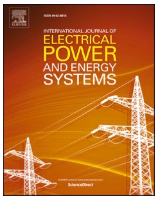
Electrical Power and Energy Systems 159 (2024) 110037

Contents lists available at ScienceDirect

International Journal of Electrical Power and Energy Systems

journal homepage: www.elsevier.com/locate/ijepes



Building and validating a Large-Scale combined transmission & distribution synthetic electricity system of Texas

Carlos Mateo [a,*], Fernando Postigo [a,1], Tarek Elgindy [b], Adam B. Birchfield [c], Pablo Dueñas [d,a], Bryan Palmintier [b], Nadia Panossian [b], Tomás Gómez [a], Fernando de Cuadra [a], Thomas J. Overbye [c], Farnaz Safdarian [c], Diana Wallison [c,2]

[a] Institute for Research in Technology (IIT), School of Engineering (ICAI), Comillas Pontifical University, Spain
[b] National Renewable Energy Laboratory, Golden, CO, USA
[c] Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX, USA
[d] Massachusetts Institute of Technology, Cambridge, MA, USA

ARTICLE INFO

Keywords:
Power system
Test system
Distribution
Transmission
Networks
Planning
Co-simulation
Power flow

ABSTRACT

Distributed energy resources, such as rooftop solar, have rapidly expanded in recent years, given declining costs and the desire to reduce carbon emissions. With more energy resources located in the lower-voltage distribution system, it is increasingly helpful to utilize combined transmission and distribution (T&D) system models to analyze interactions between these normally-distinct subsystems. This paper proposes a methodology for creating very large-scale, highly detailed, combined T&D systems that are synthetic—that is, free from non-public data—yet still realistic. The methodology creates very large-scale combined T&D systems by merging the most up-to-date techniques for creating synthetic distribution feeder networks with the latest methods for building synthetic, meshed bulk-power transmission networks. This methodology is demonstrated on a T&D system geolocated in Texas, and benchmarked with co-simulation results. Validation demonstrates that the resulting *syn*-texas-TDgrid synthetic test system realistically represents characteristics found in actual networks, addressing the lack of available T&D test systems. With over 15,000 feeders and 46 million electrical nodes, this T&D dataset has applications for research in optimal power flow algorithms, voltage control, reconfiguration, and T&D coordination schemes under high adoption of distributed energy resources.

1. Introduction

Decarbonization to avoid climate change requires achieving a more renewable energy mix. Wide-scale adoption of solar and wind on both transmission and distribution (T&D) scales provides an opportunity to realize this transformation. However, the associated variability and uncertainty pose new technical challenges for the planning and operations of T&D. Simultaneously increased electrification (e.g., electric vehicles, electric heating) and increased distributed energy resources (DERs) such as storage and responsive demand are also being integrated into power systems. These new resources offer opportunities to increase the flexibility of the system, but can introduce capacity and operational challenges.

In this context, transmission and distribution system operators (TSOs & DSOs) need to define a framework for cooperation, since large adoptions of DERs could introduce contradictory signals in their networks, such as if DERs managed for TSO needs create challenges for the DSO. Several authors have proposed TSO-DSO coordination schemes

---

* Corresponding author.
E-mail address: cmateo@comillas.edu (C. Mateo).

[1] IIT-Comillas when he collaborated in developing the work of this paper.

[2] This work was authored in part by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. Funding provided by the Advanced Research Projects Agency–Energy. A portion of the research was performed using computational resources sponsored by the Department of Energy's Office of Energy Efficiency and Renewable Energy and located at the National Renewable Energy Laboratory. The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes.

https://doi.org/10.1016/j.ijepes.2024.110037
Received 21 December 2023; Received in revised form 18 April 2024; Accepted 7 May 2024

[1,2]. However, detailed simulations and analysis of such schemes are complicated by limited access and difficult data clean-up with actual system data. These challenges can be overcome with realistic open-access test systems; however, such test systems have historically not existed at the scale comparable to the actual grid [1]. This is particularly challenging for distribution, where test systems have historically had a very narrow scope, often limited to one or a very few feeders [3]. The IEEE radial test feeders are the first test systems for U.S. distribution systems [4]. Until the late 2010′s, one of the largest public test systems was the IEEE 8500-node test feeder, with only 8,500 nodes and only two voltage levels [5]. Other test systems have been provided focusing on specific aspects of distribution systems, but also with a small scale [6,7,8]. In addition, other institutions have also provided test systems, like the EPRI Representative Feeders and the PG&E Prototypical Feeders, but they are also very limited in terms of scale, representativeness and topological layout [3]. Refs. [9,10] are examples of non-synthetic test systems, based on actual data mainly about LV systems.

In 2016, ARPA-E launched the GRID DATA call [11], funding the development of large-scale, realistic, validated, and open-access power system network models to allow the successful development and testing of transformational power system optimization and control algorithms. The program funded the creation of 2 repositories, 5 projects that created transmission models, and one that created a distribution model. As a result of the distribution model project, in 2020, several authors of this paper presented the start of a new generation of test systems, with large-scale data sets, up to about 10 million nodes [12,13], extensively validated [14], which are orders of magnitude larger and more complex than previous test feeders. However, these test systems are limited to only distribution systems, preventing detailed analysis about TSO-DSO coordination. Since then, several datasets have been provided in the literature [15,16,17]. Particularly, [17] claims to having produced a large-scale T&D system, but without the geographical layout of the network and with 22,000 buses the realism and scale is not comparable to the dataset that we present in this paper, which models the state of Texas with 46 million nodes.

The creation of large-scale models of the power system can be thought of as a planning problem, where the main objective is to minimize investment and operation costs, supply consumers, and connect generation at all scales. This effort is subject to technical and geographical constraints, including ensuring adequate levels of energy losses and quality of service. As such, synthetic test systems can be built using planning algorithms to locate and size substations, and design the network while also determining the location of additional components like switching and voltage control devices [13,18,19,20,21].

Network planning in utilities is achieved by dedicated engineers proposing candidate solutions for portions of the system and analyzing them with specialized software such as PSS/E, CYME, and DIgSILENT [22,23]. However, this type of planning approach is not practical for producing very large-scale test systems. On the other hand, the research community has traditionally used mathematical programming to obtain optimal designs [24]. Such approaches have strong computational limitations and are only feasible for small-scale systems. The major alternatives that can be used to obtain very-large systems are based on heuristics or *meta*-heuristics [25]. These algorithms can obtain a good design solution much faster. They do not necessarily reach the global optimum, but this fact can actually add to the realism since actual networks also deviate from the global optimum. Recent developments to create synthetic test systems are based also on heuristics [26], and with the advances of artificial intelligence, on generative adversarial networks [27]. However, this paper outperforms these solutions in two relevant dimensions, 1) realism of the networks and variety of equipment and voltage levels; and 2) scale of the resulting data sets.

Network planning has many sub-elements, from substation and feeder design to voltage control and reliability improvement through switching devices. As such, a vast variety of algorithms has been proposed for each of these functionalities, ranging from Tabu Search and k-means for substations [28,29], to minimum spanning trees, Delaunay triangulations, evolutionary algorithms, and branch-exchange for feeders [25,30,31]. Additionally, voltage regulators, generator voltage control, tap-changing transformers, shunt capacitors, and shunt reactors can be installed to improve voltages and/or reduce energy losses [32]. Finally, switching devices, loops, and meshed networks are required to evaluate reliability in distribution networks, or to apply N-1 criteria in transmission [33].

In [12,13], several of the authors of this paper presented a methodology to build very large-scale synthetic distribution systems, based on the U.S. Reference Network Model (RNM-US), that leverages the work of previous RNM models for Europe [34,35]. In [19,20], the remaining co-authors proposed a methodology to build synthetic transmission systems. In 2019, ARPA-E funded the collaboration of these two promising research approaches for building Transmission & Distribution (T&D) systems. As a result of this, the methodology of this paper creates very large-scale combined T&D systems by merging these most up-to-date techniques for creating synthetic distribution feeder networks with the latest methods for building synthetic, meshed bulk-power transmission networks. This builds on earlier work by the team to produce much smaller T&D datasets in [36] where we found the two paradigms must be carefully coordinated to produce a cohesive dataset that models the entire system from generator terminals to low-voltage customers. The major contribution of this paper is not the methodology itself, but rather the resulting dataset, which addresses the lack of combined T&D test systems, by producing an unprecedented very large-scale dataset covering Texas. The illustration of the methodology serves to understand the characteristics of the underlying system. The key objective behind building this synthetic network is that the dataset—in its entirety or relevant subsets—remains a challenge for a wide range of power systems applications, both today and for years to come. Thereby it can serve as a foundational test system to benchmark diverse power systems algorithms from across the T&D spectrum including bulk-system-aware DER management through DER management systems (DERMS), market and service interactions between large numbers of DERs and bulk systems, advanced power flow and optimal power flow algorithms and applications. Its large scale can also support artificial intelligence and machine learning algorithm development by providing a diverse set of training data to build a basic understanding that can then be adapted to specific real-world systems or algorithms. The research and development of such applications, which leverage the whole scale of the test system, is out of the scope of this paper, and is intended to be a challenge by itself for future research aiming at fostering the development of innovative high efficiency solutions, capable of working at the rapidly expanding scales needed by future utilities and system operators.

In Section 2, the methodology to build a combined T&D system is summarized and briefly described. Section 3 presents the proposed T&D system of Texas. In Section 4 the test system is validated. Co-simulation results are presented in Section 5. Finally, Section 6 discusses potential applications and Section 7 concludes the paper.

2. Methodology to build a combined T&D system

In this section, we describe a methodology to automatically design transmission and distribution systems. We merge the most up-to-date planning techniques for building synthetic systems, combining the U. S. Reference Network Model to plan distribution [13], with the synthetic transmission grid generation approach from [19,20].

A combined T&D system includes interconnected networks of different voltages, as illustrated in Fig. 1. The system is categorized by voltage levels which are labeled as transmission or distribution. In this paper, the sub-transmission network is planned as part of the transmission system.

Each voltage level is supplied with a voltage source, as shown in Fig. 2. The sources for the low voltage networks are distribution transformers, the sources for medium voltage networks are distribution

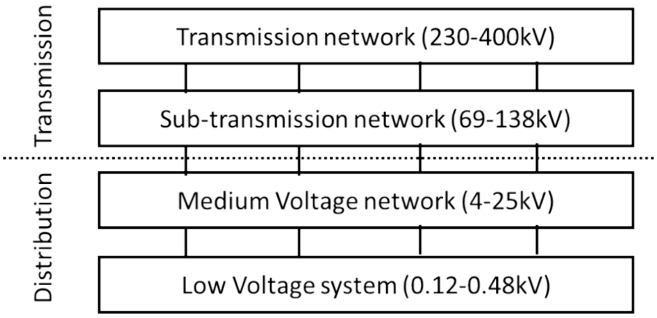


Fig. 1. Structure of a T&D System.

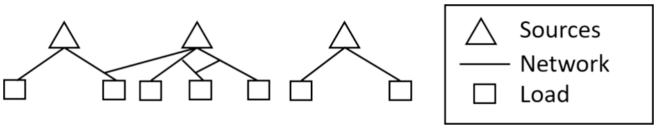


Fig. 2. Structure of a voltage level.

substations, and the sources for the sub-transmission networks are transmission substations. Additionally, generators are connected at various points in the transmission system, acting as sources of this voltage level. The loads of a voltage level can be either consumers directly connected to that voltage level or sources from the immediate lower voltage level (or even lower voltage levels).

The methodology followed to build a combined T&D system in this paper is bottom-up, starting from the lower voltage levels and designing one voltage level at a time. The steps within each voltage level are: 1) obtain the location and demand of loads, 2) plan sources locations, and 3) design the network that interconnects sources with loads. This methodology designs all the networks starting with only the information about the demand and location of consumers. The methodology is described hereinafter, and an overview of the main steps is presented in Fig. 3.

Consumers are identified by using building footprints, which are overlaid on top of parcel (or land-use) information. Public databases, like OpenStreetMap, or private vendors gather and provide building and land information worldwide. This combined information is then utilized to categorize each building by its purpose (for example, as a single-family or multi-family residential unit, retail store, or office space), by checking where each building's footprint falls within the marked land-use divisions. The location of each consumer is given by the service drop, i.e., the closest point of the building to the street, or road. The peak load is inferred from a database of building archetypes, such as ResStock or ComStock databases from NREL for the US. Another potential database is EPICOPE for Europe. The inference process allocates each real building to an archetype by considering metadata beyond the square footage and use, such as the vintage or land value. In addition, the system layout follows the layout of streets, which are obtained from publicly available street maps, like OpenStreetMap. The inputs for the next level are points of consumption with relevant electrical characteristics (peak load, voltage level, number of phases) and a layout following the street map.

For planning source assignments, clustering algorithms are required. We use a minimum spanning tree and remove branches to decide which loads from the previous stage are to be supplied from the same source [23]. For designing the network, minimum spanning trees and Delaunay triangulations are used for proposing initial designs, which are made feasible and improved using branch-exchange algorithms [37], while considering the thermal limits of power lines, the ANSI voltage limits, and the street map layout.

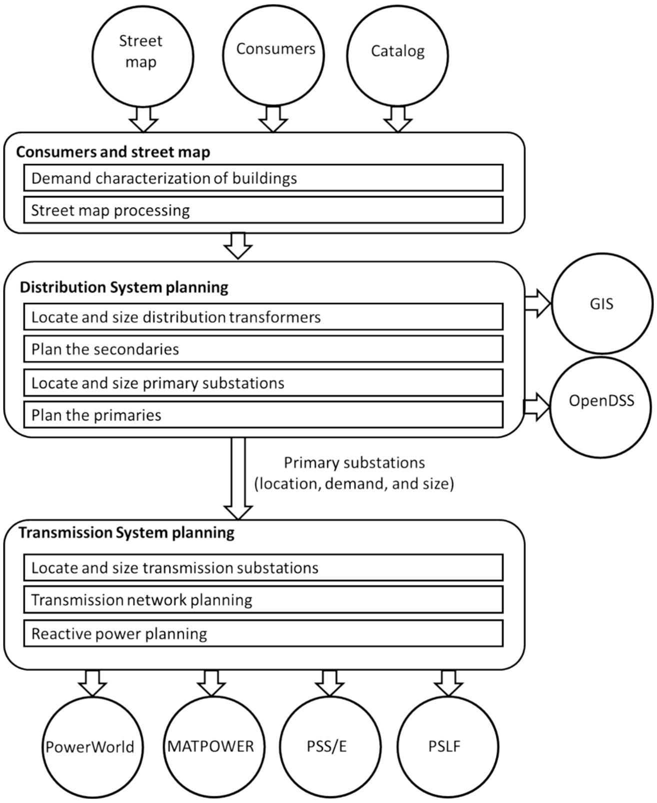


Fig. 3. Overview of the T&D methodology.

Each voltage level has distinct characteristics. For example, in a low-voltage system, a mix of single-phase center-tapped transformers and three-phase distribution transformers are placed to supply the loads, depending on their size. This system is designed as radial and is characterized by very short lines, where each distribution transformer supplies very few consumers. This level models the service drop from the distribution transformer located in the street to each building, following an engineering design by combining tree and star configurations [13].

Fig. 4 illustrates the placement of distribution transformers from the low-voltage clustering algorithm, while Fig. 5 illustrates the architecture of the secondaries, which connect the distribution transformers to each building.

In medium voltage (MV), the network is operated radially. This is designed with minimum spanning trees and branch-exchange algorithms. However, the structure of this network can contain loops with open switches, so algorithms are applied to assign switch locations and line redundancies, aiming to place them in locations that maximize reliability at a minimum cost. The MV network in the U.S. is characterized by the coexistence of single-, two- and three-phase circuit

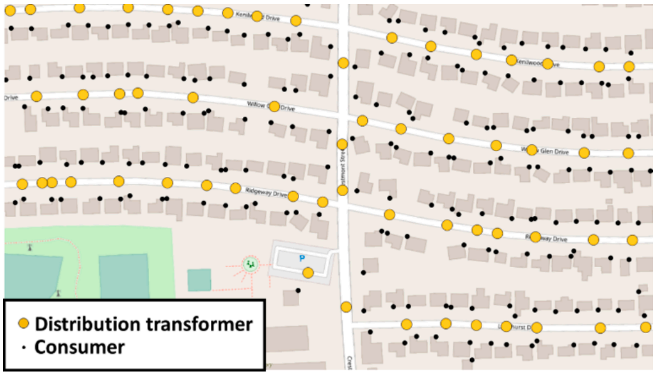


Fig. 4. Distribution transformer allocation.

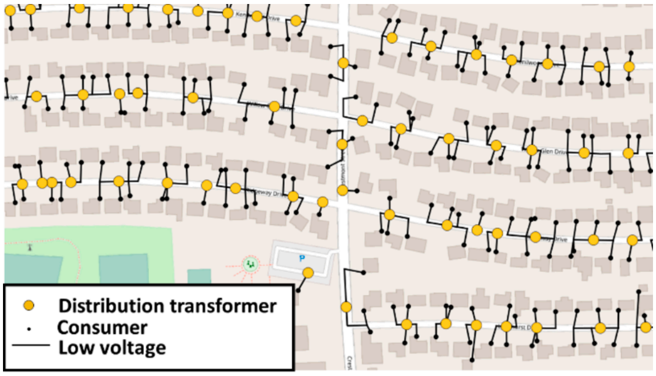


Fig. 5. Low-voltage network.

sections. This is modeled with algorithms that first assign the number of phases of each power line section, and then assign them to specific phases (A/B/C) [38]. The objective function for this purpose is the minimization of power imbalances. The MV network also must include voltage control devices (i.e. voltage regulators and capacitors). A recursive algorithm based on Depth-First-Search is applied to install voltage regulators to avoid voltage thresholds being violated, voltages being assessed with the backward-forward power flow method. Capacitor installation is guided by the reduction of energy losses, seeking to improve voltages as well [13]. After the base MV system is built, a post-processing step is used to identify control system set points and potentially additional voltage regulation devices to maintain adequate power quality across a wider range of operating conditions than are captured in the base planning model. Post-processing is also used to identify the internal topology of distribution substations by dividing the overall transformer load into multiple transformer banks and to produce a range of DER adoption scenarios, complete with time series load and solar PV production data [12].

Fig. 6 illustrates the placement of substations from the medium-voltage clustering algorithm, while Fig. 7 illustrates the architecture of the medium-voltage network. The radial network topology is used in the vast majority of distribution feeders, but in some dense urban centers, including downtown Houston and Austin, the actual system uses meshed low voltage Secondary Networks (SN). SNs are fundamentally different in their design and operation from radial distribution systems, and limited SN test feeders have been released such as the IEEE 342 node test system [39]. Moreover, algorithms for generating synthetic SNs are in their infancy and were not available to the team. For the sake of having a complete dataset, the synthetic Texas dataset, uses radial distribution feeders in all areas, including dense urban centers. Such use of radial feeders in dense downtown regions does exist in other parts of the U.S., including Los Angeles, so such designs are consistent with the general goal of producing realistic but not real synthetic grids. Still, in future versions of the Texas dataset or other similar datasets, including synthetic SNs to serve load in limited downtown areas of larger cities could provide a valuable addition to the models.

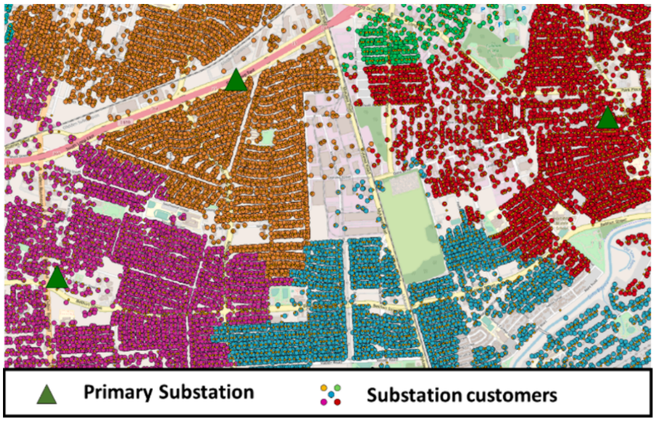


Fig. 6. Clusters and substation allocation.

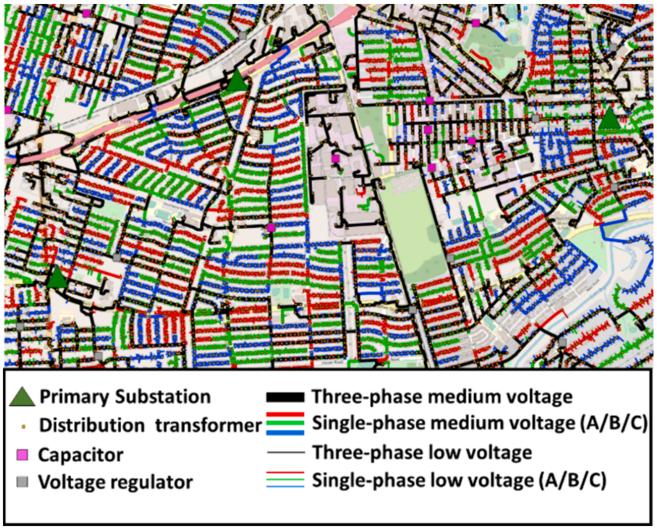


Fig. 7. Medium-voltage network.

The transmission system methodology starts with the substation planning process. In this process, the distribution feeders are considered loads and form the fragments that will make transmission substations. Each feeder can be considered an individual load, or, as is more common, they can be aggregated at a substation level. In addition to the distribution-connected load, additional load can be added to represent industrial customers who directly connect to the high-voltage network. The exact amounts and locations of transmission-connected loads depend on the area being modeled and the desired applications. Bulk power system generators also form transmission substations, with some smaller generators being added to existing transmission substations. Once the substations are defined with generator and load definitions, the substations are assigned voltage levels using the multi-area method described in the clustering process of [31]. With voltage levels assigned, the substations can be designed with buses and connecting transformers, including generator step-up transformers as appropriate.

The second step of the transmission system methodology is the transmission network planning process. This is a computationally intensive procedure because it must select the transmission lines that will form the backbone of the combined T&D network. The heuristic method defined in [30] is applied, which selects candidate lines using an iterative process that balances the desired characteristics of economic feasibility, topological realism, and engineering performance and reliability. As part of this process, the algorithm runs dc power flow analysis for N-1 contingency conditions, weighting more heavily candidate lines that are shown by sensitivity analysis to be favorable for improving N-1 contingency reliability. In the end, a selection of lines numbering about 1.1–1.2 times the number of substations is added at each voltage level, forming an interconnected mesh that serves the underlying distribution network.

The final step of the transmission system methodology involves reactive power planning, as discussed in [20]. This process moves the analysis from a DC power flow to an AC power flow by adding required reactive power and voltage control resources such as transformer tap changers, switched shunt capacitors, and inductors.

3. T&D synthetic system in Texas

Applying the methodology described in Section II, the *syn*-texas-

TDgrid dataset,[3] a synthetic T&D system of Texas has been built. Fig. 8 summarizes the estimation of residential (R), commercial (C), and industrial (I) loads in the counties that have been used to build this system.

The distribution system is comprised of urban and rural areas, with medium and low voltage levels, including primary substations (which interface with sub-transmission at 69 kV and 138 kV), medium voltage networks (4 kV, 12.47 kV, and 25 kV), distribution transformers and low voltage systems (120 V and 480 V). Besides the diversity in terms of nominal voltages, there is diversity in terms of voltage management strategies, using voltage regulators (VR), capacitors (CAP), or both. Additionally, some regions use delta-wye substations with single- and three-phase distribution circuit segments, while others are connected to delta-delta substations with two- and three-phase segments. The full Texas distribution system is split into 236 zones. Their diversity is summarized in Table 1.

The main metrics of the synthetic *T&D* system of Texas are summarized in Table 2. The power system supplies 14 million consumers, with over 4,000 transmission substations and 46 million electrical nodes. The design of the distribution system is very detailed, including distribution substations, distribution transformers, capacitors, voltage regulators, MV feeders, and LV network. The MV network comprises single-, two- and three-phase power lines.

The transmission system is depicted in Fig. 9, including the generators, with their locations and technology types, as well as the transmission power lines at three voltage levels.

The full synthetic T&D system is depicted in Fig. 10. It comprises 14,983 km of 345 kV Extra High Voltage (EHV) network, 66,080 km of 138 kV network, 16,301 km of 69 kV network, 671,316 km of MV network, and 264,016 km of LV lines.

4. Validation

4.1. Distribution

Following the methodology proposed in [14], the distribution system is validated by comparing it with generalized data and metrics derived from multiple U.S. utilities. For example, Fig. 11 compares the histograms of LV 3-phase line length from the synthetic distribution system with data from utilities.

The summary of all the validation metrics is shown in Fig. 12. In general, there is a good matching with the data from the utilities, being the major differences found in demand. This is due to the significantly higher number of rural feeders in Texas compared to the utility data used in the validation dataset. This can be seen by comparing histograms partitioned by the density of load in Fig. 13. While the distribution of load in high-density feeders in Texas matches well with that of the utility data, there are a lot more feeders with low demand in the very low population density regions of Texas than is present in the utility data. These very rural regions also impact other metrics, such as the number of switches, since switches use to be less common in such rural areas where there are fewer or no nearby feeders to interconnect with.

4.2. Transmission

The model of the T&D system is built bottom-up, starting with the demand of single consumers. This makes it necessary to verify afterwards that the total demand of the system in the actual system is consistent in aggregated terms with the demand assumed for each consumer. Fig. 14 shows the validation of the total load of the T&D system, comparing it with the ERCOT load. In general, there is a good match when the large industrial load is added to the total load of the distribution system.

The transmission system validation is done with a variety of metrics, some of which are given in [40]. The validation process recognizes that actual grids are quite diverse, and that characteristics of one grid might not match another due to differences in geography, priorities, or engineering design decisions. But the purpose is to show that the synthetic grid, considered from many angles, generally falls in the statistical ranges of actual grids. The validation process serves as a screening mechanism to flag any potential concerns about the realism of the grid for further consideration.

Several main categories of metrics are considered. First, there are statistics related to the overall system proportions. The number of buses in a substation, load versus generator substations, the amount of load present at a typical bus, as well as the distribution of generator capacities, help to ensure that the grid as a whole is structured correctly. Second, there are metrics related to the parameters of individual elements. Examples of metrics in this category include generator P and Q capacities, transmission line per-length reactance, transformer X/R ratio, and transmission line capacity as a function of its nominal voltage level. Certain metrics are physically constrained—for example, the line capacitance and inductance are related to the line length by the propagation speed, which should be quite close to the speed of light. Third, there are metrics about the transmission line network topology. These include purely topological metrics, such as the node degree distribution, clustering coefficient, and average shortest path building synthetic length. Also included are metrics related to the geographic layout of the network, such as the number of graph crossings and the overlap with the Delaunay Triangulation. One aspect of this is that there are multiple alternative paths (loops) within the transmission network, much more than in distribution. By the nature of its structure, this case has a 138 kV or 69 kV path as an alternate route for each 345 kV line, and there are 17 lines that form a direct parallel with differing voltage levels that can constrain the power flowing on the higher voltage line. Fourth, there are engineering performance metrics, which include a full contingency assessment using a contingency set with N-1 and selected N-2 contingencies. The branch loading and voltage magnitude performance in these contingency solutions are part of the analysis to assess grid realism for power flow studies. Table 3 shows a selection of the metrics checked for *syn*-texas-TDgrid-v02. More details on the validation process, along with tables against which the transmission line and transformer parameters are compared, can be found in [40]. The T-D interface is also validated by ensuring that the transmission is represented as a constant voltage source behind a reactance, appropriate for every feeder in its place in the transmission grid, and that the load P and Q in transmission are consistent with the loading and losses of the combined associated distribution feeders (plus any transmission-connected industrial loading).

5. Example co-simulation of T&D

One of the applications of a synthetic T&D system is simulating different control approaches, getting a full picture across transmission and distribution. This poses a number of computational and modeling challenges, particularly at full scale. It is first worth noting that given the long view of creating a dataset to support and challenge grid algorithm development for years to come, fitting within the limitations of today's computing was not a requirement. It is expected that computers in coming years will readily be able to compute with data sets that may be difficult to even load today. Still, the systems were designed for practical reasons to not have overly bloated representations to help manage memory issues with the large size of the datasets. For the transmission part, the memory requirements are not a concern. With 7000 nodes, even 100 MB is plenty sufficient. The distribution portion can create

---

[3] The naming convention uses hyphens to connect each uses "syn" to highlight the data is synthetic, "texas" to describe the geographic extent, "TDgrid" highlight that the system captures transmission and distribution, rather than T_only or D_only for transmission vs. distribution datasets respectively. Since we expect that some refinements and/or bug fixes over time, an identifier is added when referring to a specific version, such as *syn*-texas-TDgrid-v02.

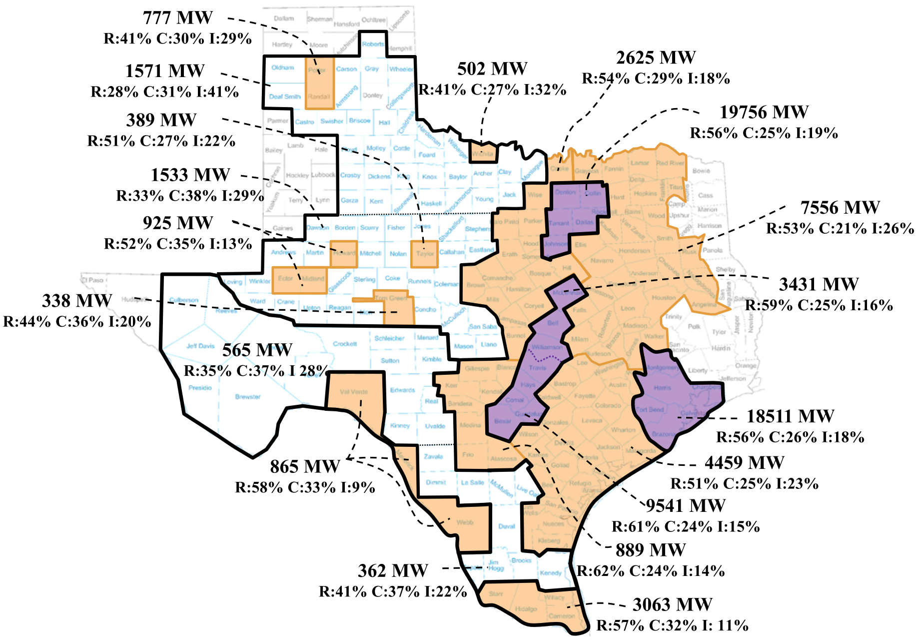


Fig. 8. Estimation of load in the *syn*-texas-TDgrid-v02 synthetic T&D system of Texas. Load per area in MW and percentage of Residential, Commercial, and Industrial load. Color legend: purple, urban areas at 69 kV/12.47 kV; brown, sub-urban areas at 115 kV/12.47 kV; and white, rural areas at 138 kV/25 kV. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 1
Distribution system diversity.

Medium Nominal Voltage	138 / 12.47 kV	138 / 25 kV	69 / 12.47 kV	69 / 4 kV	Total
# Urban zones	36	0	124	5	165
# Rural zones	55	16	0	0	71
#Zones with delta-wye substations	87	16	114	5	222
#Zones with delta-delta substations	4	0	10	0	14
#Zones with VR	38	5	43	0	86
#Zones with CAP	29	0	42	5	76
#Zones with VR and CAP	24	11	39	0	74
#Total	91	16	124	5	236

Table 2
Main metrics of the syn-texas-TDgrid-v02.

# LV Customers	14,381,174
# MV Customers	35,270
# HV Customers	584
# Substations	4,894
# Primary (MV) Feeders	15,875
# Distribution Transformers	3,860,927
Total Distribution transformer capacity (MVA)	170,312
# Capacitors	12,724
# Voltage regulators	7,346
LV Length (km)	264,016
MV Length (km)	671,316
HV Length (km)	82,381
EHV Length (km)	14,983
# Electrical Nodes	46,993,633

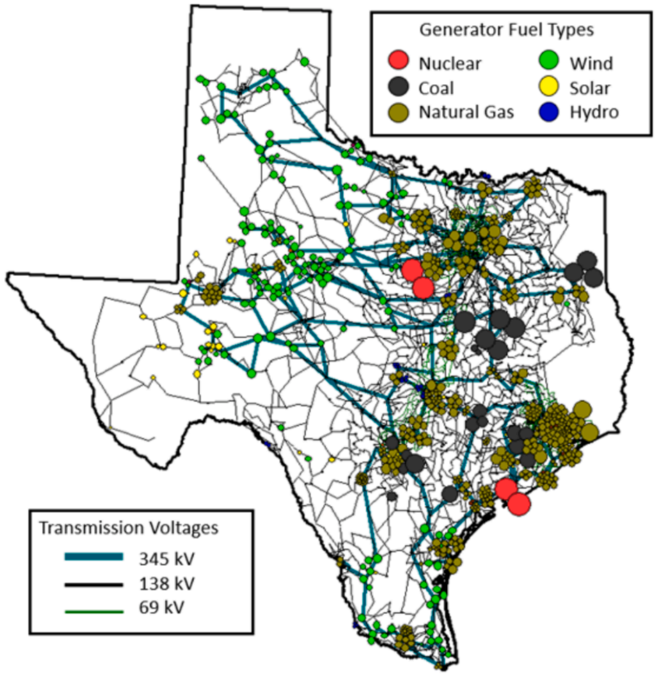


Fig. 9. Transmission network voltage levels, generator fuel types.

memory challenges for today's computers given its tens of millions of electric nodes and detailed 8760-hour profiles. As a result, the distribution dataset is constructed in a way to both share common data elements that reduce duplication and enable partial analysis for usage at smaller scales and decomposition for scaling. For instance, rather than having an 8760 profile for every single customer, we have a (large) set of relevant representative customer profiles that are referred to multiple times. And the decomposable structure makes it possible to not have to load all of the feeders at once into a single machine. For many of today's applications, working with a subset of the distribution system in combination with the transmission system may be sufficient. And for the full dataset, the data management and computing can be distributed across multiple datasets and then combined, such as through co-simulation. A second issue is slowdowns with iterations used to converge across T&D models such as local solar PV control operations. A third issue is that the transmission system is typically modeled with a positive-sequence single phase equivalent, rather than the full three-phase modeling typical for

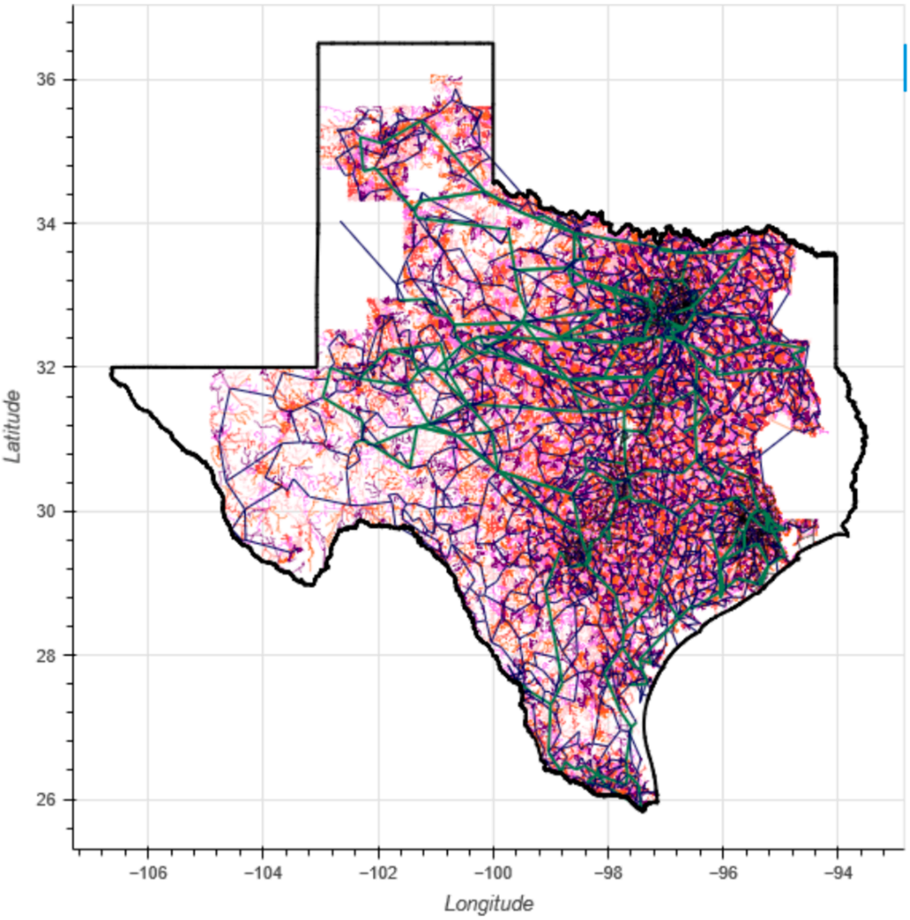


Fig. 10. The *syn*-texas-TDgrid-v02 synthetic T&D system of Texas. Transmission lines are shown in green (345 kV) and black (138/69 kV), while distribution lines are shown in orange, red, purple, yellow, and pink. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

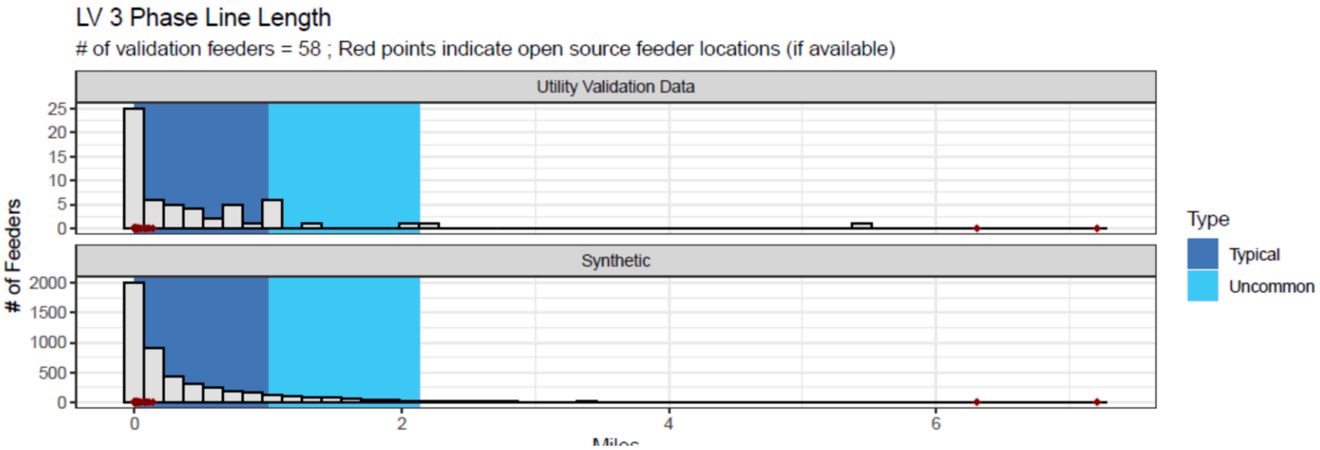
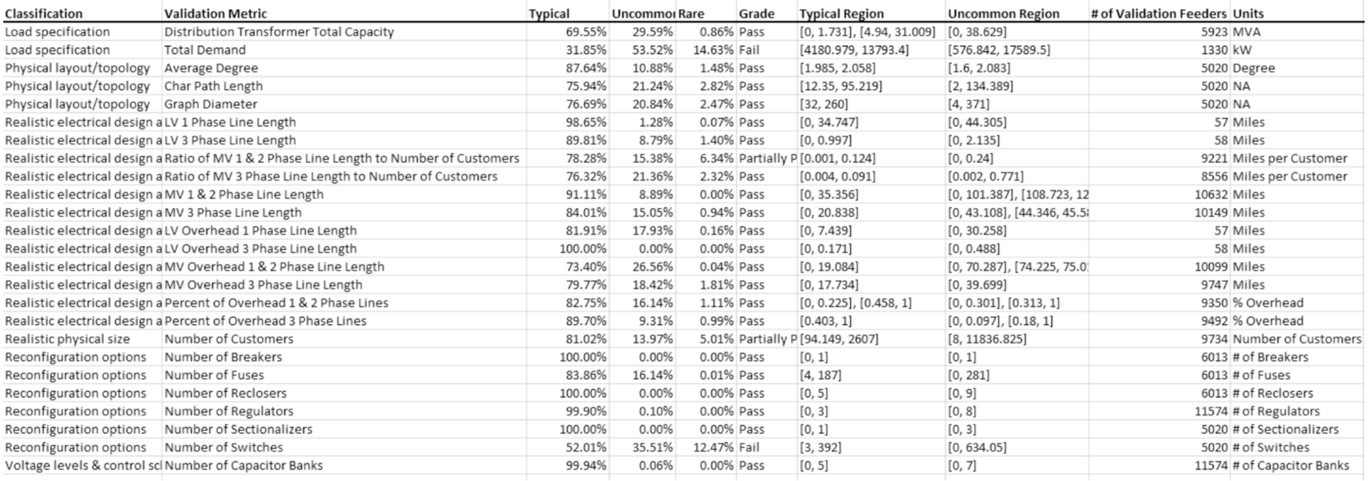


Fig. 11. Histogram of LV 3 phase line length.

distribution systems. A fourth issue is the multiple operating systems required by various simulation tools.

As a demonstration, this paper shows how many of these challenges can addressed using co-simulation, here illustrated using the HELICS framework with a portion of the overall dataset. This approach allows the distribution of the computational effort among multiple nodes, and enables cross-platform communication for T&D co-simulation [41]. In particular, [42] describes efforts to use this approach to co-simulate the smaller Austin area data set generated by the authors using the same approach described here [36]. In that effort, five computer nodes (32 cores each) are used to simulate the T&D system for Austin. The HELICS platform provides the capability to create a communication interface between our Transmission solver (PowerWorld) and Distribution solver (OpenDSS). For any loading conditions in which we want to solve the case, we start first with the loads of the distribution system and an estimation of the energy losses, and then we solve the transmission



Classification	Validation Metric	Typical	Uncommo	Rare	Grade	Typical Region	Uncommon Region	# of Validation Feeders	Units
Load specification	Distribution Transformer Total Capacity	69.55%	29.59%	0.86%	Pass	[0, 1.731], [4.94, 31.009]	[0, 38.629]	5923	MVA
Load specification	Total Demand	31.85%	53.52%	14.63%	Fail	[4180.979, 13793.4]	[576.842, 17589.5]	1330	kW
Physical layout/topology	Average Degree	87.64%	10.88%	1.48%	Pass	[1.985, 2.058]	[1.6, 2.083]	5020	Degree
Physical layout/topology	Char Path Length	75.94%	21.24%	2.82%	Pass	[12.35, 95.219]	[2, 134.389]	5020	NA
Physical layout/topology	Graph Diameter	76.69%	20.84%	2.47%	Pass	[32, 260]	[4, 371]	5020	NA
Realistic electrical design a	LV 1 Phase Line Length	98.65%	1.28%	0.07%	Pass	[0, 34.747]	[0, 44.305]	57	Miles
Realistic electrical design a	LV 3 Phase Line Length	89.81%	8.79%	1.40%	Pass	[0, 0.997]	[0, 2.135]	58	Miles
Realistic electrical design a	Ratio of MV 1 & 2 Phase Line Length to Number of Customers	78.28%	15.38%	6.34%	Partially P	[0.001, 0.124]	[0, 0.24]	9221	Miles per Customer
Realistic electrical design a	Ratio of MV 3 Phase Line Length to Number of Customers	76.32%	21.36%	2.32%	Pass	[0.004, 0.091]	[0.002, 0.771]	8556	Miles per Customer
Realistic electrical design a	MV 1 & 2 Phase Line Length	91.11%	8.89%	0.00%	Pass	[0, 35.356]	[0, 101.387], [108.723, 12	10632	Miles
Realistic electrical design a	MV 3 Phase Line Length	84.01%	15.05%	0.94%	Pass	[0, 20.838]	[0, 43.108], [44.346, 45.5	10149	Miles
Realistic electrical design a	LV Overhead 1 Phase Line Length	81.91%	17.93%	0.16%	Pass	[0, 7.439]	[0, 30.258]	57	Miles
Realistic electrical design a	LV Overhead 3 Phase Line Length	100.00%	0.00%	0.00%	Pass	[0, 0.171]	[0, 0.488]	58	Miles
Realistic electrical design a	MV Overhead 1 & 2 Phase Line Length	73.40%	26.56%	0.04%	Pass	[0, 19.084]	[0, 70.287], [74.225, 75.0	10099	Miles
Realistic electrical design a	MV Overhead 3 Phase Line Length	79.77%	18.42%	1.81%	Pass	[0, 17.734]	[0, 39.699]	9747	Miles
Realistic electrical design a	Percent of Overhead 1 & 2 Phase Lines	82.75%	16.14%	1.11%	Pass	[0, 0.225], [0.458, 1]	[0, 0.301], [0.313, 1]	9350	% Overhead
Realistic electrical design a	Percent of Overhead 3 Phase Lines	89.70%	9.31%	0.99%	Pass	[0.403, 1]	[0, 0.097], [0.18, 1]	9492	% Overhead
Realistic physical size	Number of Customers	81.02%	13.97%	5.01%	Partially P	[94.149, 2607]	[8, 11836.825]	9734	Number of Customers
Reconfiguration options	Number of Breakers	100.00%	0.00%	0.00%	Pass	[0, 1]	[0, 1]	6013	# of Breakers
Reconfiguration options	Number of Fuses	83.86%	16.14%	0.01%	Pass	[4, 187]	[0, 281]	6013	# of Fuses
Reconfiguration options	Number of Reclosers	100.00%	0.00%	0.00%	Pass	[0, 5]	[0, 9]	6013	# of Reclosers
Reconfiguration options	Number of Regulators	99.90%	0.10%	0.00%	Pass	[0, 3]	[0, 8]	11574	# of Regulators
Reconfiguration options	Number of Sectionalizers	100.00%	0.00%	0.00%	Pass	[0, 1]	[0, 3]	5020	# of Sectionalizers
Reconfiguration options	Number of Switches	52.01%	35.51%	12.47%	Fail	[3, 392]	[0, 634.05]	5020	# of Switches
Voltage levels & control sc	Number of Capacitor Banks	99.94%	0.06%	0.00%	Pass	[0, 5]	[0, 7]	11574	# of Capacitor Banks

Fig. 12. Summary of the distribution system validation metrics in Texas.

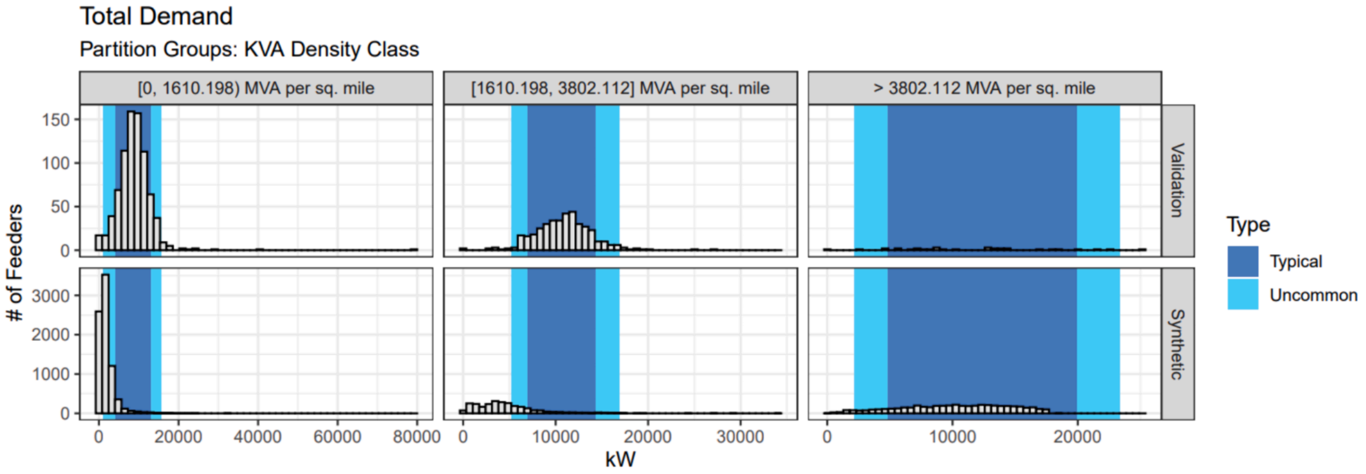


Fig. 13. Total demand partitioned by load density.

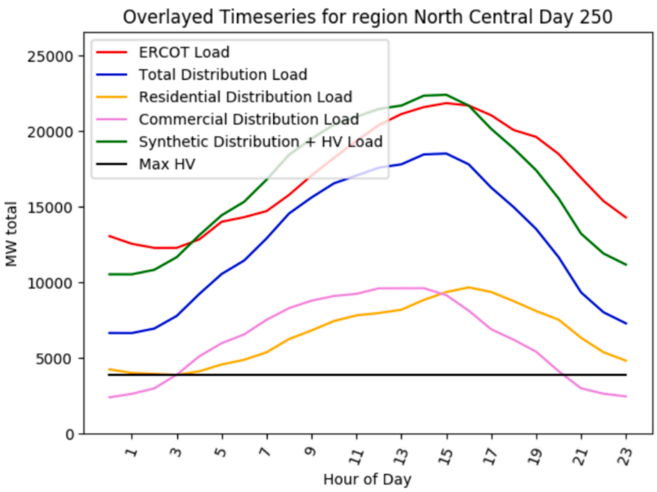


Fig. 14. Validation of load.

network. Next, in parallel, for each distribution feeder which is modeled (we can model part of the network or all of it), we pass the positive sequence voltage magnitude down to the feeder head in the model and run a distribution power flow. From here, we can calculate the effective total active and reactive power (P and Q) consumed by the distribution substation, which we translate upward to the transmission solver, and repeat. The solution converges when the power balance at the distribution substation matches that in the transmission network, within a tolerance at which point the co-simulation can advance to the next timestep. In addition to capturing the physical phenomena of power flow, this arrangement further enables combining control schemes, market operations, end use and DER simulations and even other energy systems into the same co-simulation environment. Such extensions, as well as expanding to include the entire T&D data set are left for future work. Fig. 15 and Fig. 16 depict the voltage response in the T&D system as the loads change over a 24-hour period. In Fig. 15, the T&D are simulated separately, with constant power loads on the transmission and constant voltage source at the distribution. In Fig. 16, an iterative simulation framework is used that converges upon a solution consistent with both models.

Since HELICS is a co-simulation framework, it brings together multiple existing tools and does not contain any native power flow implementation. Rather if a future researcher wanted to use HELICS for powerflow alone, they could use existing connections to OpenDSS, CYME, GridLAB-D, and others on the distribution side; PSS/E, PowerWorld and others on the transmission side; and/or introduce their own tools or algorithms. The use (or not) of HELICS does not impose any constraints on the choice of representations. It does not require or suppose a positive sequence representation for the transmission system. It

Table 3
Selected transmission system validation metrics in Texas.

Validation Metric	Typical Values	Synthetic Texas Transmission Case
Voltage magnitude	70–90 % in 0.98–1.04	91 %
	30+% in 1.00–1.02	28.1 %
Shunt capacitors and reactors	10–25 % of substations	8.8 %
	30-50 % of > 200 kV subs.	22.5 %
Contingency analysis for selected N-1 and N-2 contingencies	Power flow convergence, minimal overloads	8598 conts., converge, minimal violations
Substations with load	75–90 %	92.7 %
Generation Capacity/Load	1.2–1.6	1.21
Load per bus	Mean 6–18 MW	12.9 MW
Substations with generators	5–25 %	6.68 %
Generator MW maximum capacities	25–200 MW, 40+%	63.2 %
	200 + MW, 2–20 %	20.5 %
Transmission line per-km, per-unit reactance, within 10–90 % range from [40]	> 70 %	345 kV: 98.7 % 138 kV: 99.5 % 69 kV: 93.6 %
Transmission line X/R ratio, within 10–90 % range from [40]	> 70 %	345 kV: 100 % 138 kV: 99.9 % 69 kV: 99.4 %
Lines/Substations by kV level	1.1–1.4	345 kV: 1.34 138 kV: 1.2 69 kV: 1.2
Lines on the minimum spanning tree, by kV level	40–60 %	345 kV: 57 % 138 kV: 51 % 69 kV: 1.2 %
Lines on Delaunay triangulation	65–85 %	345 kV: 86 % 138 kV: 77 % 69 kV: 74 %
Total line length / length of minimum spanning tree	1.2–2.2	345 kV: 1.91 138 kV: 1.96 69 kV: 1.8

can and has been used to bring together 3-phase representations of the bulk transmission system.

This case study serves to illustrate, with a partial view of the system, that the provided data can be used to perform T&D co-simulation. The use of the HELICS platform allows leveraging standard commercial simulation tools (e.g. PowerWorld and OpenDSS), combining them to carry out the simulations and can be scaled to combine millions of separate simulations if desired in future work. In the literature, there are also dedicated solutions to obtain this kind of solution. For example, [43] applies a power flow algorithm to solve a 784 k node T&D system, while [44] shows a high efficiency algorithm, applied on a 2592 case study, that outperforms Newton-Raphson algorithms. The dataset provided in this paper has 46 million nodes and 679 interrelated cycles. The application of such methodologies to simulate the full T&D system of Texas is outside of the scope of this paper, instead the data set described is intended to provide a challenge for future research in these areas aiming at improving computation technologies.

6. Discussion and potential applications

The large-scale synthetic test system built in this paper aims to fill the existing gap in publicly available, realistic, full-scale, detailed T&D test systems. It aims to serve as a benchmark to improve power flow algorithms which are the core of many of the tools and analyses used by utilities to operate and plan transmission and distribution systems. There are numerous publications that have already been released which utilize both the separate distribution and transmission datasets that the authors have previously released such as [45,46,47,48]. We expect that many future studies will also be able to utilize the significantly larger combined T&D datasets that we present this paper.

Capturing T&D interactions can provide a foundation for a wide range of analyses. For example, technical aspects of integrated T&D power systems analysis, ranging from power flow analyses to optimal

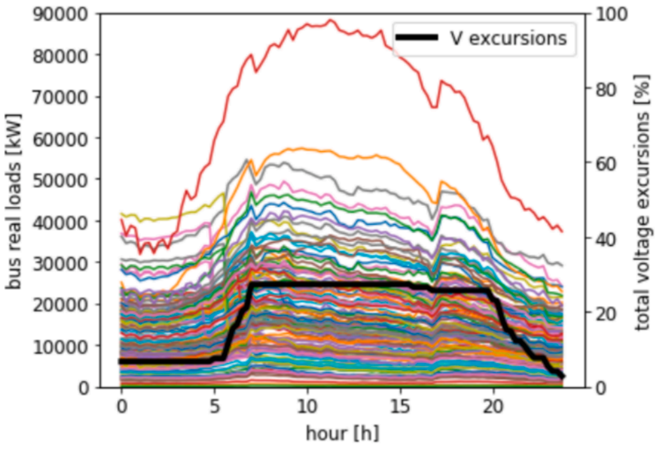


Fig. 15. Voltage response with fixed supply voltage at substations.

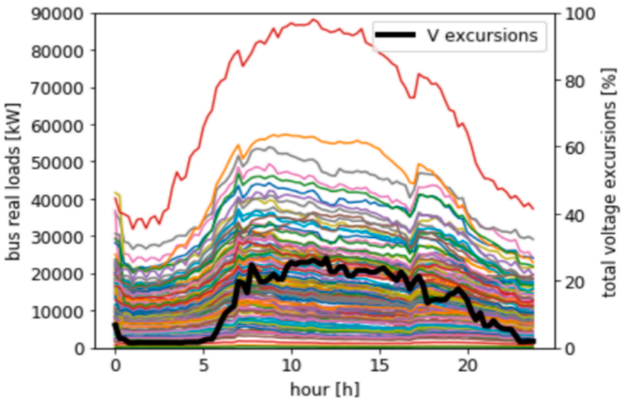


Fig. 16. Voltage response with supply voltage from closely tied transmission co-simulation.

power flows, reconfiguration, reliability analysis, voltage control, energy losses assessment, and more. It can also be used to explore future grid scenarios such as the interactions of distributed generation, storage, electric vehicles, demand response, energy efficiency measures, and more with T&D power operations. These kinds of analyses could help to identify how DERs can contribute to integrated energy and climate objectives, while making the most of flexibility and opportunities to support T&D technical constraints. The dataset also enables analyzing economic interactions, such as price-responsive demand, bid-in demand resources, or local distribution scale markets with wholesale electric markets, including the range of options for responding to FERC 2222 opportunities. Yet another class of applications includes testing, debugging, and scaling advanced algorithms that promise to help manage these and other T&D interactions in future grids.

This dataset can also provide fully open grid infrastructure datasets that can be used for detailed analyses of TSO and DSO cooperation arrangements. To do so, various market or other arrangements can be explored with the synthetic Texas T&D dataset as a test grid. Such analysis could include capturing potential challenges if DERs provide energy and grid services to the TSO without considering local distribution conditions, or vice-versa where the local customer or distribution-driven DER operations may run counter to system-wide needs.

7. Conclusions

The methodology of this paper creates very large-scale combined T&D systems by merging the most up-to-date techniques for creating synthetic distribution feeder networks with the latest methods for building synthetic, meshed bulk-power transmission networks. The unique very-large-scale synthetic T&D system built, *syn*-texas-TDgrid-v02, covers most of Texas, and is made available to the scientific community. Validation demonstrates that this synthetic test system realistically represents characteristics of actual networks.

The methodology is shown to be effective for generating combined T&D datasets over very large areas, in high resolution, and with extensive local details. Despite the different nature of the networks, it is shown that the methodologies to build transmission and distribution systems share common aspects regarding the design of multiple voltage levels. The use of a bottom-up approach allows each planning algorithm to make decisions using the aggregated results of pre-determined subsystems, ensuring a strong match and high data integrity across the entire system. The resulting synthetic test system (realistic but not real) represents an enormous leap in public test systems, providing very large-scale interconnected T&D system data that can be used for a wide range of applications. These include technical engineering analysis, forward-looking scenario analyses, and economic market interactions to test large-scale control algorithms.

Given the limited T&D interactions between utilities and system operators, we hope that this dataset will open new frontiers and encourage the scientific community and industry to think differently about how an interconnected system may be operated through concrete examples and analyses.

Moreover, the sheer size of this dataset is intended to provide a challenge for the scientific community, both presently and in the coming years. As a first step, it has been shown how to assemble the simulations of the transmission and distribution systems by leveraging HELICS. This lays the foundation for interesting future research, such as demonstrating the coordination of T&D using flexibility or distributed energy resources, and can help foster the improvement of algorithms, analysis methods, and corresponding tools for years to come.

CRediT authorship contribution statement

Carlos Mateo: Writing – review & editing, Writing – original draft, Visualization, Software, Methodology, Investigation, Conceptualization. **Fernando Postigo:** Investigation, Formal analysis. **Tarek Elgindy:** Writing – review & editing, Validation, Project administration, Methodology, Conceptualization. **Adam B. Birchfield:** Writing – review & editing, Visualization, Software, Methodology, Investigation. **Pablo Dueñas:** Writing – review & editing, Methodology, Data curation. **Bryan Palmintier:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization. **Nadia Panossian:** Formal analysis. **Tomás Gómez:** Writing – review & editing, Supervision, Conceptualization. **Fernando de Cuadra:** Investigation, Formal analysis. **Thomas J. Overbye:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization. **Farnaz Safdarian:** Writing – review & editing, Supervision, Project administration, Conceptualization. **Diana Wallison:** Writing – review & editing, Visualization, Validation, Software.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: [Carlos Mateo (and all authors) report financial support was provided by the Advanced Research Projects Agency – Energy (ARPA-E), under Work Authorization (15/CJ000/07/03) and grant agreement ARPA-E Award No. DE-AR0000713 (“Cooperative Agreement”).].

Data availability

The syn-texas-TDgrid T&D system is available to the research community from the Open Energy Data Initiative (OEDI) repository [49] and from [50].

Appendix

The *syn*-texas-TDgrid T&D system is available to the research community from the Open Energy Data Initiative (OEDI) repository [49] and from [50].

References

[1] Givisiez AG, Petrou K, Ochoa LF. A review on TSO-DSO coordination models and solution techniques. Electr Pow Syst Res 2020;189. https://doi.org/10.1016/j.epsr.2020.106659.

[2] Gerard H, Rivero Puente EI, Six D. “Coordination between transmission and distribution system operators in the electricity sector: a conceptual framework”. Util Policy Feb 2018;50:40–8. https://doi.org/10.1016/j.jup.2017.09.011.

[3] Marcos FEP, et al. A review of power distribution test feeders in the United States and the need for synthetic representative networks. Energies 2017;10:11. https://doi.org/10.3390/en10111896.

[4] Kersting WH. Radial distribution test feeders. IEEE Trans Power Syst 1991;6(3): 975–85. https://doi.org/10.1109/59.119237.

[5] Arritt RF, Dugan RC, “The IEEE 8500-node test feeder,” in *Transmission and Distribution Conference and Exposition, 2010 IEEE PES*, 2010, p. 1.

[6] Barrows C, et al. The IEEE Reliability Test System: A Proposed 2019 Update. IEEE Trans Power Syst Jan. 2020;35(1):119–27. https://doi.org/10.1109/TPWRS.2019.2925557.

[7] González-Morán C, Arboleya P, Mojumdar RR, Mohamed B. 4-node test feeder with step voltage regulators. Int J Electr Power Energy Syst Jan. 2018;94:245–55. https://doi.org/10.1016/j.ijepes.2017.06.027.

[8] Schneider KP, et al. Analytic considerations and design basis for the IEEE distribution test feeders. IEEE Trans Power Syst 2017;33(3):3181–8. https://doi.org/10.1109/TPWRS.2017.2760011.

[9] Koirala A, Suárez-Ramón L, Mohamed B, Arboleya P. Non-synthetic European low voltage test system. Int J Electr Power Energy Syst Jun. 2020;118:105712. https://doi.org/10.1016/j.ijepes.2019.105712.

[10] Taye T, Mohamed B, Suarez-Ramon L, Arboleyá P. A set of Non-Synthetic test systems of European LV Rural, LV urban and hybrid MV/LV industrial distribution networks. Int J Electr Power Energy Syst Jul. 2024;158:109941. https://doi.org/10.1016/j.ijepes.2024.109941.

[11] ARPA-E, “GRID DATA call,” https://arpa-e.energy.gov/technologies/programs/grid-data.

[12] Palmintier B, et al. Experiences developing large-scale synthetic U.S.-style distribution test systems. Electr Pow Syst Res 2021;190. https://doi.org/10.1016/j.epsr.2020.106665.

[13] Mateo C, et al. Building large-scale U.S. synthetic electric distribution system models. IEEE Trans Smart Grid 2020;11(6):5301–13. https://doi.org/10.1109/TSG.2020.3001495.

[14] Krishnan V, et al. Validation of synthetic U.S. power distribution system datasets. IEEE Trans Smart Grid 2020. https://doi.org/10.1109/TSG.2020.2981077.

[15] Taylor S, Rangarajan A, Rhodes N, Snodgrass J, Lesieutre BC, Roald LA. California test system (CATS): a geographically accurate test system based on the california grid. IEEE Trans Energy Markets, Policy Regul Mar. 2024;2(1):107–18. https://doi.org/10.1109/TEMPR.2023.3338568.

[16] Mohammadi MH, Saleh K. Synthetic benchmarks for power systems. IEEE Access 2021;9:162706–30. https://doi.org/10.1109/ACCESS.2021.3124477.

[17] Pilatte N, Aristidou P, Hug G. TDNetGen: an open-source, parametrizable, large-scale, transmission, and distribution test system. IEEE Syst J Mar. 2019;13(1):729–37. https://doi.org/10.1109/JSYST.2017.2772914.

[18] Khodr HM, Vale Z, Ramos C, Faria P. Optimization techniques for power distribution planning with uncertainties: a comparative study. In: 2009 IEEE Power & Energy Society General Meeting. IEEE; 2009. https://doi.org/10.1109/PES.2009.5275569.

[19] Birchfield AB, Xu T, Gegner KM, Shetye KS, Overbye TJ. Grid structural characteristics as validation criteria for synthetic networks. IEEE Trans Power Syst 2017;32(4):3258–65. https://doi.org/10.1109/TPWRS.2016.2616385.

[20] Birchfield AB, Xu T, Overbye TJ. Power flow convergence and reactive power planning in the creation of large synthetic grids. IEEE Trans Power Syst 2018;33(6):6667–74. https://doi.org/10.1109/TPWRS.2018.2813525.

[21] Mateo C, et al. European representative electricity distribution networks. Int J Electr Power Energy Syst 2018;99:273–80. https://doi.org/10.1016/j.ijepes.2018.01.027.

[22] Siemens, "PSS/E documentation version 32.0.0.".

[23] DIgSILENT, "DIgSILENT Powerfactory Technical Reference Documentation," *Technical Reference Documentation*. pp. 1–31, 2015.

[24] Vahidinasab V, et al. Overview of Electric Energy Distribution Networks Expansion Planning. IEEE Access 2020;8:34750–69. https://doi.org/10.1109/ACCESS.2020.2973455.

[25] Georgilakis PS, Hatziargyriou ND. A review of power distribution planning in the modern power systems era: Models, methods and future research. Electr Pow Syst Res 2015;121:89–100. https://doi.org/10.1016/j.epsr.2014.12.010.

[26] Schweitzer E, Scaglione A, Monti A, Pagani GA. Automated generation algorithm for synthetic medium voltage radial distribution systems. IEEE J Emerg Sel Top Circuits Syst 2017;7(2):271–84. https://doi.org/10.1109/JETCAS.2017.2682934.

[27] Yan R, Yuan Y, Wang Z, Geng G, Jiang Q. Active distribution system synthesis via unbalanced graph generative adversarial network. IEEE Trans Power Syst Sep. 2023;38(5):4293–307. https://doi.org/10.1109/TPWRS.2022.3212029.

[28] Pereira Junior BR, Cossi AM, Contreras J, Mantovani JRS. Multiobjective multistage distribution system planning using tabu search. IET Gener Transm Distrib 2014;8(1):35–45. https://doi.org/10.1049/iet-gtd.2013.0115.

[29] González-Sotres L, Mateo C, Sánchez-Miralles Á, Miró MA. Large-scale MV / LV transformer substation planning considering network costs and flexible area decomposition. IEEE Transactions on Power Delivery 2013;28(4):2245–53. https://doi.org/10.1109/TPWRD.2013.2258944.

[30] Birchfield AB, Overbye TJ, "Planning sensitivities for building contingency robustness and graph properties into large synthetic grids," *Proceedings of the Annual Hawaii International Conference on System Sciences*, vol. 2020, pp. 3167–3175, 2020, https://doi.org/10.24251/hicss.2020.386.

[31] Birchfield AB, Xu T, Shetye KS, Overbye TJ, "Building synthetic power transmission networks of many voltage levels, spanning multiple areas," *Proceedings of the Annual Hawaii International Conference on System Sciences*, vol. 2018, pp. 2766–2774, 2018, https://doi.org/10.24251/hicss.2018.349.

[32] Ding T, Liu S, Yuan W, Bie Z, Zeng B. A two-stage robust reactive power optimization considering uncertain wind power integration in active distribution networks. IEEE Trans Sustain Energy 2016;7(1):301–11. https://doi.org/10.1109/TSTE.2015.2494587.

[33] Trpovski A, Hamacher T, "Ring Distribution System Expansion Planning using Scenario Based Mixed Integer Programming," *Proceedings of the IEEE Power Engineering Society Transmission and Distribution Conference*, vol. 2020, 2020, https://doi.org/10.1109/TD39804.2020.9299971.

[34] Mateo Domingo C, Gómez T, Sánchez-Miralles Á, Peco JP, Candela A. A reference network model for large-scale distribution planning with automatic street map generation. IEEE Trans Power Syst 2011;26(1):190–7. https://doi.org/10.1109/TPWRS.2010.2052077.

[35] Pieltain L, Gómez T, Cossent R, Mateo C, Frías P. Assessment of the impact of plug-in electric vehicles on distribution networks. IEEE Trans Power Syst 2011;26(1):206–13. https://doi.org/10.1109/TPWRS.2010.2049133.

[36] Li H, et al. Building highly detailed synthetic electric grid data sets for combined transmission and distribution systems. IEEE Open Access Journal of Power and Energy 2020;7:478–88. https://doi.org/10.1109/oajpe.2020.3029278.

[37] Salehi J, Haghifam MR. Long term distribution network planning considering urbanity uncertainties. Int J Electr Power Energy Syst 2012;42(1):321–33. https://doi.org/10.1016/j.ijepes.2012.04.005.

[38] Postigo F, et al. Phase selection algorithms to minimize cost and imbalance in U.S. synthetic distribution networks. Int J Electr Power Energy Syst 2020;120. https://doi.org/10.1016/j.ijepes.2020.106042.

[39] Schneider K, Phanivong P, Lacroix J.-S., "IEEE 342-node low voltage networked test system," in *2014 IEEE PES General Meeting | Conference & Exposition*, IEEE, 2014, pp. 1–5. https://doi.org/10.1109/PESGM.2014.6939794.

[40] Birchfield AB, et al. A metric-based validation process to assess the realism of synthetic power grids. Energies (Basel) 2017;10(8):pp. https://doi.org/10.3390/en10081233.

[41] Bryan Palmintier; Dheepak Krishnamurthy; Philip Top; Steve Smith; Jeff Daily; Jason Fuller, "Design of the HELICS High-Performance Transmission-Distribution-Communication-Market Co-Simulation Framework," in *2017 Workshop on Modeling and Simulation of Cyber-Physical Energy Systems (MSCPES)*, 2017.

[42] Panossian N, Elgindy T, Palmintier B, Wallison D, "Synthetic, Realistic Transmission and Distribution Co-Simulation for Voltage Control Benchmarking," in *2021 IEEE Texas Power and Energy Conference, TPEC 2021*, 2021. https://doi.org/10.1109/TPEC51183.2021.9384935.

[43] Bhatti BA, Broadwater R, Dilek M. Analyzing impact of distributed pv generation on integrated transmission & distribution system voltage stability—a graph trace analysis based approach. Energies (Basel) Sep. 2020;13(17):4526. https://doi.org/10.3390/en13174526.

[44] Tbaileh A, Bhatti BA, Broadwater R, Dilek M, Beattie C. Robust Matrix Free Power Flow Algorithm for Solving T&D Systems. In: 2019 IEEE Power & Energy Society General Meeting (PESGM). IEEE; Aug. 2019. p. 1–5. https://doi.org/10.1109/PESGM40551.2019.8973953.

[45] Wang W, et al. PV Hosting Capacity Estimation: Experiences with Scalable Framework. In: 2022 IEEE 49th Photovoltaics Specialists Conference (PVSC). IEEE; Jun. 2022. p. 0967–71. https://doi.org/10.1109/PVSC48317.2022.9938653.

[46] Kroposki B, et al. Autonomous energy grids: controlling the future grid with large amounts of distributed energy resources. IEEE Power Energ Mag Nov. 2020;18(6):37–46. https://doi.org/10.1109/MPE.2020.3014540.

[47] Birchfield AB, Gegner KM, Xu T, Shetye KS, Overbye TJ. Statistical considerations in the creation of realistic synthetic power grids for geomagnetic disturbance studies. IEEE Trans Power Syst 2016:1. https://doi.org/10.1109/TPWRS.2016.2586460.

[48] Zhang G, Zhong H, Tan Z, Cheng T, Xia Q, Kang C. Texas electric power crisis 2021 warns of a new mode of blackout. CSEE J Power Energy Syst 2022. https://doi.org/10.17775/CSEEJPES.2021.07720.

[49] Open Energy Data Initiative, "Syn-texas-TDgrid T&D system," https://data.openei.org/s3_viewer?bucket=oedi-data-lake&prefix=SMART-DS/v0.9/.

[50] "Texas Combined Transmission-Distribution Test Case," https://electricgrids.engr.tamu.edu/texas7k-td/.