

PAPER • OPEN ACCESS

## An open-source framework for balancing computational speed and fidelity in production cost models

To cite this article: Kerem Ziya Akdemir *et al* 2024 *Environ. Res.: Energy* **1** 015003

View the [article online](#) for updates and enhancements.

You may also like

- [Nanoscale phase-change materials and devices](#)  
Qinghui Zheng, Yuxi Wang and Jia Zhu
- [Preparation a three-dimensional hierarchical graphene/stearic acid as a phase change materials for thermal energy storage](#)  
Xiuli Wang, Xiaomin Cheng, Dan Li et al.
- [Thermal conductivity enhanced polyethylene glycol/expanded perlite shape-stabilized composite phase change materials with Cu powder for thermal energy storage](#)  
Shanmu Xu, Xiaoguang Zhang, Zhaohui Huang et al.

# ENVIRONMENTAL RESEARCH ENERGY



## PAPER

# An open-source framework for balancing computational speed and fidelity in production cost models

### OPEN ACCESS

RECEIVED  
29 August 2023

REVISED  
14 November 2023






ACCEPTED FOR PUBLICATION  
20 December 2023

PUBLISHED  
4 January 2024

Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](#).

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.



Kerem Ziya Akdemir<sup>1,\*</sup> , Konstantinos Oikonomou<sup>2</sup> , Jordan D Kern<sup>3</sup> , Nathalie Voisin<sup>2,4</sup> , Henry Ssembatya<sup>1</sup>  and Jingwei Qian<sup>1</sup>

<sup>1</sup> Department of Civil, Construction, and Environmental Engineering, North Carolina State University, Raleigh, NC 27695, United States of America

<sup>2</sup> Pacific Northwest National Laboratory, Richland, WA 99352, United States of America

<sup>3</sup> Department of Industrial and Systems Engineering, North Carolina State University, Raleigh, NC 27695, United States of America

<sup>4</sup> Department of Civil and Environmental Engineering, University of Washington, Seattle, WA 98195, United States of America

\* Author to whom any correspondence should be addressed.

E-mail: [kakdemi@ncsu.edu](mailto:kakdemi@ncsu.edu)

**Keywords:** open-source software, production cost model, fidelity/computational speed tradeoff, model selection, hydrometeorological extremes, multisector dynamics

Supplementary material for this article is available [online](#)

## Abstract

Studies of bulk power system operations need to incorporate uncertainty and sensitivity analyses, especially around exposure to weather and climate variability and extremes, but this remains a computational modeling challenge. Commercial production cost models (PCMs) have shorter runtimes, but also important limitations (opacity, license restrictions) that do not fully support stochastic simulation. Open-source PCMs represent a potential solution. They allow for multiple, simultaneous runs in high-performance computing environments and offer flexibility in model parameterization. Yet, developers must balance computational speed (i.e. runtime) with model fidelity (i.e. accuracy). In this paper, we present Grid Operations (GO), a framework for instantiating open-source, scale-adaptive PCMs. GO allows users to search across parameter spaces to identify model versions that appropriately balance computational speed and fidelity based on experimental needs and resource limits. Results provide generalizable insights on how to navigate the fidelity and computational speed tradeoff through parameter selection. We show that models with coarser network topologies can accurately mimic market operations, sometimes better than higher-resolution models. It is thus possible to conduct large simulation experiments that characterize operational risks related to climate and weather extremes while maintaining sufficient model accuracy.

## 1. Introduction

There is growing concern about the vulnerability of bulk electric power systems to weather and climate variability and extremes [1], which can strain the grid through demand spikes, generator outages and deratings, and other equipment failures, and cause reliability impacts and volatility in electricity markets [2–5]. Correspondingly, there is growing interest in incorporating these stressors (alongside the effects of decarbonization) into operational studies using production cost models (PCMs) [3, 4, 6–8]. Numerous previous studies have used PCMs to simulate bulk power system operations on the interconnection scale [9–19], and a subset of these explicitly focus on vulnerability to weather and climate variability and extremes. Most of these previous studies employ commercial grid simulation software. Commercial models have several advantages, including shorter runtimes. However, they also have important limitations (e.g. the opacity of numerical approaches and license restrictions, etc) that do not always allow for the large stochastic simulations needed to characterize weather and climate-related operational risks in power systems.

Take, for example, a hypothetical experiment in which a commercial PCM is used to characterize uncertainty in grid performance under future climate change. Given uncertainties in future greenhouse gas

emissions pathways, along with uncertainties across climate models (including downscaling approaches), even a modestly scoped analysis may involve running the PCM for 1000 s–10 000 s of individual weather years [2, 4]. What if the commercial PCM comes with a license that limits its use to serial processing (one model run at a time)? If the modeler is further constrained by a computational budget (e.g. total wall clock runtime limit of  $L$ ), the largest simulation ensemble possible is  $L$  divided by  $R$ , where  $R$  is the average runtime of the PCM for a single year. The model runtime  $R$  becomes limiting in the experimental design if the quotient  $\frac{L}{R} \ll 10,000$ .

Thanks to the development of open-source grid asset databases like the Electric Grid Test Case Repository [20] and OpenStreetMap [21], developing open-source grid operation models is now possible and several have emerged in recent years [22–24]. Open-source models allow for flexible design and simultaneous batch processing of multiple model runs in high-performance computing (HPC) environments—both of which support stochastic simulation. For example, to increase simulation size and expand capabilities for uncertainty characterization and sensitivity analysis, developers can reduce the runtime of an open-source PCM (i.e. increase  $L/R$ ) by simplifying core process representations. Several options exist to do this, including (but not limited to) aggregating across model features and space/time, network reduction, and relaxation of binary variables [3, 25–29].

A broad literature exists on navigating the model fidelity and computational speed tradeoff in PCMs and power system expansion models [3, 30–38]. For instance, the authors in [38] developed a method to investigate the sensitivity of power system expansion models to various parameters and model structures. They found out that temporal resolution is the leading sensitivity factor in power system models. The authors in [36] analyze the effects of temporal resolution in power sector investments under different technology and policy scenarios. They argue that simplifying the temporal variability used in integrated power sector investment models might lead to deteriorations in model quality. Models should be neither oversimplified (e.g. unable to capture realistic behaviors of the grid in response to hydrometeorological stressors) nor overly complex (e.g. computational runtimes that limit experiment size and exploration of uncertainty) [39]. Yet, a comprehensive framework for customizing open-source PCMs based on these recommendations is generally lacking in the literature.

In this paper, we introduce the Grid Operations (GO) family of models, which allows for the easy training and testing of PCMs constructed around open-source synthetic grid databases. The main novelty of the GO software is that it establishes a pipeline from open-source energy data to the instantiation and calibration of PCMs. Furthermore, GO gives users the ability to vary the network topology and mathematical model formulation, as well as model parameters such as transmission line limits and economic hurdle rates. Users can then identify a model version that meets their experimental needs (e.g. resource limitations) in a robust and reproducible manner.

GO is available for all three interconnections of the U.S. (Western Interconnection, Eastern Interconnection, and Texas Interconnection). In this paper, we demonstrate our approach over the Western Interconnection, a large and diverse system with previously demonstrated sensitivity to weather and climate variability and extremes [14, 18, 40]. We conduct an evaluation of 100s of model versions and demonstrate quantitatively that models with coarser topologies and lower runtimes can capture system dynamics adequately, unlocking expanded capabilities for stochastic simulation. This paper doubles as a detailed roadmap for how to use GO's flexible platform to design scale-adaptive model versions that balance computational complexity and fidelity.

## 2. Methods

In this section, we start with a general description of the GO framework, including its approaches for calibrating PCMs. Then, we demonstrate how GO can be implemented in a real system, the U.S. Western Interconnection.

### 2.1. The generalizable GO framework

The GO software, written in Python, is based on synthetic grid databases of generators and system topologies created by Texas A&M University (TAMU) [41–44]. GO instantiates simpler modeled versions of interconnection scale PCMs in four steps: (1) nodal selection; (2) network reduction; (3) data allocation and file setup; and (4) simulation using PCMs. Figure 1 summarizes the generalizable GO framework.

#### 2.1.1. Step 1: Nodal selection

GO allows users to train, test, and select a PCM version by searching over four user-defined parameters (see blue boxes in figure 1). One of these parameters is the number of nodes that should form the final, simplified network. For example, if the full TAMU network representation has 10 000 nodes, a user could specify that

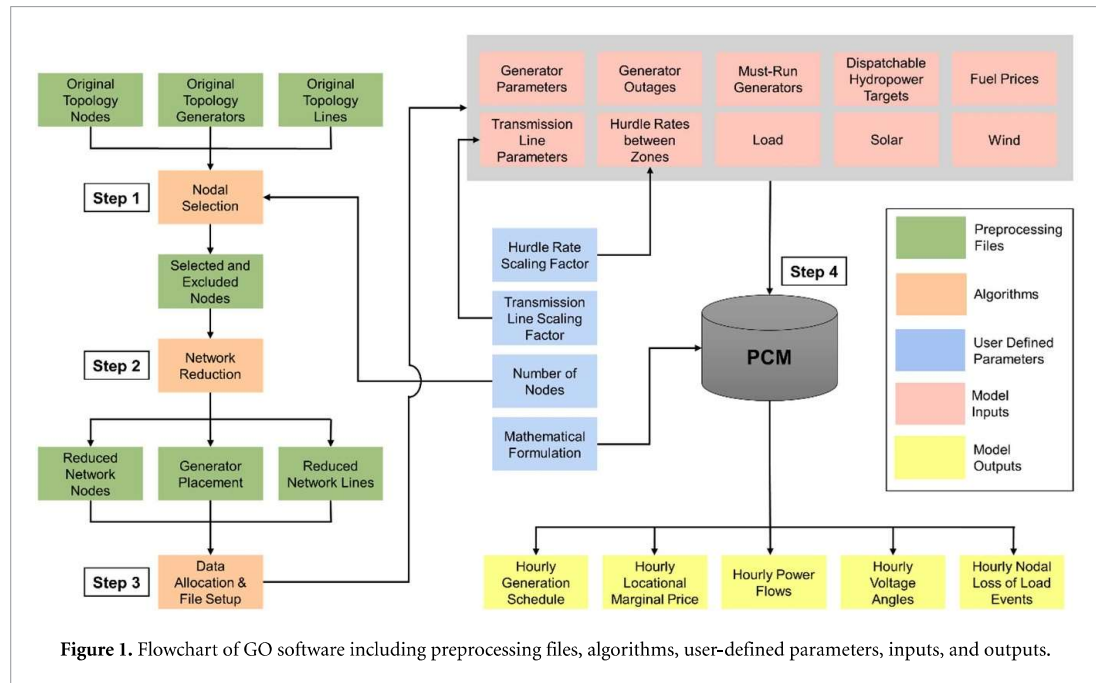


Figure 1. Flowchart of GO software including preprocessing files, algorithms, user-defined parameters, inputs, and outputs.

they want to reduce this to a simpler, ‘backbone’ network made up of 100 nodes. After specifying this number, the first step in GO is a nodal selection algorithm that identifies which nodes to preserve from the full synthetic TAMU dataset. There are three types of node classifications in the original TAMU dataset: demand nodes (each of which comes with an estimated average load), generation nodes (each of which is assigned a total generation capacity), and transmission nodes (each of which is assigned a voltage level). Several criteria are applied to identify the most critical nodes in the network and guarantee minimum requirements to electrically connect and geographically cover interconnection scale systems.

First, for each unique balancing authority (BA) and state pairing, the node with the largest average demand is selected (BAs are entities which are responsible for balancing electricity supply and demand over distinct geographical areas). After removing these from the set, the remaining nodes are selected in equal numbers across the three different node classifications, though this ratio can be altered if desired. The algorithm selects demand nodes starting from the one with the highest electricity demand and moving to nodes with lower demands. Likewise, the algorithm selects generation nodes starting from the one with the largest generation capacity and moving to nodes with smaller generation capacities. After filtering nodes by voltage level ( $>345$  kV), the algorithm selects transmission nodes starting from the node with the highest demand and moving to lower demands. Throughout the nodal selection process, a distance threshold (e.g. 5 km) restricts the algorithm from selecting any two nodes that are within that threshold from each other (this threshold can be changed by the user).

### 2.1.2. Step 2: Network reduction

After selecting which nodes to preserve from the original TAMU network, GO uses a network reduction algorithm [45] to generate a new, simplified network equivalent topology. The algorithm first moves all generators to nodes selected in step 1 based on the shortest electrical distance. It then recalculates nodal loads to compensate for the movement of the generators, such that estimated power flows on the preserved lines exactly match those in the full system. For other lines, the algorithm creates ‘equivalent’ lines and associated impedances based on Ward’s equivalent circuit calculation [46]. Using distances and impedance values typical for high-voltage transmission lines, GO computes the per-distance impedance value ( $\Omega \text{ km}^{-1}$ ), which we then associate with line flow capacities according to the transmission line loadability curve [47]. See [45] for more details about the network reduction process used by GO. For a general review of network reduction methods, also see [48].

### 2.1.3. Step 3: Data allocation and file setup

After using the network reduction algorithm to create a simplified network, a data allocation algorithm disaggregates BA-level time series inputs, such as electricity load and solar and wind generation, and assigns these values to individual nodes. BA-level time series inputs are gathered from the U.S. Energy Information



Administration (EIA) [49]. This process also creates job folders, which include all inputs and scripts necessary to run the simulations on a desktop computer or in an HPC environment.

In order to disaggregate BA-level load time series to each of the nodes in the reduced topology, nodal load weights are calculated within each BA by using the average load information from the TAMU dataset for each node in the reduced topology. In the TAMU dataset, a representative average nodal load value is assigned to each node. These nodal load values are used to calculate nodal load weights. BA nodal load weights are calculated by dividing the average nodal load of a specific node (from TAMU) by the sum of the average nodal load values of every node (from TAMU) in that specific BA. Nodal load weights are then multiplied by the load time series for their respective BAs to come up with the nodal loads for each hour. Available BA solar/wind generation is allocated to each node by calculating a weight reflecting the installed solar/wind capacity at each node. These weights are then multiplied by the solar/wind generation time series for their respective BA to get nodal solar/wind generation time series.

#### 2.1.4. Step 4: PCM

After nodal selection, network reduction, and data allocation and file setup, GO simulates GO using the resultant PCM. The PCM can simulate bulk electricity GO using either linear programming (LP) (only economic dispatch (ED)) or mixed integer LP (both unit commitment (UC) and ED) formulations. The PCM is written in the Pyomo mathematical optimization package and can be solved using open-source (e.g. HiGHS, SCIP) or commercial solvers (e.g. CPLEX, Gurobi). All results shown in this paper were produced by pairing GO's PCM with the Gurobi solver.

The objective function of the PCM is to minimize the system-wide cost of meeting fluctuating hourly electricity demand as well as cost of unserved energy, subject to several constraints such as individual generator capacities and ramp rates, thermal capacities of transmission lines, and Kirchhoff's current and voltage laws using a DC power flow approximation. The PCM iteratively minimizes costs over a user-defined operating horizon (e.g. 24–168 h) within which the modeled system operator has perfect foresight. Results presented in this paper reflect an operating horizon of 24 h. The decision variables consist of binary (if mixed-integer linear programming (MILP) is selected) and continuous electricity generation variables that control generator scheduling and dispatch, power flow between different nodes, voltage angles at each node, and loss-of-load variables at each node. Loss-of-load variables are units of last resort whose marginal cost is priced at \$2000 MWh<sup>-1</sup> (the current loss of load price in several major markets [50]). Model outputs are hourly operating generation schedules of each power plant, hourly locational marginal prices (LMPs) at each node, simulated power flows on every transmission line, hourly voltage angle at each node, and (if applicable) hourly loss of load at each node.

Nuclear power plants are regarded as must-run resources apart from forced and unforced generator outages. The availabilities of solar and wind generation are represented by exogenously defined hourly time series, though the system operator can curtail both solar and wind if necessary. Weekly hydropower generation targets are collected from EIA-923 dataset [51]. From weekly data, hourly minimum, hourly maximum, and daily allowed total hydropower generation is calculated and fed into the model. Then, PCM determines the optimal hourly hydropower schedule at each node.

Hourly generator outages are represented using data from the North American Electric Reliability Corporation's Generating Availability Data System [52]. Estimated lost capacity due to generator outages is subtracted from the nameplate generator capacities in each hour. Note that representation of both hydropower availability and unit outages can easily be substituted with alternative approaches, if desired. A more detailed discussion of GO's PCM modeling approach can be found in supplementary information.

## 2.2. Model calibration and selection process

GO allows users to calibrate and select a PCM version by searching over four user-defined parameters (see blue boxes in figure 1):

### 2.2.1. Number of nodes

Users can select different numbers of nodes that will be in the final, simplified network representation of the PCM, thereby affording flexibility in the level of system granularity. As the number of nodes in the system increases, model runtimes generally increase as well because linear and mixed integer linear power system problems exhibit polynomial time complexity.

### 2.2.2. Mathematical formulation

Users can choose to implement only ED processes by using only LP or they can model both UC and ED processes by using MILP. MILP formulations entail higher runtimes due to the presence of binary generator on/off decision variables. On the other hand, MILP formulations allow for greater fidelity with respect to the

operations of power plants (inclusion of startup costs, no load costs, and minimum-up and minimum-down times of dispatchable generators). Capturing these generator characteristics can be important in accurately capturing the generation mix, power plant emissions, and electricity prices.

### 2.2.3. Transmission line capacity scaling factors

In GO, initial transmission line capacities are initially estimated by the network reduction algorithm. However, users can uniformly adjust transmission line capacities ( $\pm$  in MW) in the reduced network to increase model fidelity.

### 2.2.4. Hurdle rate scaling factors

Hurdle rates represent the cost of delivering 1 MWh of electricity from one BA (i.e. a set of geographically clustered nodes representing quasi-independent grid operators) to another. Users can search over a range of different percentage scaling factors ( $\pm$  in %) to uniformly alter hurdle rates among BAs.

GO tracks the performance of the PCM by comparing simulated LMPs and generation mixes with historical data. Model performance is measured in terms of three fidelity metrics: mean absolute percentage errors in the interconnection scale yearly generation mix, and  $R^2$  and root mean square error (RMSE) of daily LMPs at major pricing hubs/wholesale electricity markets.

Daily LMPs at major pricing hubs are calculated by taking a demand-weighted average of nodal LMPs within the geographic boundaries of each pricing hub (see figure 3 for boundaries of pricing hubs). In order to calculate fidelity metrics on interconnection scale (i.e. for Eastern Interconnection and Western Interconnection), daily LMPs for multiple pricing hubs are utilized to calculate an aggregate (i.e. interconnection scale)  $R^2$  and RMSE score by the demand-weighting average LMPs across all pricing hubs. For example, if the average loads in three different pricing hubs are: 10 GWh, 5 GWh, and 15 GWh, and simulated vs. historical price comparisons in the same hubs yield  $R^2$  values of 0.90, 0.85, and 0.95, the interconnection scale  $R^2$  score for the PCM would be 0.9167.

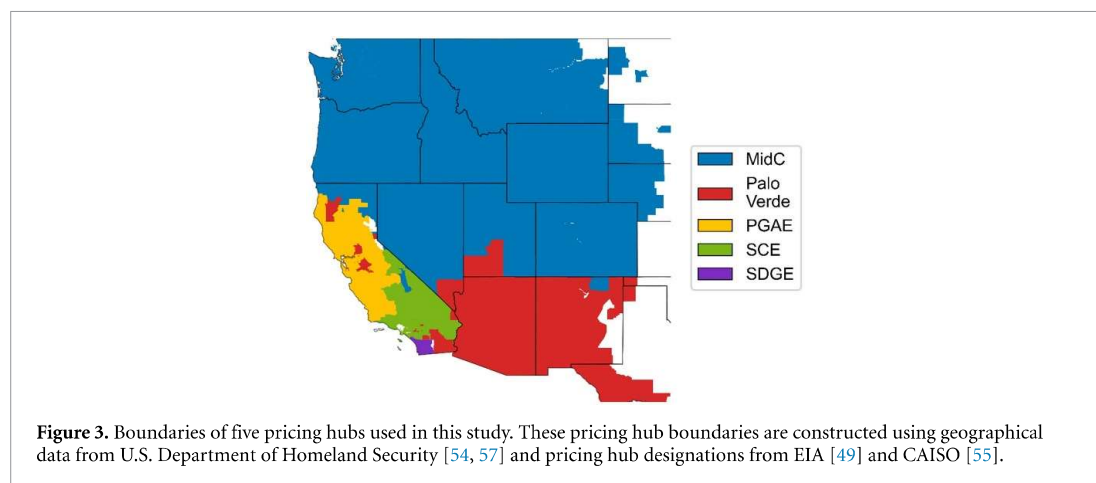
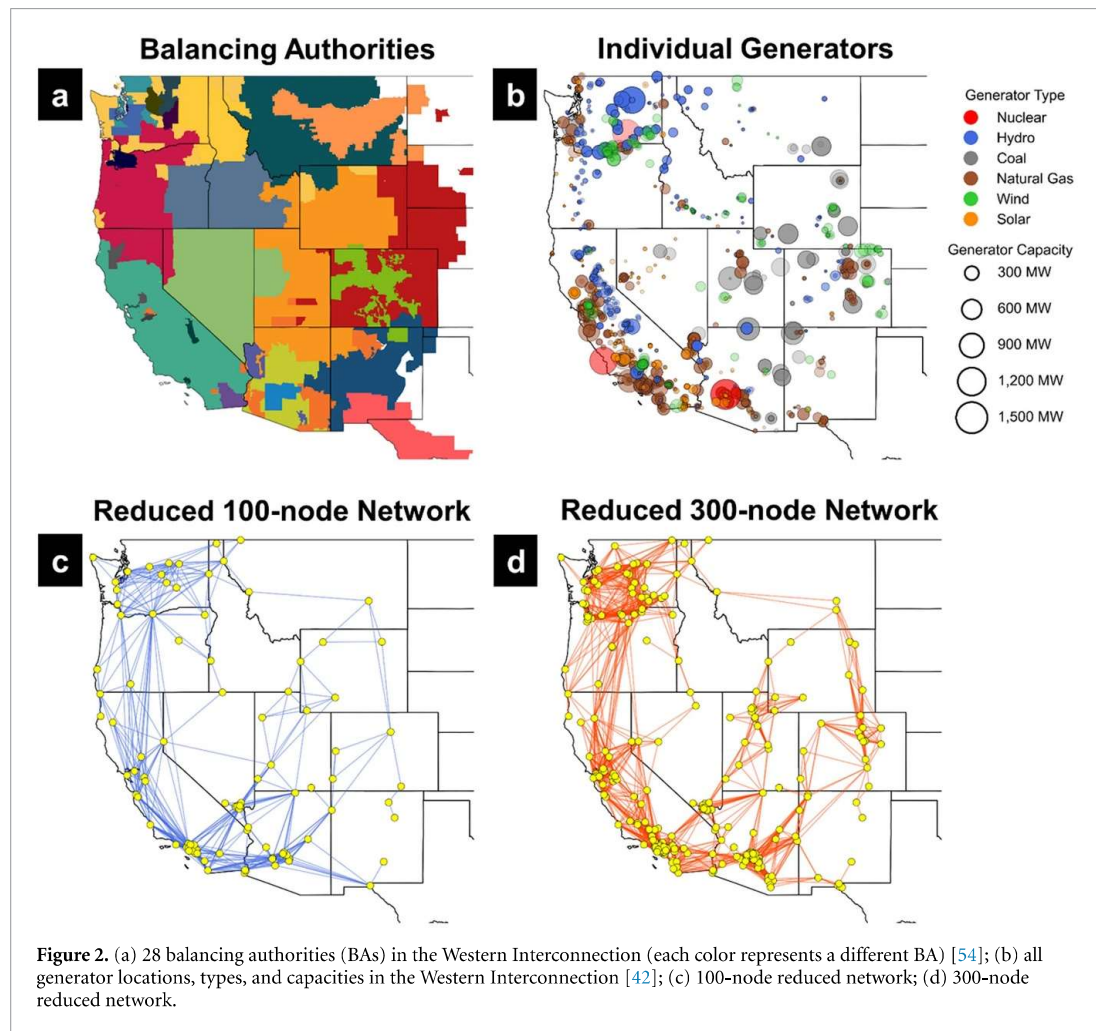
GO numerically ranks each PCM tested (1 = best) with respect to each individual metric (LMP  $R^2$ , LMP RMSE, and generation mix error), and then assigns each PCM an overall numerical ranking based on the sum of individual metric rankings. In this case, the model version with lowest sum of rankings would be the most accurate model among all created model versions. The model versions are sorted with respect to their sum of individual metric rankings and are given an ultimate fidelity ranking. This way, users can identify the highest-fidelity version of the PCM.

## 2.3. Demonstration of GO in critical test bed: U.S. Western Interconnection

GO facilitates the quick development of open-source PCMs and evaluation of model performance over a wide parameter space, allowing users to find versions that sufficiently balance the tradeoff between model fidelity and computational speed. In the remainder of this paper, we demonstrate these capabilities in a test bed of critical importance: the U.S. Western Interconnection, an interconnected system of 28 separate balancing authorities across the states of California, Oregon, Washington, Idaho, Nevada, Arizona, Utah, Wyoming, Montana, Colorado, and New Mexico (figure 2). In this paper, we use GO to (a) instantiate 540 different PCM models over a wide combinatorial parameter set; (b) measure model performance; and (c) identify models that demonstrate high fidelity and sufficient simulation speed. The 540 model versions are the combinations of:

- 9 different reduced networks (containing 100, 125, 150, 175, 200, 225, 250, 275, and 300 nodes) (see figures 2(c) and (d))
- 2 different mathematical formulations (LP and MILP)
- 6 different transmission limit scaling factors (+0 MW (baseline), +500 MW, +1000 MW, +1500 MW, +2000 MW, +2500 MW)
- 5 different hurdle rate scaling factors (−100%, −50%, 0% (baseline), +50%, +100%). Baseline values of BA-to-BA hurdle rates are taken from the 2030 Anchor Data Set developed by the Western Electricity Coordinating Council (WECC) [53].

Due to lack of open-source data and losing one-to-one node correspondence in a reduced network representation, it is not possible to conduct an LMP comparison for each node in the Western Interconnection. In this sense, LMP comparison is carried out for each pricing hub. We have access to historical prices at five major pricing hubs in the Western Interconnection. Three are in California: Pacific Gas and Electric (PG&E), Southern California Edison (SCE), and San Diego Gas & Electric (SDGE) [55]. The other two pricing hubs are the informal Mid-Columbia (MidC) trading hub in the Northwest and the Palo Verde trading hub in the Southwest [56]. Since the temporal resolution of historical LMPs in MidC and

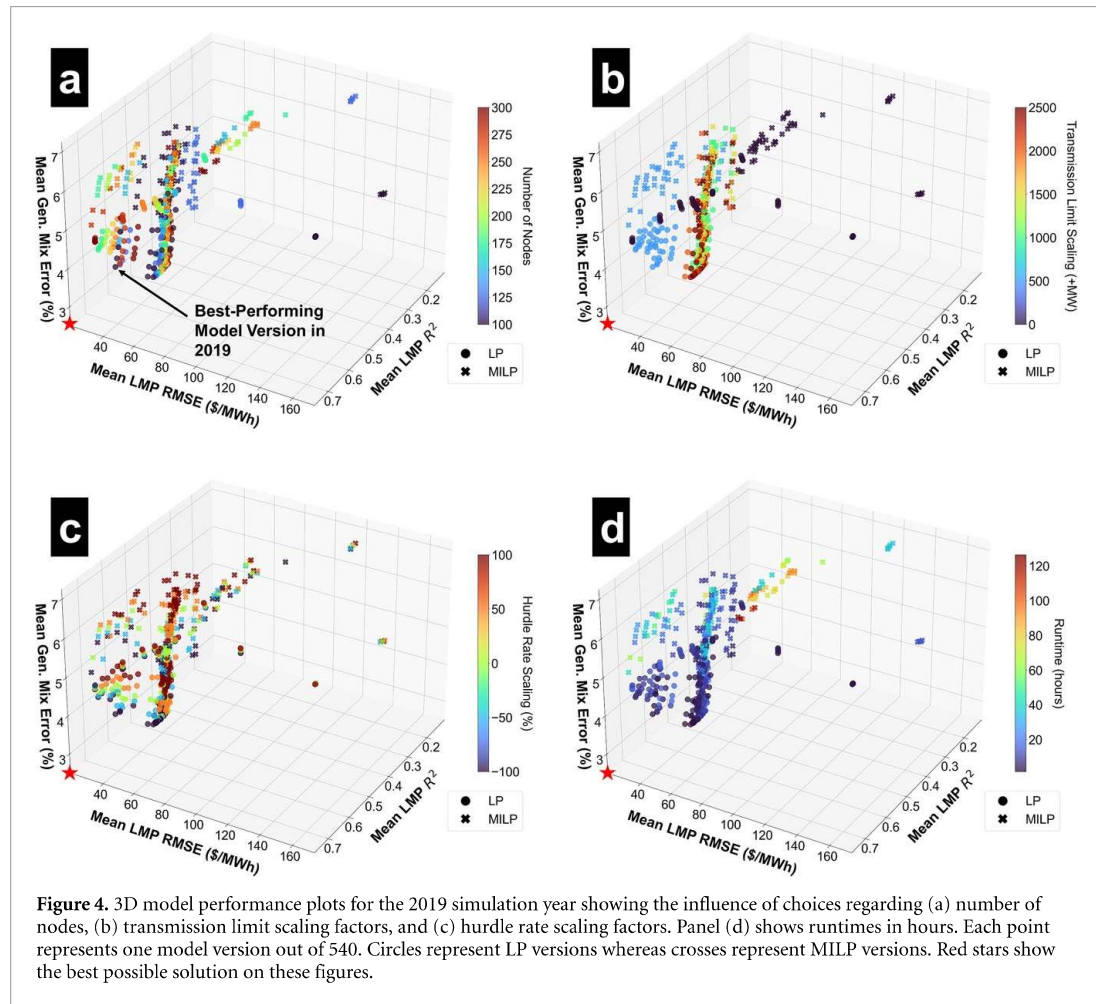


Palo Verde is daily, we aggregate LMPs simulated by the GO PCM to a daily time step for comparison. The boundaries of these five pricing hubs are shown in figure 3.

### 3. Results and discussion

#### 3.1. Yearly model calibration and selection results

Historical operating data for every BA in the Western Interconnection are available for the years 2019–2021, so our parameter search (i.e. model calibration) focuses on this period. Figures 4(a)–(d) show the



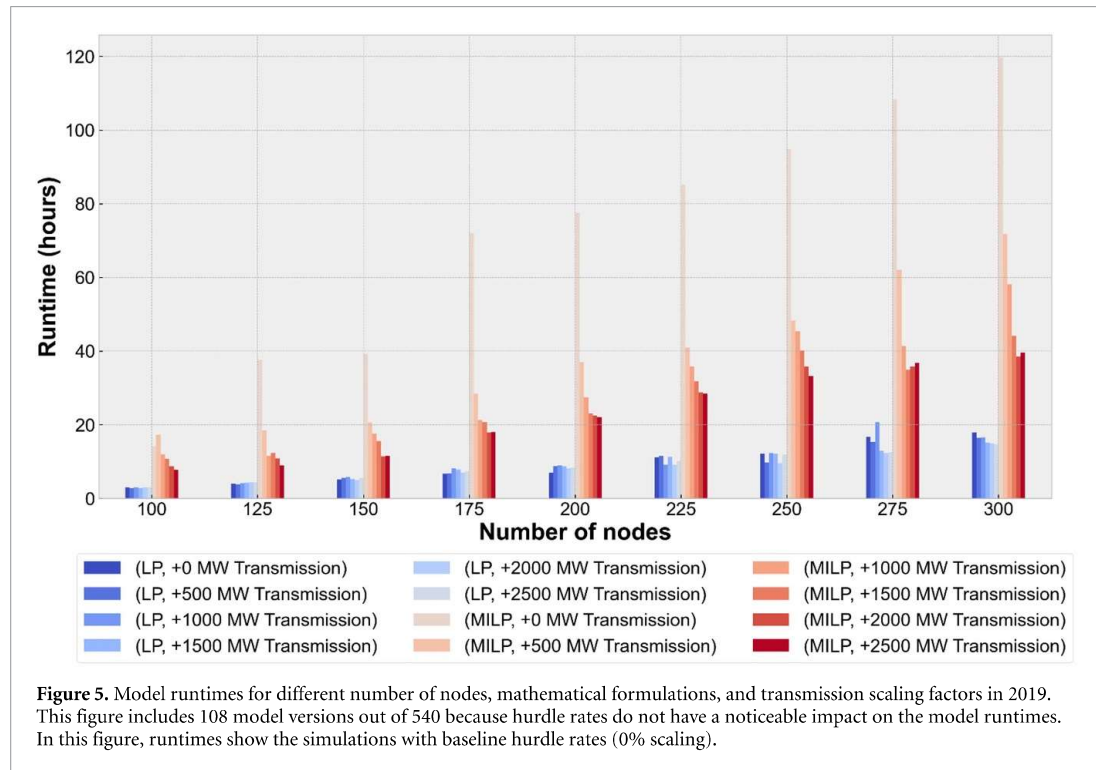
performance of all 540 model versions for the 2019 simulation year. The three dimensions measure model performance, and red stars at the origin show the ideal point (minimized LMP RMSE and generation mix error, maximizing LMP  $R^2$ ). LP formulations are shown with circles whereas MILPs are designated with crosses. In each respective panel, colors signify the number of nodes in the reduced network, transmission limit scaling factors, hurdle rate scaling factors, and wall-clock runtime.

Figure 4(a) shows that there is not a straightforward relationship between the number of nodes and model fidelity. While the expectation is generally that greater model complexity (i.e. larger numbers of nodes) should yield greater model accuracy, we see that some models with fewer nodes (blue colors tones) can also mimic GO relatively well. Figure 4(b) shows that lower transmission scaling factors (like +500 MW) do better in terms of LMP  $R^2$  and RMSE values, but higher transmission scaling factors seem to be better at capturing generation mix. In general, we observe that a higher number of nodes can give better results when coupled with lower transmission scaling factors, and models with lower numbers of nodes require higher transmission scaling factors. This is because having more nodes in the reduced topology increases the electrical connectivity of individual nodes so that lower transmission scaling factors can suffice. When we evaluate model versions by hurdle rate scaling (figure 4(c)), we can see that there is a consistent gradient in colors, indicating that lower hurdle rates increase model fidelity in 2019.

Figure 4(d) shows the same 540 model versions evaluated in terms of wall-clock runtime. Runtimes for a single year (8760 h) vary between 2 and 120 h depending on the number of nodes, mathematical formulation, and transmission scaling factors. Clearly, there are some less complex (lower node) LP formulations with shorter runtimes that can also do a good job of capturing LMPs and generation mix, indicated by numerous dark blue circles close to the ideal point (red star). In fact, the best-performing model version for 2019 only takes 2 h to finish simulating 1 year of hourly GO.

Figure 5 compares runtimes for different model versions (varying the number of nodes in the reduced network, mathematical formulation, and transmission scaling factors), all assuming the same default hurdle rates. In general, increasing the number of nodes increases the model runtimes whereas increasing





transmission line capacities decreases runtimes. However, after a certain threshold (around +2000 MW), increasing transmission line capacities does not significantly affect the runtimes because the model becomes mostly free of transmission capacity limitations. In our simulations, changing hurdle rates did not have a significant impact on the runtimes.

Figure 6 shows an example of a generation mix comparison for the best-performing model version for the 2019 simulation year (which is also indicated in figure 4(a)). Although there are some differences between historical and simulated daily generation mix, this model version captures yearly generation mix trends in the Western Interconnection to a certain extent.

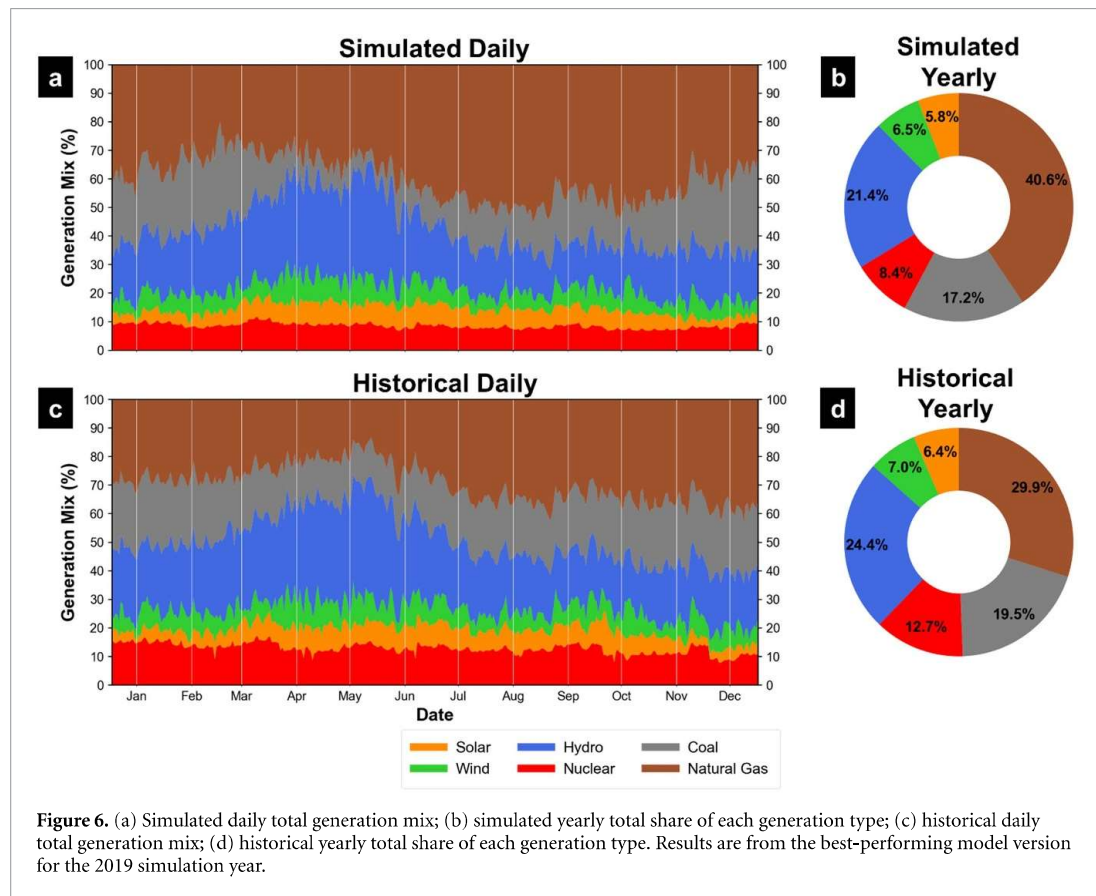
Figure 7 shows a visual comparison example of historical and simulated daily LMP time series from the same, best-performing model version for the 2019 simulation year. This model version does a good job capturing LMP variations in PG&E, SCE, and SDGE but it misses some oscillations in MidC and Palo Verde LMPs.

We calibrated the PCM separately for 2019, 2020, and 2021, and then for all combinations of those years with a leave-one-year-out approach (e.g. train on 2019–2020, test on 2021; train on 2019 and 2021, test on 2020, etc). We then compared the resultant LMPs and generation mixes with historical data and selected the best model version for each year combination. For 2019, a model with a 125-node topology, LP formulation, +500 MW transmission limit scaling, and –100% hurdle rate scaling yielded the best results. The best model version for 2020 has the same characteristics, except for the number of nodes (100 vs 125). The best-ranked model for 2021 includes a 225-node topology, a MILP formulation, +500 MW transmission limit scaling, and –100% hurdle rate scaling. See figures S2 and S3 in the supplementary information for 3D model performance plots for 2020 and 2021. Table 1 lists all selected parameters of the best-ranked model versions for each set of training year(s). In general, an LP formulation with +500 MW transmission scaling and –100% hurdle rate scaling is the most robust model version.

### 3.2. Impact of user-defined parameters on model fidelity and different model selection methods

GO's flexibility allows users to instantiate hundreds of different model versions. In this section, we mine these simulation results for larger patterns in how parameters affect model performance. The panels in figure 8 slice the 540 different model versions in several ways. Each column shows data for a different simulation year (2019, 2020, and 2021). The rows isolate the effects of changing each user-defined parameter. For example, in the first row, each box plot (color) shows distribution of performance rankings (1 = best, 540 = worst) of 60 unique model versions that share the same number of nodes but differ in terms of mathematical formulation, transmission line scaling, and/or hurdle rate scaling. There are 9 box plots,



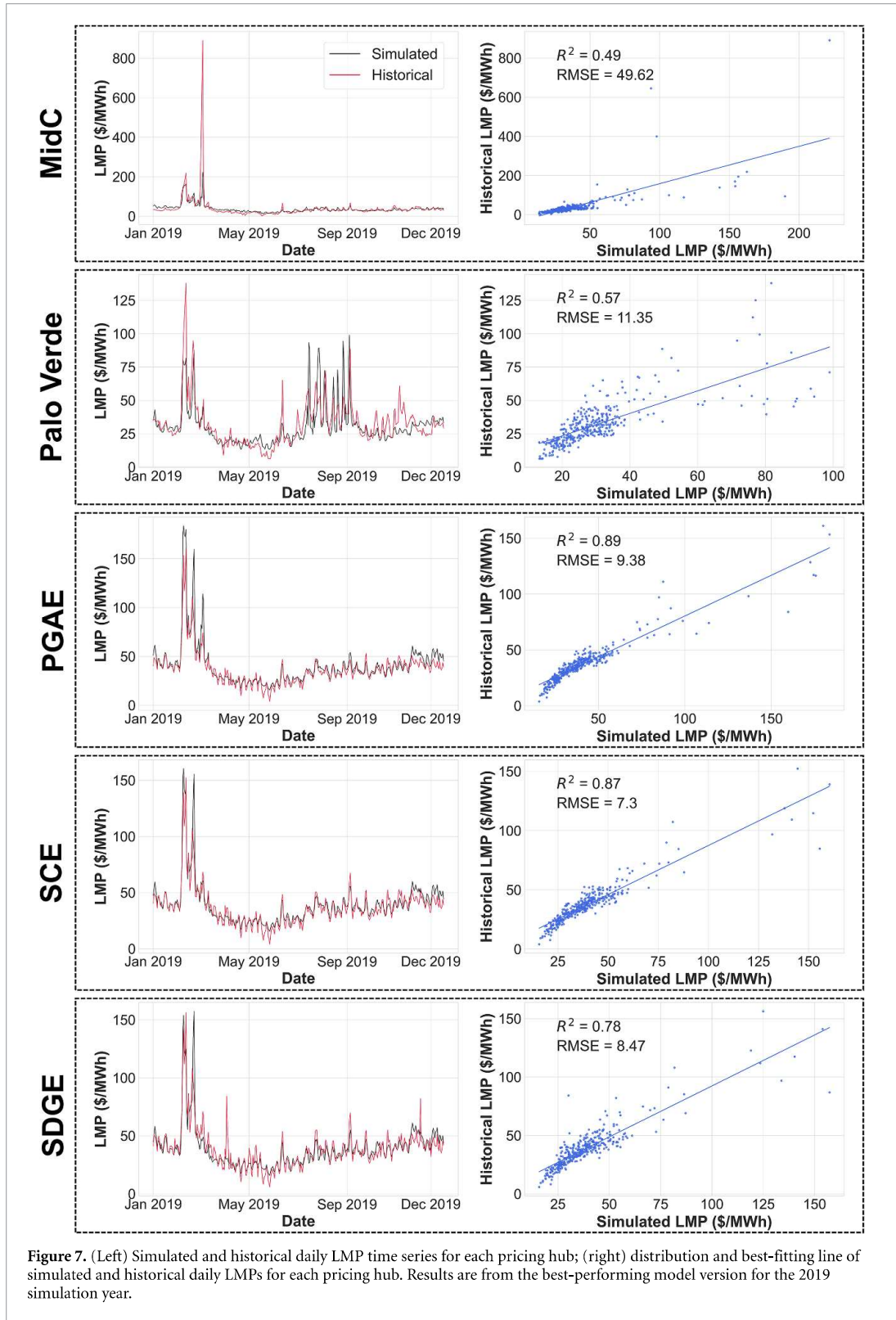


equaling the 540 total versions tested. Likewise, in the second row, each box plot (color) represents 270 out of 540 model versions, all of which share the same mathematical formulation (either LP or MILP).

The first row of figure 8 shows a fairly consistent trend across all three simulation years: as the number of nodes increases, the *median* model ranking decreases (i.e. model fidelity increases). At first glance, this may seem to conflict somewhat with our finding that the best-ranked models generally have a smaller number of nodes (see table 1). However, the lower whisker of each box plot falls close to zero, indicating that higher performing models are possible for any number of nodes selected. The interquartile ranges (IQRs) of the boxplots also vary. For example, in 2019, models with between 150–250 nodes have smaller IQRs, indicating that the performance of these models is more stable across values of other parameters than model versions with a smaller/larger number of nodes. Looking at 2021, the opposite is true. In addition, the second row of figure 8 indicates that LP versions did a better job in 2019 and 2020 but MILP versions performed better in 2021 on average.

Another finding illustrated by figure 8 (third row) is that moderate levels of transmission line scaling typically yield more accurate models. In fact, we observe that high-performing (low-ranked) models become impossible to find for certain transmission line parameterizations. Zero transmission scaling consistently yields the least-accurate models. In addition, as we increase transmission line capacities above +500 MW to +1000 MW, the model becomes free of transmission restrictions and starts to utilize lower marginal cost power plants and transfer electricity more freely throughout the network, which causes inaccuracies in generation mix and LMPs in different regions. In general, we find that model performance is the least sensitive to hurdle rate scaling (see figure S4 in the supplemental information shows small changes in model selection metrics). Thus, we primarily focus our discussion on the other three, more impactful, parameters.

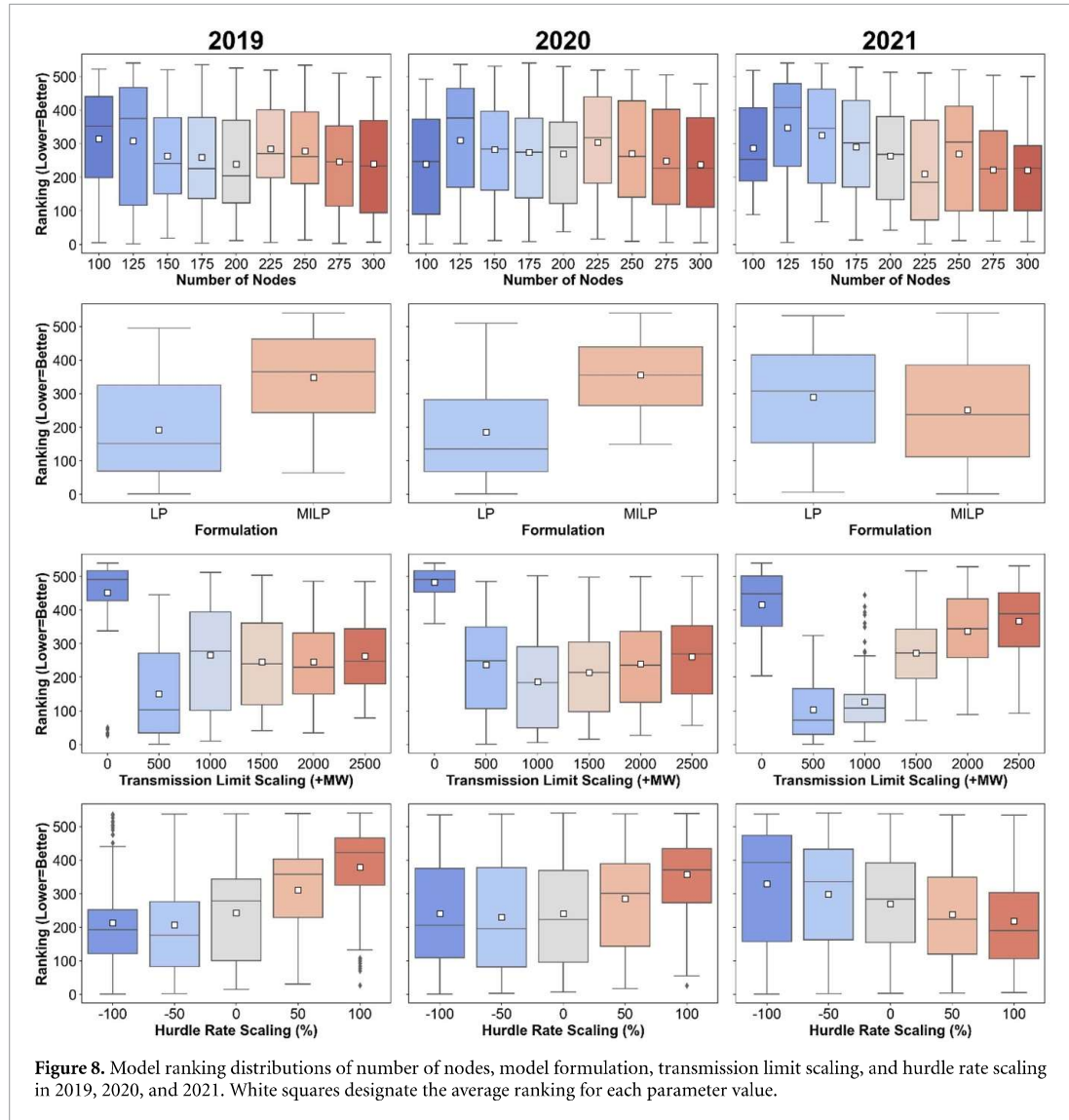
Figure 9 provides a deeper examination of the interactive effects of two key model parameters: the number of nodes in the network and transmission line scaling. In particular, we observe that having a higher number of nodes (warmer colored boxplots) becomes much more important to model accuracy if transmission lines are not scaled-up (panels on the left). Increasing the number of nodes enhances the connectivity of the reduced network, which prevents LMP spikes caused by activation of loss of load variables. However, if we scale transmission lines by more than +500 MW, we can create the same



connectivity conditions with fewer nodes. In fact, at higher transmission scaling levels, networks with lower numbers of nodes yield accurate models. In general, we find that transmission line and hurdle rate scaling factors interact like supporting calibration parameters (e.g. hyperparameters), with the nature of this interaction also dependent on the number of nodes and mathematical formulation selected.

**Table 1.** Selected parameters of the best model versions for each training set.

Training set	Number of nodes	Mathematical formulation	Transmission line limit scaling factor	Hurdle rate scaling factor
2019	125	LP	+500 MW	-100%
2020	100	LP	+500 MW	-100%
2021	225	MILP	+500 MW	-100%
2019 and 2020	125	LP	+500 MW	-100%
2019 and 2021	125	LP	+500 MW	-100%
2020 and 2021	125	LP	+500 MW	-100%
2019, 2020, and 2021	125	LP	+500 MW	-100%



**Figure 8.** Model ranking distributions of number of nodes, model formulation, transmission limit scaling, and hurdle rate scaling in 2019, 2020, and 2021. White squares designate the average ranking for each parameter value.

### 3.3. Influence of training and testing data on model selection

Table 2 lists model performance metrics and rankings (1 = best) for all possible training and test year combinations, allowing us to observe how the choice of this data could impact model selection.

For example, when the 540 different versions of the PCM are used to simulate 2019, the version that performs the best also performs second best when tested in 2020 and the 16th best version when tested in 2021. When we train the PCM on all two and three-year combinations, the best-performing model version is the exact same version that performs best when simulating 2019 (see table 1).

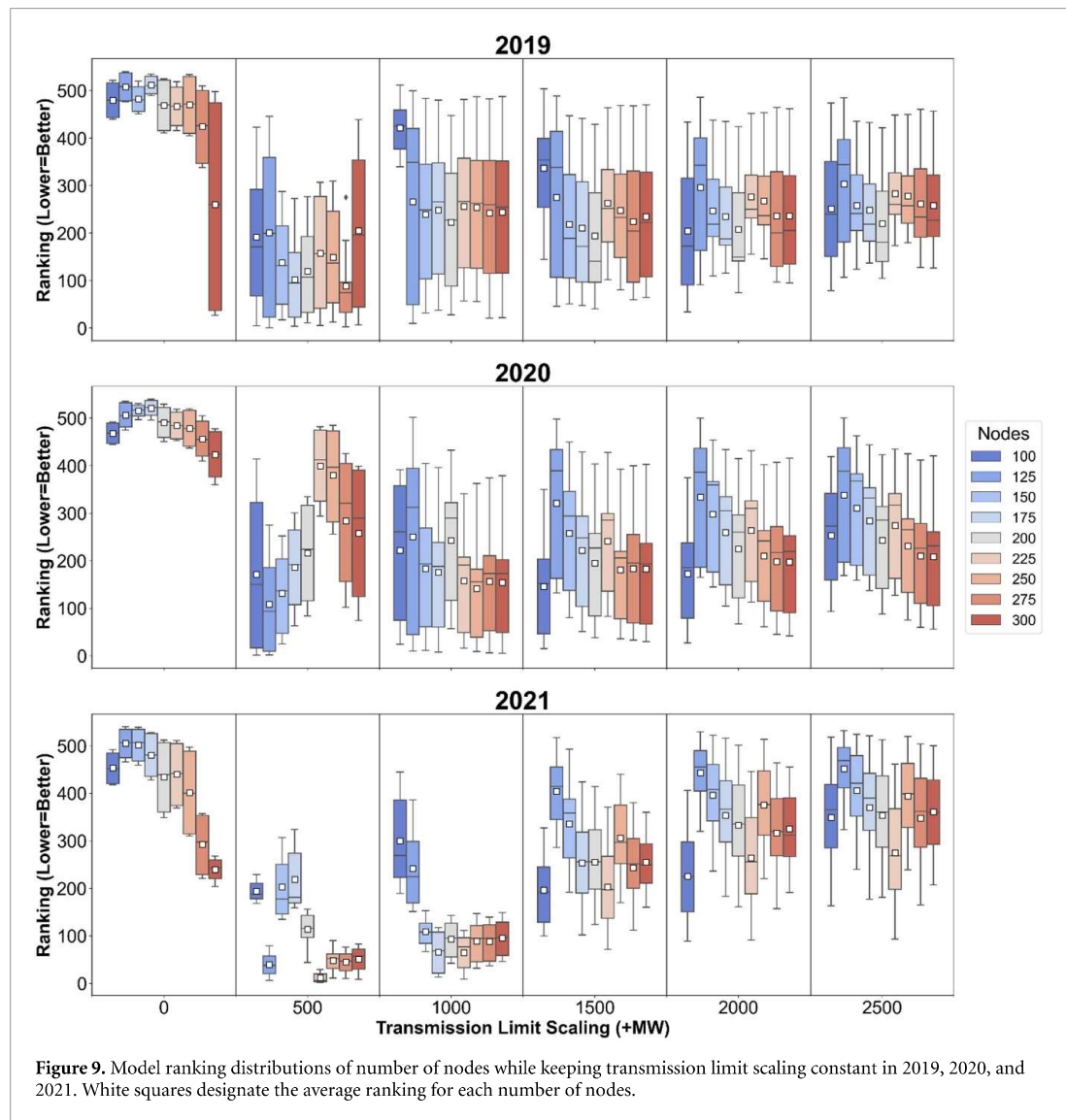


Figure 9. Model ranking distributions of number of nodes while keeping transmission limit scaling constant in 2019, 2020, and 2021. White squares designate the average ranking for each number of nodes.

Table 2. Model performance metrics and rankings for different training and test years.

	Test year												
	2019				2020				2021				
	LMP $R^2$	RMSE	Gen. mix error (%)	Rank	LMP $R^2$	RMSE	Gen. mix error (%)	Rank	LMP $R^2$	RMSE	Gen. mix error (%)	Rank	
Training Years	2019	0.62	30.35	3.56	1	0.62	29.69	2.91	2	0.31	52.64	3.04	16
	2020	0.66	33.31	3.73	5	0.63	31.48	2.75	1	0.28	55.95	2.41	213
	2021	0.58	32.19	5.58	221	0.42	35.48	5.08	468	0.27	52.22	1.73	1
	2019–2020	0.62	30.35	3.56	1	0.62	29.69	2.91	2	0.31	52.64	3.04	16
	2019–2021	0.62	30.35	3.56	1	0.62	29.69	2.91	2	0.31	52.64	3.04	16
	2020–2021	0.62	30.35	3.56	1	0.62	29.69	2.91	2	0.31	52.64	3.04	16
	2019–2020–2021	0.62	30.35	3.56	1	0.62	29.69	2.91	2	0.31	52.64	3.04	16



When the 540 different versions of the PCM are used to simulate 2020, the version that performs the best also does a good job of capturing 2019 grid dynamics, but it performs significantly worse at simulating GO in 2021. Likewise, when the PCM is trained only on 2021 data, it does a worse job simulating GO in 2019 and 2020. A possible reason is that the best model version in 2021 uses a MILP formulation. Among the 3 years, the highest average daily natural gas prices are observed in 2021 (5.42\$/MMBtu), which increases the marginal cost of natural gas power plants and generally favors coal plants in the dispatch order. However, without binary variables controlling the on/off status of coal power plants in LP versions, coal plants overproduce. An MILP formulation introduces new costs (like startup) and additional constraints (like minimum up and down time) for coal generators, which decreases their usage in the model yielding a closer match with the historical generation mix.

There are numerous possible ways to select the 'best' model version, including using other metrics, changing the weights of each metric in the selection stage, and selecting different model versions for each pricing hub. For example, although the best model for 2019 accurately mimics the LMPs in three California pricing hubs, it misses some price oscillations in the MidC and Palo Verde hubs (figure 7). If we only considered the MidC hub while selecting the best version, we would have chosen a model with 300-node topology, LP formulation, +0 MW transmission limit scaling, and 0% hurdle rate scaling. This would have increased the LMP  $R^2$  of MidC from 0.49 to 0.74. On the other hand, a model with 100-node topology, LP formulation, +2000 MW transmission limit scaling, and 0% hurdle rate scaling would work better for the Palo Verde hub. This would have increased the LMP  $R^2$  of Palo Verde from 0.57 to 0.7.

In addition, changing how we weigh the individual model performance metrics and/or which metrics are considered at all can strongly affect the PCM selection process. For instance, figure 10 shows  $R^2$  values from different model versions for each pricing hub as well as the demand-weighted average  $R^2$  value for the entire Western Interconnection. The 'best' model version for 2019 (i.e. the lowest ranked model when LMP  $R^2$ , LMP RMSE, and average generation mix are all considered) is designated with a red square. However, there are different model versions (like the 175-node MILP version with +500 MW transmission scaling or the 300-node LP version with +0 MW transmission scaling) with worse overall rankings but higher LMP  $R^2$  values (these are indicated by orange squares).

An important feature of GO is its ability to help power system modelers navigate the tradeoff between model fidelity and computational speed. Figure 11 illustrates this tradeoff for all 540 model versions in 2021. It took nearly 40 h for the best-performing model version to simulate 2021 grid dynamics on an hourly time step (cross in green box). As we can see, there are other possible versions in the orange box which have similar rankings (i.e. fidelity) but require much lower runtimes. The modeler might prefer another version in orange box to save computational time if the main interest of the research entails running the model hundreds or thousands of times to characterize system performance, e.g. under weather and climate variability and extremes. We can also see that if we also considered runtimes in model selection, the algorithm may have a tendency to select a version in orange box as those versions are closer to the red star which is the ideal point in this figure.

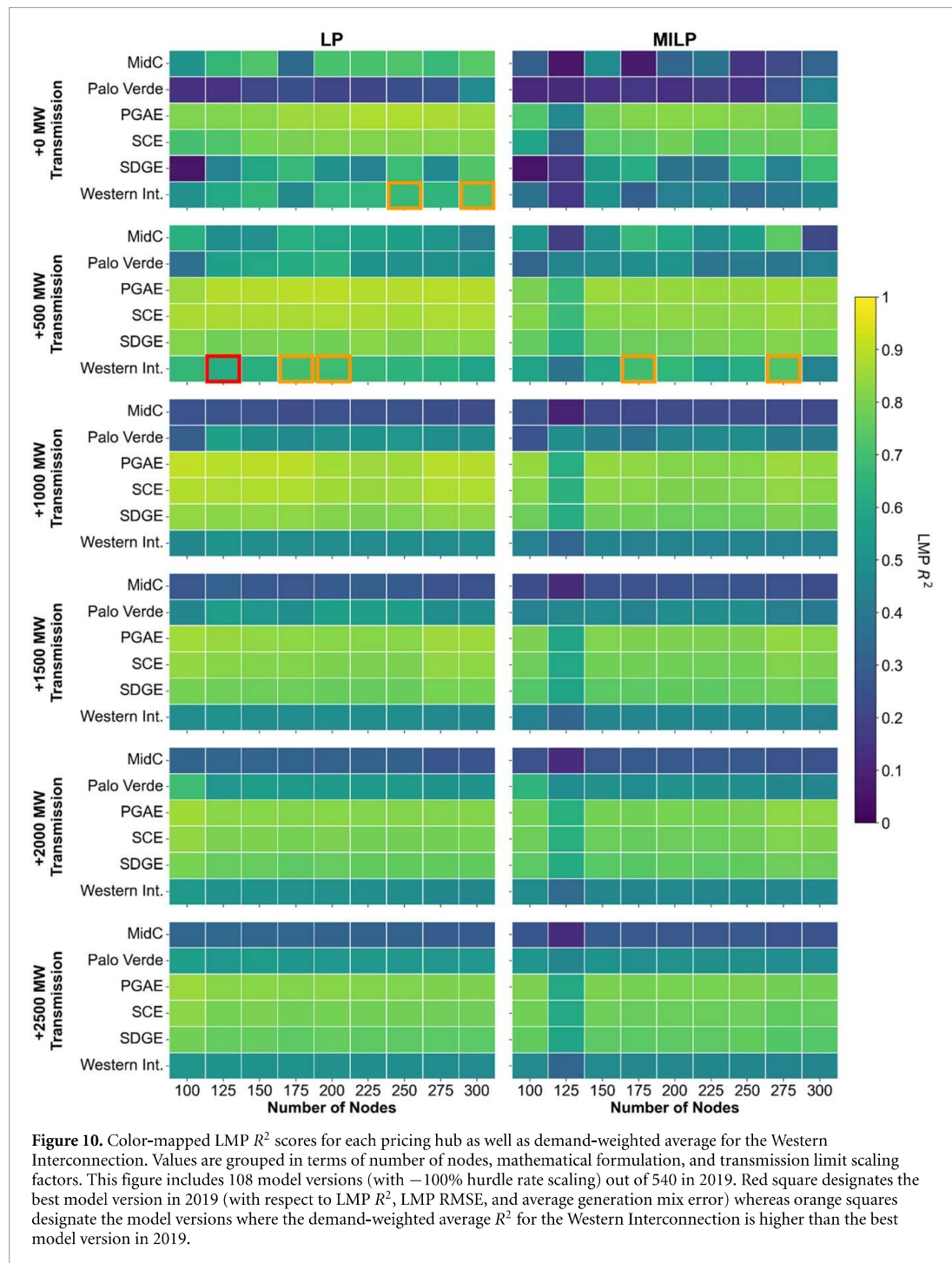
### 3.4. Experimental aims and computational limits

Ultimately, model selection should consider both the need for model fidelity and limitations on researchers' computational resources (e.g. available number of cores, memory size, runtime limitations, etc). The GO framework supports researchers in developing specialized s for their research questions. For instance, if a user is interested in individual generator operations or air pollution emissions, choosing a MILP formulation may be more appropriate. Exploring weather and climate uncertainty may recommend and faster LP. To expand on this, we revisit the hypothetical example described in the introduction. Specifically, a researcher would like to explore the impacts of future climate change on operations of the U.S. Western Interconnection, using an ensemble of different general circulation models (GCMs), and exploring a range of different shared socioeconomic pathways (SSPs) and representative concentration pathways (RCPs). In addition, for each unique climate state (GCM + SSP combination), the researcher would like to simulate GO using a 100 year Monte Carlo weather ensemble to explore stationary uncertainty.

The feasible scale of this experiment is likely to be constrained by available computational resources. For example, what if the hypothetical researcher can run 10 simulations simultaneously on an HPC cluster and has 45 d (or 1080 h) to complete the experiment? Table 3 shows how GO can be used to identify the best PCM parameterization for five different experimental designs (columns in table 3).

Moving left to right, the number of runs required by each experiment increases. To meet this requirement while staying within the allotted computational time, a researcher would need faster models that utilize lower number of nodes and sometimes LP formulations. For example, moving from an experiment that requires 200 runs to one that requires 400, the best available model (that is also feasible) uses an MILP formulation, but the topology must be reduced from 225 nodes to 125. For any experiment

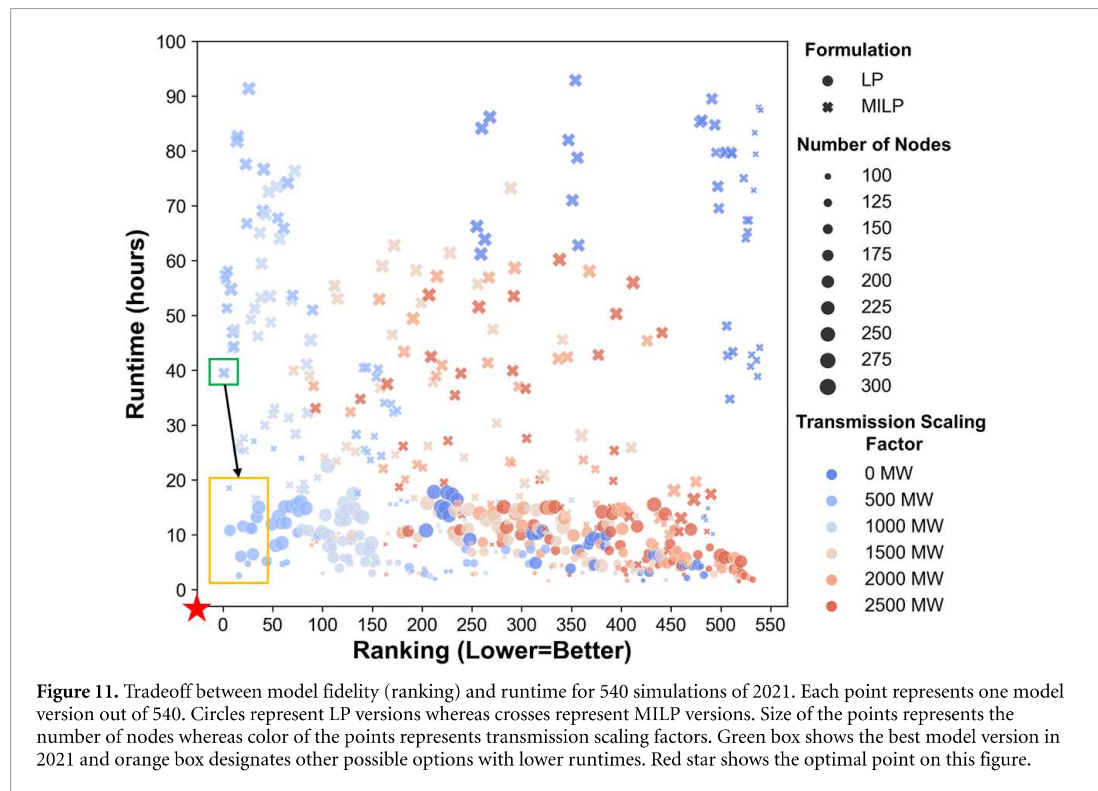




involving at least 800 runs, an LP model is needed to match the computational resource requirements. As the number of runs required increases further beyond this point, the best available models must utilize simpler networks. Transmission scaling factors increase slightly to create a less transmission-constrained system that can be solved more quickly.

#### 4. Limitations and future work

GO comes with some limitations which double as areas for future work. First, transmission and hurdle rate scaling factors uniformly impact all lines and BA-to-BA power transactions. Although it may require more costly calibration efforts to search for an optimal scaling factor for each individual line and BA-to-BA transaction, selectively altering the individual transmission line capacities and hurdle rates might create a



**Table 3.** Number of available model versions and parameters of the best available model version under varying numbers of runs for uncertainty analysis. The data in this table refers to 2021 simulation results.

	1 GCM, 1 RCP, 2SSPs, 100 weather years	1 GCM, 2 RCPs, 2 SSPs, 100 weather years	2 GCMs, 2 RCPs, 2 SSPs, 100 weather years	1 GCM, 4 RCPs, 4 SSPs, 100 weather years	4 GCMs, 4 RCPs, 4 SSPs, 100 weather years
Runs required	200	400	800	1600	6400
Number of feasible model versions	540	535	534	525	214
Number of nodes	225	125	225	125	100
Mathematical Formulation	MILP	MILP	LP	LP	LP
Transmission Scaling Factor	+500 MW	+500 MW	+500 MW	+500 MW	+1500 MW
Hurdle Rate	-100%	-100%	100%	-100%	-100%
Scaling Factor					
1 year runtime (hours)	39.5	18.5	10.8	2.6	1.6

more accurate model version. Second, the GO PCM assumes a single, central operator with perfect foresight of load and generation resources beside the probabilistic outage. Integrating forecast errors and balancing (real-time) markets could improve realism in certain cases. Last, fuel price information for many BAs is limited. A spatial algorithm was used to generate fuel price time series for some BAs (see supplementary information for more details). Using more granular and reliable fuel price information would most likely enhance our ability to replicate LMPs more accurately.

### 5. Conclusion

Accurately representing power system dynamics over a wide range of operating conditions is critical for performing vulnerability analysis. This imposes a challenge on researchers with computational budgets to strike a balance between model fidelity and computational speed.

To that end, this paper introduces GO, a framework for training and testing scale-adaptive open-source PCMs on the U.S. interconnection scale. GO allows users to search over many parameterizations of a PCM and identify versions that adequately balance model fidelity and computational speed. In an application over the Western U.S., our results show that simplified PCMs utilizing LP formulations and significantly reduced networks can adequately capture LMPs and generation mix. This result quantitatively supports the potential for large, stochastic simulation experiments using open-source PCMs, including experiments designed to characterize risks from climate and weather variability and extremes. Other findings of note include the interplay between transmission line scaling and network reduction in model calibration; essentially, transmission line scaling allows users to maintain network dynamics (and model accuracy) while reducing system complexity (and model runtimes). We also show that model selection is sensitive to choices around training data (e.g. weather year) and testing data (e.g. market or regional subsystem of interest). Finally, we provide a salient example of how experimental design (e.g. the scale of an uncertainty analysis), in the presence of limits on computational resources, can lead researchers to choose different versions of a PCM [58] serves as an example of how GO can be utilized for a renewable energy integration study in Western Interconnection.

## 6. Software and data availability

The model is open-source and publicly available. All codes of the model and data used are available under MIT free software license [59]. All model outputs utilized in this study are available under Creative Commons Attribution 4.0 International license [60].

### Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: <https://doi.org/10.57931/1923267>.

### Acknowledgments

This research was supported by the US Department of Energy, Office of Science, as part of research in the MultiSector Dynamics, Earth and Environmental System Modeling Program.

### ORCID iDs

Kerem Ziya Akdemir  <https://orcid.org/0000-0002-4532-3319>

Konstantinos Oikonomou  <https://orcid.org/0000-0003-0785-2102>

Jordan D Kern  <https://orcid.org/0000-0002-1999-0628>

Nathalie Voisin  <https://orcid.org/0000-0002-6848-449X>

Henry Ssembatya  <https://orcid.org/0009-0005-1141-3163>

### References

- [1] Akdemir K Z, Kern J D and Lamontagne J 2022 Assessing risks for New England's wholesale electricity market from wind power losses during extreme winter storms *Energy* **251** 123886
- [2] Hill J, Kern J, Rupp D E, Voisin N and Characklis G 2021 The effects of climate change on interregional electricity market dynamics on the U.S. West Coast *Earth's Future* **9** e2021EF002400
- [3] Oikonomou K, Tarroja B, Kern J and Voisin N 2022 Core process representation in power system operational models: gaps, challenges, and opportunities for multisector dynamics research *Energy* **238** 122049
- [4] Su Y, Kern J D, Denaro S, Hill J, Reed P, Sun Y, Cohen J and Characklis G W 2020 An open source model for quantifying risks in bulk electric power systems from spatially and temporally correlated hydrometeorological processes *Environ. Modelling Softw.* **126** 104667
- [5] van Vliet M T H, Wiberg D, Leduc S and Riahi K 2016 Power-generation system vulnerability and adaptation to changes in climate and water resources *Nat. Clim. Change* **6** 375–80
- [6] Mideksa T K and Kallbekken S 2010 The impact of climate change on the electricity market: a review *Energy Policy* **38** 3579–85
- [7] Panteli M and Mancarella P 2015 Influence of extreme weather and climate change on the resilience of power systems: impacts and possible mitigation strategies *Electr. Power Syst. Res.* **127** 259–70
- [8] Pleßmann G and Blechinger P 2017 How to meet EU GHG emission reduction targets? A model based decarbonization pathway for Europe's electricity supply system until 2050 *Energy Strategy Rev.* **15** 19–32
- [9] Samaan N, Milligan M, Hunsaker M and Guo T 2015 Three-stage production cost modeling approach for evaluating the benefits of intra-hour scheduling between balancing authorities 2015 *IEEE Power & Energy Society General Meeting (IEEE)* pp 1–5
- [10] Nsanzeza R, O'Connell M, Brinkman G and Milford J B 2017 Emissions implications of downscaled electricity generation scenarios for the western United States *Energy Policy* **109** 601–8
- [11] Voisin N, Kintner-Meyer M, Skaggs R, Nguyen T, Wu D, Dirks J, Xie Y and Hejazi M 2016 Vulnerability of the US western electric grid to hydro-climatological conditions: how bad can it get? *Energy* **115** 1–12

- [12] Jorgenson J, Denholm P and Mai T 2018 Analyzing storage for wind integration in a transmission-constrained power system *Appl. Energy* **228** 122–9
- [13] Deane J P, Drayton G and Ó Gallachóir B P 2014 The impact of sub-hourly modelling in power systems with significant levels of renewable generation *Appl. Energy* **113** 152–8
- [14] Cohen S M, Dyreson A, Turner S, Tidwell V, Voisin N and Miara A 2022 A multi-model framework for assessing long- and short-term climate influences on the electric grid *Appl. Energy* **317** 119193
- [15] Brouwer A S, van den Broek M, Zappa W, Turkenburg W C and Faaij A 2016 Least-cost options for integrating intermittent renewables in low-carbon power systems *Appl. Energy* **161** 48–74
- [16] Ibanez E, Magee T, Clement M, Brinkman G, Milligan M and Zagana E 2014 Enhancing hydropower modeling in variable generation integration studies *Energy* **74** 518–28
- [17] Behboodi S, Chassin D P, Djilali N and Crawford C 2017 Interconnection-wide hour-ahead scheduling in the presence of intermittent renewables and demand response: a surplus maximizing approach *Appl. Energy* **189** 336–51
- [18] O’Connell, Voisin N, Macknick and Fu 2019 Sensitivity of Western U.S. power system dynamics to droughts compounded with fuel price variability *Appl. Energy* **247** 745–54
- [19] Ding Y, Li M, Abdulla A, Shan R, Gao S and Jia G 2021 The persistence of flexible coal in a deeply decarbonizing energy system *Environ. Res. Lett.* **16** 064043
- [20] Electric Grid Test Case Repository Texas A&M University Electric Grid Datasets (available at: <https://electricgrids.engr.tamu.edu/>) (Accessed 2 March 2023)
- [21] OpenStreetMap Foundation, OpenStreetMap (available at: [www.openstreetmap.org/](http://www.openstreetmap.org/)) (Accessed 6 June 2023)
- [22] Brown T, Hörsch J and Schlachtberger D 2018 PyPSA: python for power system analysis *J. Open Res. Softw.* **6** 4
- [23] Exascale Computing Project EXASGD (available at: [www.exascaleproject.org/research-project/exasgd/](http://www.exascaleproject.org/research-project/exasgd/)) (Accessed 6 June 2023)
- [24] NREL *Sienna* (available at: [www.nrel.gov/analysis/sienna.html](http://www.nrel.gov/analysis/sienna.html)) (Accessed 6 June 2023)
- [25] Oh H 2010 A new network reduction methodology for power system planning studies *IEEE Trans. Power Syst.* **25** 677–84
- [26] Biener W and Garcia Rosas K R 2020 Grid reduction for energy system analysis *Electr. Power Syst. Res.* **185** 106349
- [27] Svendsen H G 2015 Grid model reduction for large scale renewable energy integration analyses *Energy Proc.* **80** 349–56
- [28] Shawhan D L et al 2014 Does a detailed model of the electricity grid matter? Estimating the impacts of the regional greenhouse gas initiative *Resour. Energy Econ.* **36** 191–207
- [29] Galván A, Haas J, Moreno-Leiva S, Osorio-Aravena J C, Nowak W, Palma-Benke R and Breyer C 2022 Exporting sunshine: planning South America’s electricity transition with green hydrogen *Appl. Energy* **325** 119569
- [30] Dranka G G, Ferreira P and Vaz A I F 2021 A review of co-optimization approaches for operational and planning problems in the energy sector *Appl. Energy* **304** 117703
- [31] Jiang Z, Tong N, Liu Y, Xue Y and Tarditi A G 2020 Enhanced dynamic equivalent identification method of large-scale power systems using multiple events *Electr. Power Syst. Res.* **189** 106569
- [32] Krishnan V, Ho J, Hobbs B F, Liu A L, McCalley J D, Shahidehpour M and Zheng Q P 2016 Co-optimization of electricity transmission and generation resources for planning and policy analysis: review of concepts and modeling approaches *Energy Syst.* **7** 297–332
- [33] LaRocca S, Johansson J, Hassel H and Guikema S 2015 Topological performance measures as surrogates for physical flow models for risk and vulnerability analysis for electric power systems *Risk Anal.* **35** 608–23
- [34] O’Neill R P, Krall E A, Hedman K W and Oren S S 2013 A model and approach to the challenge posed by optimal power systems planning *Math. Program* **140** 239–66
- [35] Hamilton W T, Husted M A, Newman A M, Braun R J and Wagner M J 2020 Dispatch optimization of concentrating solar power with utility-scale photovoltaics *Optim. Eng.* **21** 335–69
- [36] Bistline J E T 2021 The importance of temporal resolution in modeling deep decarbonization of the electric power sector *Environ. Res. Lett.* **16** 084005
- [37] Daraeepour A, Patino-Echeverri D and Conejo A J 2019 Economic and environmental implications of different approaches to hedge against wind production uncertainty in two-settlement electricity markets: a PJM case study *Energy Econ.* **80** 336–54
- [38] Schyska B U, Kies A, Schlott M, Von Bremen L and Medjroubi W 2021 The sensitivity of power system expansion models *Joule* **5** 2606–24
- [39] Lara J D, Lee J T, Callaway D S and Hodge B-M 2020 Computational experiment design for operations model simulation *Electr. Power Syst. Res.* **189** 106680
- [40] Voisin N, Kintner-Meyer M, Wu D, Skaggs R, Fu T, Zhou T, Nguyen T and Kraucunas I 2018 Opportunities for joint water–energy management: sensitivity of the 2010 Western U.S. Electricity Grid Operations to climate oscillations *Bull. Am. Meteorol. Soc.* **99** 299–312
- [41] Birchfield A B, Xu T, Gegner K M, Shetye K S and Overbye T J 2017 Grid structural characteristics as validation criteria for synthetic networks *IEEE Trans. Power Syst.* **32** 3258–65
- [42] Electric Grid Test Case Repository ACTIVSg10k: 10000-bus synthetic grid on footprint of western United States (available at: <https://electricgrids.engr.tamu.edu/electric-grid-test-cases/activsg10k/>) (Accessed 9 January 2023)
- [43] Electric Grid Test Case Repository ACTIVSg70k: 70,000 bus synthetic grid on footprint of eastern United States (available at: <https://electricgrids.engr.tamu.edu/electric-grid-test-cases/activsg70k/>) (Accessed 3 May 2023)
- [44] Electric Grid Test Case Repository ACTIVSg2000: 2000-bus synthetic grid on footprint of Texas (available at: <https://electricgrids.engr.tamu.edu/electric-grid-test-cases/activsg2000/>) (Accessed 3 May 2023)
- [45] Shi D et al 2012 Optimal generation investment planning: pt. 1: network equivalents 2012 *North American Power Symp. (NAPS)* (IEEE) pp 1–6
- [46] Ward J B 1949 Equivalent circuits for power-flow studies *Trans. Am. Inst. Electr. Eng.* **68** 373–82
- [47] Kundur P S and Malik O P 2022 *Power System Stability and Control* (McGraw Hill) (available at: [www.mhprofessional.com/power-system-stability-and-control-second-edition-9781260473544-usa](http://www.mhprofessional.com/power-system-stability-and-control-second-edition-9781260473544-usa))
- [48] Zhu Y and Tylavsky D 2018 An optimization-based DC-network reduction method *IEEE Trans. Power Syst.* **33** 2509–17
- [49] EIA Hourly Electric Grid Monitor (available at: [www.eia.gov/electricity/gridmonitor/dashboard/electric\\_overview/US48/US48](http://www.eia.gov/electricity/gridmonitor/dashboard/electric_overview/US48/US48)) (Accessed 2 October 2023)
- [50] CAISO 2022 *Annual Report on Market Issues & Performance*
- [51] EIA Form EIA-923 detailed data with previous form data (EIA-906/920) (available at: [www.eia.gov/electricity/data/eia923/](http://www.eia.gov/electricity/data/eia923/)) (Accessed 4 October 2023)

- [52] NERC *Generating Availability Data System (GADS)* (available at: [www.nerc.com/pa/RAPA/gads/Pages/GeneratingAvailabilityDataSystem-\(GADS\).aspx](http://www.nerc.com/pa/RAPA/gads/Pages/GeneratingAvailabilityDataSystem-(GADS).aspx)) (Accessed 26 September 2023)
- [53] WECC 2030 ADS PCM Release Notes (available at: [www.wecc.org/Reliability/2030ADS\\_PCM\\_ReleaseNotes\\_GV-V2.3\\_6-9-2021.pdf](http://www.wecc.org/Reliability/2030ADS_PCM_ReleaseNotes_GV-V2.3_6-9-2021.pdf)) (Accessed 10 January 2023)
- [54] Homeland Infrastructure Foundation-Level Data Control Areas (available at: <https://hifld-geoplatform.opendata.arcgis.com/datasets/geoplatform::control-areas/explore?location=39.644053%2C-112.615515%2C4.91>) (Accessed 9 February 2023)
- [55] CAISO California ISO OASIS (available at: <http://oasis.caiso.com/mrioasis/logon.do>) (Accessed 2 October 2023)
- [56] EIA Wholesale electricity and natural gas market data (available at: [www.eia.gov/electricity/wholesale/](http://www.eia.gov/electricity/wholesale/)) (Accessed 2 October 2023)
- [57] Homeland Infrastructure Foundation-Level Data Electric retail service territories (available at: <https://hifld-geoplatform.opendata.arcgis.com/datasets/geoplatform::electric-retail-service-territories-2/explore?location=38.884462%2C-114.257812%2C5.75>) (Accessed 9 February 2023)
- [58] Akdemir K Z, Robertson B, Oikonomou K, Kern J, Voisin N, Hanif S and Bhattacharya S 2023 Opportunities for wave energy in bulk power system operations *Appl. Energy* **352** 121845
- [59] Akdemir K Z, Oikonomou K, Kern J and Voisin N GO-WEST v1.0.0 *Zenodo* (<https://doi.org/10.5281/ZENODO.10067714>) (Accessed 2 November 2023)
- [60] Akdemir K Z, Oikonomou K, Kern J D and Voisin N 2023 IM3 GO WEST parameter search dataset *MSD-LIVE* (<https://doi.org/10.57931/1923267>)