. Bank, J. Dong, Z. Yuan, Z. Lin, and Y. Liu, "Visualization of wide area measurement information from the fnet system," in *IEEE Power and Energy Society General Meeting*. IEEE, 2011, Conference Proceedings.
[57] W. Trinh, K. Shetye, I. Idehen, and T. Overbye, "Iterative matrix pencil method for power system modal analysis," in *Proceedings of the 52nd Hawaii International Conference on System Sciences*, 2019, Conference Proceedings.
[58] I. Idehen, B. Wang, K. Shetye, T. Overbye, and J. Weber, "Visualization of large-scale electric grid oscillation modes," in *North American Power Symposium (NAPS)*. IEEE, 2018, Conference Proceedings.
[59] J. Glover, T. Overbye, and M. Sarma, *Power System Analysis and Design*, 6th ed. Boston, MA: Cengage Learning, 2017.
[60] J. D. Weber and T. J. Overbye, "Voltage contours for power system visualization," *IEEE Transactions on Power Systems*, vol. 15, no. 1, pp. 404–409, 2000.
[61] T. J. Overbye and J. D. Weber, "Visualization of power system data," in *Proceedings of the 33rd Annual Hawaii International Conference on System Sciences*, 2000, Conference Proceedings.

[62] "Southwest power pool (spp) price contour map," Tech. Rep. [Online]. Available: pricecontourmap.spp.org/pricecontourmap/

[63] B. Shneiderman, *The eyes have it: A task by data type taxonomy for information visualizations*. Elsevier, 2003, pp. 364–371.

[64] C. Ware, *Information visualization: perception for design*. Morgan Kaufmann, 2019.

[65] F. P. Preparata and M. I. Shamos, *Computational geometry: an introduction*. Springer Science & Business Media, 2012.

[66] T. J. Overbye, G. Gross, M. J. Laufenberg, and P. W. Sauer, "Visualizing power system operations in an open market," *IEEE Computer Applications in Power*, vol. 10, no. 1, pp. 53–58, 1997.

[67] P. M. Mahadev and R. D. Christie, "Envisioning power system data: concepts and a prototype system state representation," *IEEE Transactions on Power Systems*, vol. 8, no. 3, pp. 1084–1090, 1993.

[68] J. Cohn, "When the grid was the grid: the history of north america's brief coast-to-coast interconnected machine [scanning our past]," *Proceedings of the IEEE*, vol. 107, no. 1, pp. 232–243, 2019.

[69] N. Cohn, S. B. Biddle, R. G. Lex, E. H. Preston, C. W. Ross, and D. R. Whitten, "On-line computer applications in the electric power industry," *Proceedings of the IEEE*, vol. 58, no. 1, pp. 78–87, 1970.

[70] A. L. Figueroa-Acevedo, "Opportunities and benefits for increasing transmission capacity between the us eastern and western interconnections," Thesis, 2017.

[71] Y. Li and J. D. McCalley, "Design of a high capacity inter-regional transmission overlay for the u.s," *IEEE Transactions on Power Systems*, vol. 30, no. 1, pp. 513–521, 2015.

[72] M. A. Elizondo, N. Mohan, J. O'Brien, Q. Huang, D. Orser, W. Hess, H. Brown, W. Zhu, D. Chandrashekhara, and Y. V. Makarov, "Hvdc macrogrid modeling for power-flow and transient stability studies in north american continental-level interconnections," *CSEE Journal of Power and Energy Systems*, vol. 3, no. 4, pp. 390–398, 2017.

[73] A. Bloom, J. Novacheck, G. Brinkman, J. McCalley, A. Figueroa-Acevedo, A. Jahanbani-Ardakani, H. Nosair, A. Venkatraman, J. Caspary, D. Osborn, and J. Lau, "The value of increased hvdc capacity between eastern and western u.s. grids: The interconnections seam study," *IEEE Transactions on Power Systems*, vol. 37, no. 3, pp. 1760–1769, 2022.

[74] Western Area Power Administration, "East/west ac intertie feasibility study," Tech. Rep., 1994.

[75] WAPA, "Western area power adminstration's first 25 yeas as a power marketing agency," Tech. Rep., 2002. [Online]. Available: https://www.wapa.gov/newsroom/Publications/Documents/25yr-history_2.pdf

[76] J. Caspary, J. McCalley, S. Sanders, and M. Stoltz, "Proposed eastern interconnection and western electricity coordinating council seams study," in *CIGRE US National Committee 2015 Grid of the Future Symposium*, 2015, Conference Proceedings, pp. 1–11.

[77] "Ieee test cases," IEEE Power and Energy Society. [Online]. Available: https://site.ieee.org/pes-tccwg/links-to-test-cases/

[78] J. M. Snodgrass, "Tractable algorithms for constructing electric power network models," Ph.D. dissertation, The University of Wisconsin-Madison, 2021.

[79] "University of wisconsin-madison synthetic electric network models," 2022. [Online]. Available: https://github.com/WISPO-POP/Synthetic-Electric-Network-Models

[80] C. Barrows, "Reliability Test System - Grid Modernization Lab Consortium," February 2021. [Online]. Available: https://github.com/GridMod/RTS-GMLC

[81] S. Babaeinejadsarookolaee, A. Birchfield, R. D. Christie, C. Coffrin, C. DeMarco, R. Diao, M. Ferris, S. Fliscounakis, S. Greene, R. Huang, C. Josz, R. Korab, B. Lesieutre, J. Maeght, D. K. Molzahn, T. J. Overbye, P. Panciatici, B. Park, J. Snodgrass, and R. Zimmerman, "The power grid library for benchmarking ac optimal power flow algorithms," *arXiv preprint arXiv:1908.02788*, 2019.

[82] S. Babaeinejadsarookolaee, J. Snodgrass, S. Acharya, S. Greene, B. Lesieutre, and C. L. DeMarco, "Comparison of real and synthetic network models of the western united states with respect to new realism measures," in *2021 IEEE Power and Energy Conference at Illinois (PECI)*. IEEE, 2021, Conference Proceedings, pp. 1–8.

[83] C. Barrows, A. Bloom, A. Ehlen, J. Ikäheimo, J. Jorgenson, D. Krishnamurthy, J. Lau, B. McBennett, M. O'Connell, and E. Preston, "The ieee reliability test system: A proposed 2019 update," *IEEE Transactions on Power Systems*, vol. 35, no. 1, pp. 119–127, 2019.

[84] N. Rhodes, L. Ntaimo, and L. A. Roald, "Balancing wildfire risk and power outages through optimized power shut-offs," *IEEE Transactions on Power Systems*, vol. 36, pp. 3118–3128, 2021.

[85] "Form eia-860 detailed data with previous form data (eia-860a/860b)," Energy Information Agency, 2019. [Online]. Available: https://www.eia.gov/electricity/data/eia860/

[86] "Production and curtailments data," California Independent Systems Operator (CAISO), 2019. [Online]. Available: http://www.caiso.com/informed/Pages/ManagingOversupply.aspx

[87] S. Babaeinejadsarookolaee, Tech. Rep., 2018. [Online]. Available: https://github.com/WISPO-POP/SyntheticElectricNetworkModels/tree/master/Synthetic%20Generator%20Cost%20Curves

[88] H. Zaininger, A. Wood, H. Clark, T. Laskowski, and J. Burns, "Synthetic electric utility systems for evaluating advanced technologies. final report," Power Technologies, Inc., Schenectady, NY (USA), Tech. Rep., 1977.

[89] L. Wang, J. McCalley, C. L. DeMarco, F. Akhavizadegan, A. Venkatraman, and J. Snodgrass, "Development of expansion planning methods and tools for handling uncertainty," Power Systems Engineering Research Center, Tech. Rep., May 2020.

[90] "Form 1 - electric utility annual report," Federal Energy Regulatory Commission, Mar 2022. [Online]. Available: https://www.ferc.gov/general-information-0/electric-industry-forms/form-1-electric-utility-annual-report

[91] "Form no. 715 - annual transmission planning and evaluation report," Federal Energy Regulatory Commission. [Online]. Available: https://www.ferc.gov/industries-data/electric/electric-industry-forms/form-no-715-annual-transmission-planning-and-evaluation-report

[92] S. Company, "Acsr aluminum conductor. steel reinforced. bare." Tech. Rep., 2012. [Online]. Available: https://overheadtransmission.southwire.com/wp-content/uploads/2017/06/acsr.pdf

[93] J. J. LaForest, *Transmission line reference book, 345 kv and above*, second edition. ed. Palo Alto, Calif.: General Electric Company. Large Transformer Division and Energy Systems and Technology Divisions Electric Power Research Institute., 1982.
[94] S. Company, "Acss aluminum conductor, steel supported. bare." Tech. Rep., 2017. [Online]. Available: https://overheadtransmission.southwire.com/wp-content/uploads/2017/06/acss.pdf
[95] C. Coffrin, R. Bent, K. Sundar, Y. Ng, and M. Lubin, "Powermodels.jl: An open-source framework for exploring power flow formulations," in *2018 Power Systems Computation Conference (PSCC)*. IEEE, 2018, Conference Proceedings, pp. 1–8.
[96] S. Taylor and L. A. Roald, "A framework for risk assessment and optimal line upgrade selection to mitigate wildfire risk," *Electric Power Systems Research*, vol. 213, p. 108592, 2022.
[97] Y. Liu, T. J. Overbye, W. Wang, X. Fang, J. Wang, H. Cui, and F. Li, "Transmission-distribution dynamic co-simulation of electric vehicles providing grid frequency," 2022.
[98] D. Wallison, M. Gaskamp, Z. Mao, Y. Liu, K. S. Shetye, and T. Overbye, "Design considerations for operational power system simulation scenarios," in *2020 52nd North American Power Symposium (NAPS)*. IEEE, 2021, Conference Proceedings, pp. 1–6.
[99] S. Tripathi, "A study of cyber-attack resilience in a der-integrated synthetic grid based on industry standards and practices," Thesis, 2022.
[100] "Statistics," National Interagency Fire Center, 2020. [Online]. Available: https://www.nifc.gov/fire-information/statistics
[101] S. Martinuzzi, A. J. Allstadt, A. M. Pidgeon, C. H. Flather, W. M. Jolly, and V. C. Radeloff, "Future changes in fire weather, spring droughts, and false springs across U.S. national forests and grasslands," *Ecological Applications*, vol. 29(5), 2019.
[102] A. D. Syphard and J. E. Keeley, "Location, timing and extent of wildfire vary by cause of ignition," *Int. J. of Wildland Fire*, vol. 24, no. 1, pp. 37–47, 2015.
[103] B. D. Russell, C. L. Benner, and J. A. Wischkaemper, "Distribution feeder caused wildfires: Mechanisms and prevention," in *65th Annual Conference for Protective Relay Engineers*, 2012, pp. 43–51.
[104] S. Jazebi, F. de León, and A. Nelson, "Review of Wildfire Management Techniques—Part I: Causes, Prevention, Detection, Suppression, and Data Analytics," *IEEE Trans. Power Delivery*, vol. 35, no. 1, pp. 430–439, 2020.
[105] M. Mohler, "CAL FIRE Investigators Determine Cause of the Camp Fire," May 2019. [Online]. Available: https://www.fire.ca.gov/media/5121/campfire_cause.pdf
[106] C. Miller, M. Plucinski, A. Sullivan, A. Stephenson, C. Huston, K. Charman, M. Prakash, and S. Dunstall, "Electrically caused wildfires in Victoria, Australia are over-represented when fire danger is elevated," *Landscape and Urban Planning*, vol. 167, pp. 267–274, 2017.
[107] "Learn about public safety power shutoffs," Pacific Gas and Electric Company, 2021. [Online]. Available: https://www.pge.com/en_US/residential/outages/public-safety-power-shuttoff/learn-about-psps.page
[108] "CPUC PSPS Rollup: Oct. 2013 Through Dec. 31, 2020," California Public Utility Commission, [Accessed : 1 May, 2021]. [Online]. Available: https://bit.ly/3irnSvm

[109] S. H. B. M. Sotolongo, C. Boloon, "California power shut-offs: Deficiencies in data and reporting," [Accessed : 18 Jan, 2021]. [Online]. Available: https://bit.ly/3B8GO9B

[110] Pacific Gas and Electric Company, "2021 Wildfire Mitigation Plan Report," *RULEMAKING 18-10-007*, 5 Feb. 2021, [Online]. Available: https://bit.ly/3aTv3ZD

[111] "Press release: PG&E Announces Major New Electric Infrastructure Safety Initiative to Protect Communities from Wildfire Threat; Undergrounding 10,000 Miles of Power Lines in Highest Fire-Threat Areas," Pacific Gas and Electric Company, 2021. [Online]. Available: https://bit.ly/3ithm7y

[112] Y. Zhang, F. You, H.-X. Chen, L.-H. Zhang, J.-P. Zhu, and J.-J. Zhou, "Statistical analysis of wildfire accidents inducing flashover of transmission lines," *Xiaofang Kexue yu Jishu*, vol. 30, no. 12, pp. 1177–1180, 2011.

[113] J. A. Sathaye, L. L. Dale, P. H. Larsen, G. A. Fitts, K. Koy, S. M. Lewis, and A. F. P. de Lucena, "Rising temps, tides, and wildfires: Assessing the risk to california's energy infrastructure from projected climate change," *IEEE Power and Energy Magazine*, vol. 11, no. 3, pp. 32–45, 2013.

[114] Pacific Gas and Electric Company, "Pacific Gas and Electric Company Amended 2019 Wildfire Safety Plan," Tech. Rep., 2019.

[115] G. Delille, B. Francois, and G. Malarange, "Dynamic frequency control support by energy storage to reduce the impact of wind and solar generation on isolated power system's inertia," *IEEE Trans. Sustainable Energy*, vol. 3, pp. 931–939, 2012.

[116] Y. Liu, Z. Mao, H. Li, K. S. Shetye, and T. J. Overbye, "Integration of renewable generators in synthetic electric grids for dynamic analysis," *arXiv preprint arXiv:2101.02319*, 2021.

[117] X. Fang, H. Yuan, and J. Tan, "Secondary frequency regulation from variable generation through uncertainty decomposition: An economic and reliability perspective," *IEEE Trans. Sustainable Energy*, 2021.

[118] W. Wang, X. Fang, H. Cui, F. Li, Y. Liu, and T. J. Overbye, "Transmission-and-distribution dynamic co-simulation framework for distributed energy resource frequency response," *IEEE Transactions on Smart Grid*, 2021.

[119] V. V. Viswanathan and M. Kintner-Meyer, "Second use of transportation batteries: Maximizing the value of batteries for transportation and grid services," *IEEE Trans. Vehicular Technology*, vol. 60, no. 7, pp. 2963–2970, 2011.

[120] "Generic Models (PV Plants)". [Online]. Available: https://www.esig.energy/wiki-main-page/generic-models-pv-plants/

[121] H. Cui, F. Li, and K. Tomsovic, "Hybrid symbolic-numeric framework for power system modeling and analysis," *IEEE Trans. Power Systems*, vol. 36, no. 2, 2020.

[122] W. Wang, X. Fang, and A. Florita, "Impact of DER communication delay in agc: Cyber-physical dynamic co-simulation," in *IEEE 48th Photovoltaic Specialists Conference (PVSC)*, 2021, pp. 2616–2620.

[123] R. C. Dugan and T. E. McDermott, "An open source platform for collaborating on smart grid research," in *2011 IEEE Power and Energy Society General Meeting*, 2011, pp. 1–7.

[124] B. Palmintier, D. Krishnamurthy, P. Top, and et al, "Design of the helics high-performance transmission-distribution-communication-market co-simulation framework," in *2017 Workshop on MSCPES*. IEEE, 2017, pp. 1–6.

[125] R. Podmore and M. R. Robinson, "The role of simulators for smart grid development," *IEEE Transactions on Smart Grid*, vol. 1, no. 2, pp. 205–212, 2010.

[126] S. Gissinger, P. Chaumes, J. . Antoine, A. Bihain, and M. Stubbe, "Advanced dispatcher training simulator," *IEEE Computer Applications in Power*, vol. 13, no. 2, pp. 25–30, 2000.

[127] T. J. Overbye, P. W. Sauer, C. M. Marzinzik, and G. Gross, "A user-friendly simulation program for teaching power system operations," *IEEE Transactions on Power Systems*, vol. 10, no. 4, pp. 31 725 – 1733, 1995.

[128] M. Foley, A. Bose, W. Mitchell, and A. Faustini, "An object based graphical user interface for power systems," *IEEE Transactions on Power Systems*, vol. 8, no. 1, pp. 97–104, 1993.

[129] R. Podmore, J. C. Giri, M. P. Gorenberg, J. P. Britton, and N. M. Peterson, "An advanced dispatcher training simulator," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-101, no. 1, pp. 17–25, 1982.

[130] J. E. Tate, T. J. Overbye, J. Sebestik, and G. C. Reese, "Interactive lessons for pre-university power education," *IEEE Transactions on Power Systems*, vol. 23, no. 3, pp. 824–830, 2008.

[131] T. Strasser, M. Stifter, F. Andrén, and P. Palensky, "Co-simulation training platform for smart grids," *IEEE Transactions on Power Systems*, vol. 29, no. 4, pp. 1989–1997, 2014.

[132] A. Chakrabortty and A. Bose, "Smart grid simulations and their supporting implementation methods," *Proceedings of the IEEE*, vol. 105, no. 11, pp. 2220–2243, 2017.

[133] T. J. Overbye, Z. Mao, A. Birchfield, J. D. Weber, and M. Davis, "An interactive, stand-alone and multi-user power system simulator for the pmu time frame," in *2019 IEEE Texas Power and Energy Conference (TPEC)*, 2019, pp. 1–6.

[134] T. J. Overbye, Z. Mao, K. S. Shetye, and J. D. Weber, "An interactive, extensible environment for power system simulation on the pmu time frame with a cyber security application," in *2017 IEEE Texas Power and Energy Conference (TPEC)*, 2017, pp. 1–6.

[135] A. B. Birchfield, T. J. Overbye, and K. R. Davis, "Educational applications of large synthetic power grids," *IEEE Transactions on Power Systems*, vol. 34, no. 1, pp. 765–772, 2018.

[136] Z. Mao, H. Huang, and K. Davis, "W4ips: A web-based interactive power system simulation environment for power system security analysis," *Proceedings of the 53rd Hawaii International Conference on System Sciences*, 2020.

[137] P. W. Sauer, M. A. Pai, and J. H. Chow, *Power system dynamics and stability: with synchrophasor measurement and Power System Toolbox*. IEEE Press, 2018.

[138] D. L. Fletcher and J. G. Coleman, "The dispatcher training simulator at peco energy company," *IEEE Transactions on Power Systems*, vol. 13, no. 3, pp. 1150–1155, 1998.

[139] G. D. Irisarri, M. Rafian, B. N. Miller, and E. J. Dobrowolski, "Heuristic scenario builder for power system operator training," *Proceedings of the IEEE*, vol. 80, no. 5, pp. 698–711, 1992.

[140] C. Y. Teo, H. B. Gooi, and T. W. Chan, "An innovative pc based simulator for power system studies," *Electric Power Systems Research*, vol. 38, no. 1, pp. 33–42, 1996.

[141] M. W. Shah and R. L. Biate, "Design and simulation of solar pv model using matlab/simulink," *International Journal of Scientific & Engineering Research*, vol. 7, no. 3, pp. 551–554, 2016.

[142] H. Li, J. L. Wert, A. B. Birchfield, T. J. Overbye, T. G. S. Roman, C. M. Domingo, F. E. P. Marcos, P. D. Martinez, T. Elgindy, and B. Palmintier, "Building highly detailed synthetic

electric grid data sets for combined transmission and distribution systems," *IEEE OAJPE*, vol. 7, pp. 478–488, 2020.

[143] A. Popp and J. Hughes, "Austin sustainable and holistic integration of energy storage and solar pv [austin shines]. final report, version 2," City of Austin, TX (United States), Tech. Rep., 2020.

[144] J. L. Wert, K. S. Shetye, H. Li, J. H. Yeo, X. Xu, A. Meitiv, Y. Xu, and T. J. Overbye, "Coupled infrastructure simulation of electric grid and transportation networks," in *2021 IEEE Power Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, 2021, pp. 1–5.

[145] M. Sengupta, Y. Xie, A. Lopez, A. Habte, G. Maclaurin, and J. Shelby, "The national solar radiation data base (nsrdb)," *Renewable and sustainable energy reviews*, vol. 89, pp. 51–60, 2018.

[146] J. Schmall, "Ercot experience with the der_a model," 2020. [Online]. Available: https://www.nerc.com/comm/PC/System%20Planning%20Impacts%20from%20Distributed%20Energy%20Re/Mod%20-%20DER_A-ERCOT-SPIDERWG_v1%20-%20Schmall.pdf

[147] J. L. Wert, H. Chang, F. Safdarian, T. J. Overbye, and Y. Xu, "The economic and technical impacts of houston's electric vehicles on the texas transmission system: A case study." IEEE, 2022, Conference Proceedings.

[148] K. S. Shetye, H. Li, J. L. Wert, X. Xu, A. Meitiv, Y. Xu, and T. J. Overbye, "Generation dispatch and power grid emission impacts of transportation electrification," in *2021 North American Power Symposium (NAPS)*. IEEE, 2021, Conference Proceedings.

[149] M. Muratori and T. Mai, "The shape of electrified transportation," *Environmental Research Letters*, vol. 16, no. 1, p. 011003, dec 2020. [Online]. Available: https://doi.org/10.1088/1748-9326/abcb38

[150] P. Grahn, K. Alvehag, and L. Söder, "Phev utilization model considering type-of-trip and recharging flexibility," *IEEE Transactions on Smart Grid*, vol. 5, no. 1, pp. 139–148, 2014.

[151] Y. Kongjeen and K. Bhumkittipich, "Impact of plug-in electric vehicles integrated into power distribution system based on voltage-dependent power flow analysis," *Energies*, vol. 11, no. 6, 2018. [Online]. Available: https://www.mdpi.com/1996-1073/11/6/1571

[152] N. O'Connell, Q. Wu, and J. Østergaard, "Efficient determination of distribution tariffs for the prevention of congestion from ev charging," in *2012 IEEE Power and Energy Society General Meeting*, 2012, pp. 1–8.

[153] "Electric Grid Test Case Repository". [Online]. Available: https://electricgrids.engr.tamu.edu/

[154] T. Xu, A. Birchfield, K. Gegner, K. Shetye, and T. Overbye, "Application of large-scale synthetic power system models for energy economic studies," in *Hawaii International Conference on System Sciences (HICSS)*, 2017.

[155] H. Li, J. H. Yeo, A. L. Bornsheuer, and T. J. Overbye, "The creation and validation of load time series for synthetic electric power systems," *IEEE Transactions on Power Systems*, vol. 36, no. 2, pp. 961–969, 2021.

[156] H. Li, J. H. Yeo, J. L. Wert, and T. J. Overbye, "Steady-state scenario development for synthetic transmission systems," in *2020 IEEE Texas Power and Energy Conference (TPEC)*, 2020, pp. 1–6.

[157] Texas Department of Motor Vehicles, *2021 Texas Alternative Fueled Vehicle Report*, 2021.
[158] "U.S. Plug-in Electric Vehicle Sales by Model." [Online]. Available: https://afdc.energy.gov/data/10567