

Received September 27, 2021, accepted October 22, 2021, date of publication December 10, 2021, date of current version December 16, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3124477

Synthetic Benchmarks for Power Systems

M. HOSSAIN MOHAMMADI^{ID}, (Member, IEEE), AND KHALED SALEH^{ID}, (Member, IEEE)

CanmetENERGY Research Centre, Natural Resources Canada, Varennes, QC J3X 1P7, Canada

Corresponding author: M. Hossain Mohammadi (hossain.mohammadi@nrcan-rncan.gc.ca)

This work was supported by the Program of Energy Research and Development at Natural Resources Canada.

ABSTRACT Power system benchmarks are transmission and distribution networks used to evaluate novel control algorithms and simulate grid evolution scenarios. These benchmarks range in size, system characteristics, and use cases. Although active working groups have created and published many benchmarks, these networks are not all representative of a given region and may not consider certain aspects such as increased penetration levels of distributed energy resources. To address these issues, synthetic benchmark networks and methodologies for generating them have been developed by various research groups. This paper provides a comprehensive survey of procedures commonly used to generate synthetic networks and a detailed account of the various metrics used to define and validate benchmarks. Existing models are categorized into different approaches, including expert design, anonymized clustering, statistical sampling, and heuristic algorithms. Deep graph generation based techniques are also presented and recommended for the network generation problem. A comparative summary is provided to highlight the different existing works in this area and reveal research gaps, along with a list of published datasets and their characteristics.

INDEX TERMS Graph theory, machine learning, modeling, network topology, neural networks, power distribution, power grids, power system modeling, power transmission, statistics.

NOMENCLATURE

a_t	Action taken by an agent at time t .	$D_{KS}(\cdot)$	Kolmogorov–Smirnov divergence.
\mathbf{b}	Betweenness centrality vector.	D_{ij}	Shortest distance to traverse from nodes i to j .
\mathbf{c}	Clustering coefficient vector.	D_{\max}	Graph diameter.
$d(\cdot)$	Discriminator function.	\mathcal{E}	Set of graph edges.
e_i	Number of edges in local cluster of node i .	$\mathbb{E}[\cdot]$	Expected value of a random variable.
$g(\cdot)$	Generator function.	\mathcal{G}	Graph structure.
\mathbf{k}	Node degree vector.	\mathbf{K}	Node degree matrix.
\mathbf{l}	Characteristic path length vector.	\mathbf{L}	Laplacian or lower-triangular matrix of \mathbf{A} .
m	Number of features per node.	$\mathcal{N}(\cdot)$	Gaussian distribution function.
p_{fake}	Probability output value of a discriminator.	$P(\cdot), Q(\cdot)$	Cumulative density function.
$p(\cdot), q(\cdot)$	Probability density function.	\mathbb{R}	Real coordinate space.
r_t	Reward awarded to an agent at time t .	S_{ij}	Expected savings by looping nodes i and j .
\mathbf{t}	Bus type assignment vector.	\mathcal{V}	Set of graph vertices.
$w(\cdot)$	Bus type entropy distribution.	\mathbf{X}	Feature matrix of a graph.
\mathbf{z}	Latent variable vector.	$\delta_{ij}(v)$	Number of shortest paths from i to j crossing v .
\mathbf{z}_{in}	Input properties of a generative model.	θ, ϕ	Neural network parameters.
\mathcal{A}	Action space of an agent.	$\lambda_2(\mathbf{L})$	Second smallest eigenvalue of Laplacian.
A_{ij}	Adjacency matrix entry of nodes i and j .	$\boldsymbol{\mu}$	Vector of mean values.
\mathbf{A}	Adjacency matrix of a graph.	$\pi(\cdot)$	Policy function of an agent.
$D_H(\cdot)$	Hellinger distance.	ρ	Degree assortativity.
$D_{KL}(\cdot)$	Kullback–Leibler divergence.	$\boldsymbol{\sigma}^2$	Vector of variance values.
		(\cdot)	Tuple or sequence of variables.
		$ \cdot $	Size of a given set or vector.
		$\langle \cdot \rangle$	Average value of a given vector.

The associate editor coordinating the review of this manuscript and approving it for publication was Ravindra Singh.

I. INTRODUCTION

The electric grid must undergo significant changes and modernization to improve reliability and facilitate decarbonisation in the face of climate change. Renewable and distributed energy resources (DERs) will help achieve these targets and decentralize the power sector while satisfying rising energy demands [1], [2]. Load flexibility from all sectors (e.g. residential, industrial, commercial) will contribute to grid services, thereby reducing energy loss and contributing to peak shaving [3], [4].

Despite the benefits, adopting these technologies introduce challenges due to the complex and interconnected structure of future grids. A grid's performance under new planning and operational paradigms must be well understood and predicted to achieve the targets for clean energy and grid resilience. For example, widespread deployment of DERs requires developing novel control and protection schemes to enable smooth operation and avoid disruptions [5]. To address such challenges, researchers and industry experts rely on using existing benchmark networks to investigate transmission or distribution systems under different operating conditions. These benchmarks can be modeled in various simulation environments depending on the analysis type and can facilitate validating developed algorithms and techniques.

Active working groups in IEEE and CIGRÉ have created and published openly available test cases for both transmission [6]–[8] and distribution [9]–[11] systems. These benchmarks are used widely for different power system applications, e.g. testing new power flow algorithms and their scalability, evaluating protection schemes, implementing grid control, and initiating planning studies. The network sizes range from a few buses or nodes to a couple of thousand and consist of standard power system components (e.g. transformers, switches, overhead lines, underground cables). Practical aspects such as topological and system differences in modeling North American and European grids were considered by developing alternative versions. Despite the extensive use and applications of these standard benchmarks in power system research, most of them omit important considerations such as modern technologies (e.g. power electronic converters) and variable energy sources, thereby questioning their suitability for grid transformation studies [12], [13].

Attempts to modify existing benchmarks are often not based on real-world systems, among other issues explained below. First, not all the available test cases are *representative* of a given location. Regional differences affect the design of power systems and their benchmarks. These differences include phase unbalance in North American grids compared to Europe, the evolving role of grid infrastructures, rural vs. urban districts, system frequency and voltage levels, i.e. low (LV), medium (MV), high (HV), and extra high (EHV), network topology and feeder structures, transformer sizes and the number of connected customers, line types, and grounding, among others [11]. Additionally, the different

service types and clients (e.g. residential, industrial, commercial, agricultural, critical loads) must be identified to better consider flexibility potential and set appropriate service levels. The presence of natural or manmade boundaries (e.g. reserves, forests, rivers, lakes, buildings, neighborhoods) can also impose certain restrictions on a planning study. Hence, grid evolution must be evaluated using more *representative* benchmarks for each region that consider such factors.

Second, utilities are generally not willing to share the true network information due to *confidentiality* and security concerns. Revealing the underlying topology and device locations can have severe risks and consequences; publishing detailed data can violate consumer privacy (e.g. personal information, energy consumption) or cause grid vulnerability. It is more common to share aggregated information for research and development since it does not reveal the actual network. This lack of data access can also complicate validating benchmarks for their practicality and operational closeness to real grids. When creating a benchmark network, the tradeoff between *representativeness* and *confidentiality* must be considered.

Third, classical benchmarks are simplistic in terms of their modeled components and scenarios. They were originally developed for various applications (e.g. power flow convergence) and do not necessarily transfer well for grid evolution studies. Benchmark applications are many, such that each one needs different simulation environments and assumptions. With the increasing integration of DERs and the participation of consumers in demand response programs, these aspects need careful consideration during grid analysis and planning. The lack of modern technologies such as DERs and communication systems incentivize developing new benchmarks instead of adapting standard ones [12]. Newer variations of benchmarks must incorporate these *modeling complexities*.

Due to these issues, there has been a growing research interest to develop and create test cases for power systems, known as “*synthetic benchmark networks*”. Although some differences exist to protect sensitive information of real grids, overall operation and regional representation are preserved as much as possible. Several groups have developed methods to generate synthetic benchmarks for various applications.

This paper provides a comprehensive survey and discussions on procedures for creating synthetic power systems (known as generative models) and categorizes them into different approaches. Each generative model is compared in terms of its requirements, assumptions, selection criteria, and results. Various metrics for characterizing and validating representative networks are also described and clustered into distinct groups. While recognizing that there are benchmarks for microgrids designed by experts [14]–[16] (for purposes such as testing control and protection algorithms, etc.), real-world microgrids are not yet mature, nor widespread; it would be premature to propose or survey methods for creating *representative* benchmarks.

The paper is structured as follows: Section II introduces the general procedure for creating synthetic benchmarks. Sections III–VII describe individual steps of this generative process by discussing details of published research. The various generation methods are compared in Section VIII for different aspects, before tabulating available datasets. Novel artificial intelligence-based techniques are also suggested and recommended for the generation problem.

Note that the terms “graph” and “network” are used interchangeably in this paper, and some words must be interpreted in context, e.g. “distribution” as network or statistics and “generation” for the creation of synthetic networks or output of a power generation unit.

II. GENERATION OF SYNTHETIC NETWORKS

A power system benchmark consists of multiple modeling layers as shown in Figure 1. At the lowest level, the topological layer consists of basic network or graph structures such as nodes and edges to represent the connectivity. Next, the geographical layer maps each node to a corresponding location that may or may not be real, depending on the required level of anonymization. The electrical layer then assigns sufficient information to the underlying network to run power system simulations such as optimal power flow. Other high-level aspects can be appended on a benchmark depending on the modeling complexity required, e.g. time series load and DER profiles.

Figure 2 illustrates the general procedure for generating a synthetic benchmark network. Step 1 of problem specification defines a network’s requirements, e.g. application or use case, regional constraints, and available or future assets. Failure to specify them clearly may lead to infeasible or unrealistic results in later stages that can delay the overall process. Next, data representative of a given network is collected in Step 2, which comprises the graph topology (e.g. nodes, edges), system characteristics (e.g. transformers, lines, substations), and/or statistical distributions of actual grids (e.g. number of buses per substation, total line length per feeder). This data must be pre-processed in Step 3 for ease of use and fed into a generative model. For a given application, the data may be anonymized to hide sensitive details while preserving the main characteristics. Step 4 consists of model development that can be considered as a separate standalone process; the adopted approach in grid modeling impacts the generated network and its quality. Once a suitable model is developed, synthetic networks are generated in Step 5 and then validated in Step 6. The metrics used are related to the initial specifications set in Step 1.

While a linear process can be followed for Figure 2, a systematic approach is more common for navigating between the different steps and incrementally adjusting the inputs and assumptions based on the outcomes of each step. Considering an unsatisfactory network validation in the final step, either the specifications and assumptions in Step 1 can be adjusted or more representative data can be collected in Step 2.

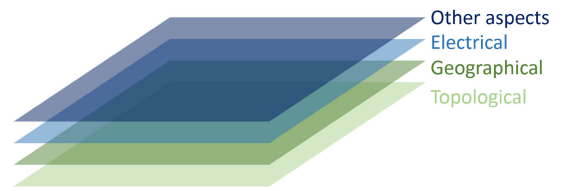


FIGURE 1. Multiple modeling layers in a power system benchmark.

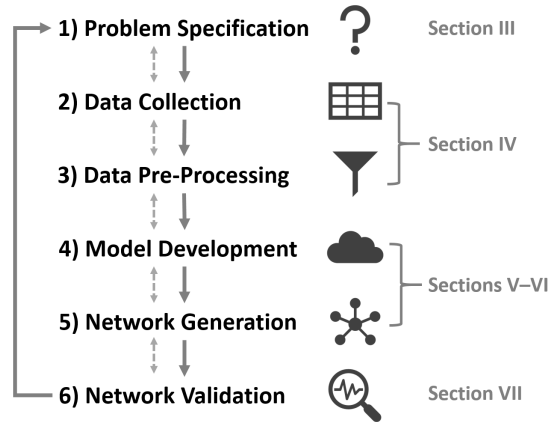


FIGURE 2. General procedure for generating synthetic networks.

III. PROBLEM SPECIFICATION

A design or generative process starts by defining the problem specifications. In creating a synthetic network for power systems, the benchmark must be *representative* of an actual grid in terms of its topology, system, and operation, and be capable of incorporating *modeling complexities* for grid evolution studies toward high penetration of renewable energy and DER integration. These high-level requirements depend on the problem specifications described below.

Initially, the application or use case of a synthetic benchmark must be determined to accurately model system components in a simulation environment. Power system applications can be divided into different categories: power quality, operation, protection, control, planning, and design. Each falls under transient or steady-state analyses, where the latter considers models with slow dynamics and timescales longer than seconds. Choosing a specific application affects subsequent steps of the generation process and the type of benchmark created as discussed in this section.

Different selection criteria must be considered when creating or using benchmarks such as differences between transmission and distribution systems and regional variations. According to CIGRÉ Task Force C6.04.02 [11], typical differences of a distribution grid in North America are its service, phase unbalance, voltage level, and system frequency when compared to a European one, as shown in Table 1. Considering these regional differences affect a benchmark’s type, structure and operation, it is challenging to evaluate a network based on these high-level considerations. Generally, each research work that focuses on creating synthetic

TABLE 1. Regional differences in distribution networks [11].

Properties	North American	European
Original service	Rural farms	Dense urban loads
Structure	Mostly radial	Mostly ring
Symmetry	Unbalanced	Balanced
Frequency	60 Hz	50 Hz
Typical MV feeder	12.47 or 24 kV 3-phase, 4-wire	11 or 20 kV 3-phase, Δ
Typical MV/LV transformer	15–75 kVA (1-phase), <8 customers	Urban 300–500 kVA Rural 50–200 kVA
Typical LV feeder	Mostly 1-phase, 2-wire 120/240 V (1-phase)	Mostly 3-phase, 4-wire 400 V (3-phase)

benchmarks prioritizes different criteria and typically uses a subset of them to validate generated networks in the final stage.

These criteria can be categorized into three main classes of characteristics for power systems: (i) topological, (ii) system, and (iii) operational. The individual criteria or metrics used in previous works are described in the following subsections and summarized in Table 2 under their referenced use: ‘Input’ represents the information required to generate benchmarks, ‘Constraint’ denotes conditions to check during the generation step, ‘Output’ corresponds to reported properties upon analyzing power systems (useful in identifying patterns), and ‘Validation’ relates to the performance comparison between synthetic and actual grids.

A. TOPOLOGICAL CHARACTERISTICS

Topological characteristics apply to graph-based problems in most disciplines through well-known properties. Within power systems, they can be used to distinguish transmission or distribution systems located in different regions. For example, radial networks, common in rural areas, have different topologies than ring or meshed networks, which are typically observed in densely populated urban districts for better resilience. Central hubs connected to many buses or nodes, such as HV-MV substations or pad-mounted transformers with high kVA ratings, can be identified with these characteristics. As seen in Table 2, topological metrics are mostly used to validate networks rather than for model development. Several references focused on analyzing and reporting these properties are shown in the ‘Output’ column.

Before discussing each one, fundamental graph theory must be defined. An undirected graph, \mathcal{G} , comprises a set of vertices, \mathcal{V} , a set of edges, \mathcal{E} , and a feature matrix, \mathbf{X} , as in (1). It is convenient to use an adjacency matrix, \mathbf{A} , such that the presence of edges is described by node indexing. Here, $|\mathcal{V}|$ is the total number of nodes and m is the number of features per node. Some graphs can be weighted, so any real value between 0 and 1 can represent an edge’s strength [70].

$$\begin{aligned} \mathcal{G} &= (\mathcal{V}, \mathcal{E}, \mathbf{X}) = (\mathbf{A}, \mathbf{X}) \\ \mathbf{A} &\in [0, 1]^{|\mathcal{V}| \times |\mathcal{V}|} \\ \mathbf{X} &\in \mathbb{R}^{|\mathcal{V}| \times m} \end{aligned} \quad (1)$$

1) NODE DEGREE

The node degree, k_i , defined in (2) uses the adjacency matrix to sum the number of edges connected to node i . The node degree can be described by a statistical distribution, its average value $\langle \mathbf{k} \rangle$ related to the numbers of nodes and edges, or the maximum degree [30], [31].

$$\begin{aligned} k_i &= \sum_{j=1}^{|\mathcal{V}|} A_{ij} \\ \langle \mathbf{k} \rangle &= \frac{1}{|\mathcal{V}|} \sum_{i=1}^{|\mathcal{V}|} k_i = \frac{2|\mathcal{E}|}{|\mathcal{V}|} \end{aligned} \quad (2)$$

In [21]–[24], [28], and [32], the average degree of transmission systems was reported to be 2–3 regardless of the region, grid size and voltage level under study. The range of $\langle \mathbf{k} \rangle$ is found to be consistently lower, i.e. 1.4–2, for MV distribution systems [17], [18], [30], [36]. The maximum degree corresponding to the most connected node of MV distribution networks was 4–10 [30], significantly lower than that of other large-scale networks. Power grids are structurally different from internet and social networks due to the lack of high-degree central hubs.

Instead of relying on single-point values such as an average, statistical distributions capture relative occurrences of a metric’s values based on sample data and help modeling networks in detail. For example, it is common to use empirical distributions [30], [34] or fit statistical models to the nodal degree, including the exponential [20], [25], [31], [32], Poisson [27], or mixture models [19], [26].

2) CLUSTERING COEFFICIENT

Another metric is the clustering coefficient, c_i , of node i defined in (3). Here, e_i is the number of edges in the local cluster of node i , and c_i is normalized with respect to the total number of possible edges in that neighborhood. A high value indicates the tendency of nodes to group together in dense clusters or communities.

Previous works [20], [22], [32] calculate the average coefficient over all nodes, $\langle c \rangle$, to ensure realism of generated networks. Transmission systems are reported to have a high value uncorrelated to the network size, generally 1–20% [20]–[22], [31], [35]. This range is comparable to standard graphs, such as a small-world or regular lattice [23], [24].

$$\begin{aligned} c_i &= \frac{e_i}{k_i(k_i - 1)/2} \\ \langle c \rangle &= \frac{1}{|\mathcal{V}|} \sum_{i=1}^{|\mathcal{V}|} c_i \end{aligned} \quad (3)$$

The clustering coefficient can also measure the vulnerability of power systems, i.e. a drop in nominal performance due to a disruptive event [20], [38]. For radial distribution feeders, the clustering coefficient becomes irrelevant since it converges to zero as a result of its tree structure [39].

TABLE 2. Referenced use of characteristics in modeling power system benchmarks.

	Characteristic	Input	Constraint	Output	Validation
Topological	Node Degree	[17]–[22]	N/A	[23]–[32]	[19], [33]–[36]
	Clustering Coefficient		N/A	[21], [23], [31], [32], [37]–[39]	[20], [22], [35]
	Charac. Path Length	[17]	N/A	[29]–[32], [37], [40], [41]	[20], [22], [34], [35]
	Graph Diameter	[17]	N/A	[29], [32], [35], [37]	[20], [34], [35]
	Degree Assortativity		N/A	[30]–[32], [42]	[35]
	Centrality	[17]	N/A	[28], [32], [40], [43]–[45]	[20]
	Other Measures	[19]	N/A	[21], [31], [42], [46], [47]	[19], [36], [39], [48]
	Geography	[49]–[53]	N/A		
System	Asset Statistics		N/A	[54]	[34], [49], [55]–[58]
	Line Lengths	[17], [19], [33]	N/A	[30], [54]	[19], [33], [34], [55], [59], [60]
	Line Ratings	[19], [22], [49], [50]	N/A	[21]	[55], [59]
	Transformer Ratings	[49], [50], [58]	N/A	[54]	[34], [55], [56], [59]
	Load Ratings	[17], [22], [51], [61]	N/A		[19], [34], [55]
	Generation Ratings	[22], [61]	N/A		[19], [55]
Operational	Power Quality	N/A	[19], [22], [53], [62]	[22]	[19], [34], [48], [56], [59], [62]–[64]
	Operation & Reliability	N/A	[49], [56], [58], [65]	[54]	[55], [62], [64]
	Protection & Control	N/A		[66], [67]	[34], [48], [62], [68], [69]

3) CHARACTERISTIC PATH LENGTH

This metric, $\langle l \rangle$, represents how close two nodes are to one another by taking the average of all the possible shortest paths between two nodes as given in (4). Here, D_{ij} is the minimum length needed to traverse from nodes i to j , calculated using Dijkstra’s algorithm [71].

$$\langle l \rangle = \frac{1}{|\mathcal{V}|(|\mathcal{V}| - 1)} \sum_{\forall i, j \in \mathcal{V}} D_{ij} \quad (4)$$

Typically, $\langle l \rangle$ is compared to well-known graphs; various studies [20], [22], [29], [31], [32], [37] found that it scales faster than $\log |\mathcal{V}|$ but slower than $\sqrt{|\mathcal{V}|}$ for transmission systems, signifying that it falls between standard graphs (small-world, regular 2D). A weighted approach can also be used for actual line lengths [40]. For example, [17], [34] computed $\langle l \rangle$ of generated synthetic networks in miles and compared with actual distribution feeders. Studied MV distribution networks in China were observed to be 4.5–10.5 km for suburban and 2–4.5 km for urban regions [30]. When nodes are removed incrementally, the metric’s relative increase can help reveal a system’s vulnerability as explored in [41].

4) GRAPH DIAMETER

The graph diameter, D_{\max} , finds the largest end-to-end distance over all pairs of nodes as described in (5). Similar to the characteristic path length, the diameter of transmission networks for different voltage levels and regions scales on the order of $\sqrt{|\mathcal{V}|}$ [20], [29], [32], [35], [37]. In the case of distribution feeders, a typical range of 32–260 miles was used in [34] for validating U.S. synthetic networks, with the graph diameter of most U.S. utility feeders in [17] also found to be below 300 miles.

$$D_{\max} = \max_{\substack{i, j \in \mathcal{V} \\ i \neq j}} D_{ij} \quad (5)$$

5) DEGREE ASSORTATIVITY

The assortativity, ρ , represents the Pearson correlation in the nodal degrees on opposite ends of each edge. A high

coefficient value indicates that a random node prefers to connect to other nodes with a similar degree. Transmission networks are observed to have a small negative assortativity, e.g. -0.1 for the Eastern (EI), Western (WI) and Texas Interconnects (TI), which differ from standard graphs [31], [32], [42]. This arises due to many connections from a central substation and variation in the nodal degree values. For EHV systems, ρ tends to decrease even further [35]. On the other hand, the assortativity was found to be moderately positive, i.e. 0.5 on average, for MV distribution networks in China, indicating similar degree connectivity at lower voltage levels [30].

6) CENTRALITY

The most central node, e.g. a substation, can be found using various metrics. The betweenness, b_v , in (6) denotes the number of times a node v appears in the shortest path between any pair [20], [40]. Here, $\delta_{ij}(v)$ is the number of shortest paths that cross node v . This metric can also be defined for edges instead of nodes [43]. A high value indicates the node’s relative importance and its vulnerability. For transmission networks [20], [28], the statistical distribution of \mathbf{b} followed a power law with most nodes being close to zero. The average and maximum values were on the order of 10^3 for an unnormalized \mathbf{b} that excludes the denominator sum. This metric was not reported to be used for distribution systems.

$$b_v = \sum_{i \neq j \neq v \in \mathcal{V}} \frac{\delta_{ij}(v)}{\sum_{v \in \mathcal{V}} \delta_{ij}(v)} \quad (6)$$

Other centrality measures account for electrical properties, including the degree, eigenvector and closeness centrality, each one with different statistical distributions [45]. The electric centrality [32], [44] is based on the bus impedance matrix and explains how cascading effects occur in grids.

7) OTHER METRICS

Beyond these characteristics, there are uncommon ones defined in the literature that were beneficial in extracting certain properties of power grids.

The first useful metric is the *hop distance* from the source, e.g. a substation node, denoting the number of edges along the path. Due to the tree-like structure of most distribution feeders, this metric can help distinguish different networks. For example, a negative binomial statistical distribution was fit and used for various radial MV systems in [19] and [36].

Second, the *minimum cycle basis* relates to mesh analysis in circuit theory [72]. It represents the collection of the smallest cycles that form a given graph's basis. While the minimum cycle basis is not suitable for studying radial structures, it can be used for analyzing most transmission networks. Its statistical distribution denotes the number of cycles per size representing a graph. For example, the negative binomial statistical distribution was fit well for different North American transmission grids [39], [48].

Third, the *algebraic connectivity* is defined as the second-smallest eigenvalue of the graph Laplacian, $\lambda_2(\mathbf{L})$ representing the overall connectivity [73]. Here, $\mathbf{L} = \mathbf{K} - \mathbf{A}$, where \mathbf{K} is the degree matrix. A network is almost disconnected if its value is close to 0, whereas a unit value of $\lambda_2(\mathbf{L})/|\mathcal{V}|$ represents the fully connected case. The smallest eigenvalue is always 0 so is ignored. For transmission systems, $\lambda_2(\mathbf{L})$ was found to scale as a power function with respect to the number of nodes, lying between 1-D and 2-D regular lattices [21], [31], [42]. It should be noted that distribution systems were not studied with this metric.

Fourth, the *bus type entropy*, $w(\mathbf{t})$, was used to correlate different bus type assignments, \mathbf{t} (e.g. generation, load, connection), and a grid topology [46], [47]. It is formulated using the mathematical definition of entropy based on the relative probabilities of bus and link types across a network. Randomly permuting these assignments can cause significant deviation in a power grid's dynamics. For statistical comparison, randomized bus assignments, $\tilde{\mathbf{t}}$, of studied transmission systems were used to obtain empirical distributions of $w(\tilde{\mathbf{t}})$, which converged to a Gaussian for large sample sizes (due to the central limit theorem). The normalized distance between the empirical mean value and original $w(\mathbf{t})$ was then computed. This distance continuously increased for larger network sizes, from which a scaling property was fit and used for creating synthetic buses. In other words, larger power grids tend to deviate from randomized bus type assignments, indicating an inherent, non-trivial topological structure.

8) GEOGRAPHY

Geographical tools have gained interest by the research community and electric utilities since they can help model more accurate power systems by considering the topography and boundaries of a given region. In its simplest form, a geometrical calculation can verify whether two transmission lines at the same voltage level intersect each other using the substation coordinates. For example, the number of geographical line intersections was computed in [53] to impose a penalty for placing synthetic lines. The crossing of natural or manmade boundaries can also be verified through geographic information system (GIS) tools as in [50]–[52].

This constraint was implicitly imposed in the generation of synthetic distribution networks by accounting for settlements, environmental factors, and street maps.

B. SYSTEM CHARACTERISTICS

While topological metrics analyze networks using graph theory, system characteristics correspond to features that vary from one domain to another. For example, they can distinguish different power grids, such as transmission and distribution, and their geographical regions, based on connected assets and their ratings. Table 2 indicates that system characteristics are mostly used to generate and validate networks. They are predefined before a power system simulation (refer to the 'Input' column) and comprise statistics on utility assets, component ratings, and census information. To avoid any privacy breach, the values can be anonymized with the help of statistical distributions.

1) ASSET STATISTICS

Power grids can be identified by their number of components, such as substations, lines, and transformers. Various studies rely on statistics of asset information to quantify their availability and density (relative or spatial).

In an analysis of selected North American transmission systems, including EI, WI and Federal Electricity Regulatory Commission (FERC) form No. 715, the mean number of buses connected to a substation was 1.7–3.5 and its statistical distribution followed an exponential decay [55]. The number of lines per substation was 1.1–1.4 for each voltage level, and the percentages of substations with load and generation were 75–90% and 5–25% respectively. A similar study analyzed the percentages of substations with shunt reactive devices (e.g. capacitors, reactors) per voltage level for comparing a synthetic network with actual grids [56].

Typical ranges of a distribution feeder were extracted from thousands of U.S. utility data and reported in [34] and [49]: 94–2607 customers, 4–187 fuses, 3–392 switches, 0–5 reclosers, 0–3 regulators, and 0–5 capacitor banks. These values were used to compare synthetic grids and categorize them into different validation classes: good, marginal and to be checked. Uncommon ranges and empirical distributions were presented to justify any data outliers. Another study [61] used similar metrics as well as the number of solar photovoltaic (PV) systems and different customer types (e.g. residential, industrial) to remove correlated parameters and reduce the size of their problem.

Anonymized data was provided by 79 distribution system operators (DSOs) as part of the Joint Research Centre and covered almost 75% of connected customers in the European Union (EU) [54]. This dataset comprised main DSO indicators and their empirical distributions, such as the number of LV consumers per MV consumer (401 mean) or per distribution transformer (86 mean) and the number of MV supply points per substation (1210 mean), all of which were used to validate synthetic distribution grids in [57].

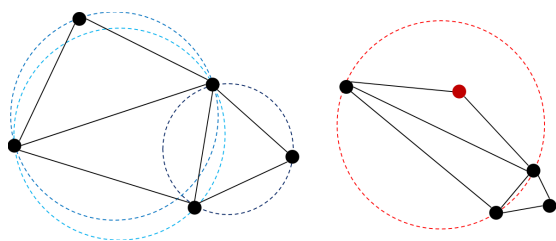


FIGURE 3. Triangulation for a set of five points on a 2-D plane: left is a Delaunay triangulation since none of the triangle circumcircles contain another point, while the right does not satisfy this condition (red point in a circumcircle).

2) LINE LENGTHS

Related to the characteristic path length (refer to Section III-A), the actual line lengths can reveal underlying structures of power grids. Transmission line lengths were used to validate synthetic grids [33], while [59] fit a generalized extreme value statistical distribution. In [55], the percentage of transmission lines on the minimum spanning tree (MST) was 45–55%, where the MST is a tree structure with the minimum sum of edge weights. This metric accounts for installation costs and geographical constraints. Another metric indicates that the ratio of total line length to the MST length is 1.2–2.2. Around 65–80% of transmission lines fall along their Delaunay triangulation, i.e. connections that maximize the smallest triangulation angles as shown in Figure 3.

Both typical and uncommon ranges of U.S. distribution line lengths were reported in [34] for various systems: LV 1-phase, LV 3-phase, MV 1-phase (overhead, total), and MV 3-phase (overhead, total). The ratios of MV line lengths to the number of customers were also used to validate synthetic feeders. Similarly, the lengths of overhead and underground lines were collected to cluster U.S. feeders [17]. In the Chinese distribution dataset of [30], the average edge length (suburban 0.2–0.4 km, urban 0.1–0.2 km) and the network length per unit area (suburban 0.6–2 km, urban 2–4.2 km) were analyzed to distinguish different sub-networks. For the EU dataset [54], the following DSO indicators were reported: LV circuit length per LV consumer (30 m mean), MV circuit length per MV supply point (1.04 km mean), and LV and MV underground ratios (66% and 60% mean). Cable lengths of a DSO dataset from the Netherlands were fit well to a modified Cauchy distribution for synthesizing MV feeders [19].

3) LINE RATINGS

Studying a power system requires defining line ratings, such as the impedance and thermal limit. To create synthetic transmission grids, the distributed line reactance and resistance (Ω/km), per-unit reactance and MVA capacity were all modeled using FERC data in [59]. Line impedances of standard test cases (IEEE-300, NYISO) were also fit to various distributions [21]. Another study [55] validated a synthetic grid if at least 70% of its transmission line ratings fell within the 10–90 percentile range, specifically the per-unit

reactance, X/R ratio and MVA limit (each per voltage level). This “rule-of-thumb” range was based on analyzing the EI and WI datasets. Similar approaches can be used for distribution line ratings.

4) TRANSFORMER RATINGS

Ratings of power system devices such as transformers can help understand a typical grid’s structure. Around 80% of per-unit reactances for EI and WI were within 0.05–0.2 [55], while the X/R ratios and MVA limits ranged between 0.1–1.4 and 10–1400 respectively. The latter two were directly correlated with the voltage level. Also, the percentages of transmission transformers with off-nominal tap ratio, phase angle regulation, LV voltage magnitude regulation, and impedance correction table were given in [56]. Statistical distributions were fit well to per-unit reactances and MVA capacities of substation transformers for each kV level [59]. For distribution transformers, the typical range and empirical distribution of their total MVA size per feeder was reported in [34]. Similarly in the EU DSO dataset [54], the distribution transformer capacities per LV consumer (4.76 kVA mean) and in urban/rural areas were provided.

5) LOAD RATINGS

From EI and WI, the mean load per bus was reported to be 6–18 MW for transmission systems [55]. Another study [22] used EI to calculate the active and reactive power consumption per capita (2 kW, 0.57 kVAR) and assign geographical coordinates to synthetic load buses. For distribution systems, the total kW load per U.S. feeder was modeled in [34], while the kW peak demands (winter, summer) and the consumed kVA and kVAR for various load types were used to cluster feeders in [17] and [61]. Since a utility aims to evenly distribute loads across a feeder [19], [74], the power consumed was modeled as a uniform distribution. The main characteristics of Italian feeders were also given in [51], such as the GWh consumption per year across agricultural, industrial and residential loads for rural, urban and industrial feeders. Daily profiles were reported for different load types across seasons, weekdays and climate zones, all of which were used to create modern benchmarks of distribution networks.

6) GENERATION RATINGS

Generator capacities in the EI and WI datasets were within 25–200 MW for more than 40% of generators (includes up to 1000 MW), while their capacity per load was 1.2–1.6 indicating reserves in case of contingency [55]. The ratio of the maximum reactive to active power was found to be 0.40–0.55 for more than 70% of generators. Generation details specific to conventional, wind, solar and other types beyond 1 MW based in the U.S. can be retrieved from [75], e.g. geographical location, grid voltage, MW capacity, and nameplate power factor. Similar surveys for other regions, e.g. [76] in Europe, can be found and used.

In [61], the kW generation capacity of PV systems and summer kVA capability were reported for U.S. distribution

feeders. Active nodes injecting power were modeled through a statistical distribution [19]. Compared to bulk generation and transmission, the active power output of distributed generation units is significantly lower, in the order of kW.

C. OPERATIONAL CHARACTERISTICS

After setting the topology and component ratings, simulation results can help predict the operational behavior within modeling limits and assumptions. Table 2 demonstrates that operational characteristics are mostly used as constraints when generating benchmarks (e.g. checking for power flow violations to reinforce a transmission network) or validating simulation results (e.g. to conform with operation standards).

Each application has certain requirements [77], [78]; for instance, an optimal power flow assumes prior knowledge of electric parameters. Various characteristics discussed below can be treated as constraints upon formulating an optimization problem (refer to Section V), since simulation results are unknown in advance. They provide a range of permissible or unacceptable values to ensure realism and integrity of a synthetic grid's performance.

1) POWER QUALITY

The most common application analyzes electrical quantities for safe operation. Typical checks include the voltages at different buses and active/reactive power flows [79].

The study of [22] computed geomagnetically induced currents of substation transformers along with bus voltages and reactive power losses. These steady-state results were used to analyze a grid's voltage stability due to solar activity affecting the Earth's magnetic field. Another paper on reactive power planning [56] compared the empirical distribution of voltage magnitudes for load buses with those of the EI, WI and TI datasets to validate test case operation. Other works [59], [64] studied or fit statistical distributions of line active power flows per voltage level to ensure consistency with actual grids. Another metric is the line loading defined as the ratio of apparent power flow with respect to the MVA rating at surge impedance loading,¹ which followed an exponential decay for EI and WI [48], [63]. The power flows across all branches were statistically modeled and used in [62].

For distribution networks, the per-unit voltage drops, power flows and estimated current ratios to the nominal cable ratings were fit to various statistical models [19]. Krishnan *et al.* [34] followed the service voltage criteria (0.95–1.05 p.u.) in the ANSI C84.1 standard [80], [81] to limit violations for all feeder nodes and avoid any overload or voltage sag/swell. The voltage unbalanced factor was limited to 3% based on [82], as well as keeping system losses below 10%, loading transformers up to 100% of their kVA rating and monitoring the power flow convergence (fewer than 20 iterations).

Recommended practices for inhibiting harmonics discussed in the IEEE 519-2014 standard [83] can also be used.

For power systems operating below 69 kV, the total harmonic voltage distortion must be limited to 5%, while each component should not exceed 3%. These tolerances are more stringent for higher voltage levels. Individual current distortion limits are set for various harmonic orders and short-circuit (SC) to load current ratios at the point of common coupling, e.g. 4% up to the 10th harmonic and an SC ratio below 20. International standards such as IEC 61727 [84] define harmonic distortion limits of injected currents for distributed PV systems, helpful in designing current controllers to compensate the harmonics and improve the power quality of grid-tied systems [66].

2) OPERATION AND RELIABILITY

When timescales exceed minutes and hours, studies such as load forecasting and optimal power flow are used to dispatch generation units in regular intervals while considering demand variations, power interchange, system economics, and other constraints [85].

More than 50% of the generators are dispatched beyond 80% of their MW capacity and 60–80% are committed for EI and WI [55]. While these values may vary by utility, they provide insights on typical operating conditions of large-scale power generation. Upon considering 12,000 contingencies of single-element outages on a synthetic 10k-bus system [56], around 300 violations were observed (overloads, voltage limits) and the power grid was secured by manual adjustments. To evaluate the robustness of synthetic transmission networks, cascading failures were simulated [64] to compute the yield,² the total number of failed lines and the number of connected components at the end of a cascade. These were compared to those of WI for different line factors of safety, i.e. extra margins for line flow capacities.

Various reliability and supply indices are defined [86] and used to evaluate a distribution system [65]. For example, the average system interruption frequency index (ASIFI) uses the total connected kVA of loads interrupted with respect to those served, while the duration index (ASIDI) relies on the interruption period. The EU DSO report [54] provides statistical distributions of SAIDI and SAIFI indicators for long unplanned interruptions. To demonstrate economic operation, locational marginal pricing was considered to compare optimal power flows of synthetic grids with the Polish 2383-bus system [62]. The range can indicate the realism of a grid's operation; negative values occurred for congested lines and a cheap marginal cost generator could not operate at its maximum limit.

3) PROTECTION AND CONTROL

Short timescales on the order of milliseconds include protection and control studies, e.g. device coordination, faults and frequency control, which are used to countereffect transients, avoid damages to equipment and ensure resilience [87].

¹Terminating impedance of a lossless line giving a flat voltage profile [77].

²Ratio of demand supplied at the end of a cascade to original demand.

For distribution systems, international standards [88] have established criteria and requirements for DER interconnection and their associated interfaces. Modeling specifications from these criteria can be set such as controlled islanding and low-voltage ride-through. In [34], various ranges of SC currents were used to validate large-scale synthetic networks: 20–40 kA for 138 kV or more (transmission) and 69–138 kV (sub-transmission), 0.3–40 kA for 1–69 kV (MV distribution), and 0.5–100 kA for less than 1 kV (LV feeders) according to [82], [89]. SC current levels and impedances of power transformers based on their thermal limits can also be referred from IEC 60076-5 [90].

Transient angle and frequency stability relate to rotor angle dynamics of synchronous generators [87], [91]. Fast disturbances such as lightning and switching need simulations in the order of milliseconds or lower. Voltage stability aims to maintain acceptable limits at all buses after a system disturbance. A synthetic transmission model was extended for transient stability [68], and several metrics were developed for its validation following N–1 contingency events. These metrics include the successive positive peak ratio (less than 1°) and the minimum damping ratio of the rotor angle (below 3%), the minimum/maximum values (59.5 and 60.5 Hz) and rates of change of bus frequency (less than 0.5 Hz/s), and the minimum ratio of the bus minimum voltage to pre-contingency levels (at least 75%). Simulated frequency responses of EI, WI and TI after a sudden loss of generation were reported in [67] and used as reference for the grid modeling. Also, the voltage angle difference across a branch, which is proportional to the active power transfer, was modeled in [48], [62] for synthetic transmission grids.

IV. DATA COLLECTION AND PRE-PROCESSING

After problem specification, Steps 2–3 of Figure 2 correspond to collecting and pre-processing datasets for generating synthetic benchmarks. Sufficient and relevant information must be gathered from various sources as an ongoing task. A power system representation generally comprises the graph topology (nodes, edges, connectivity), system characteristics (e.g. line and transformer ratings) and statistical distributions that capture a grid's structure or behavior. Some examples include census information for residential loads and households, and anonymized network data from utilities. Geospatial representation of regional districts can be extracted through GIS tools to consider geographical aspects and appended on other layers as abstracted in Figure 1.

However, it should be noted that a collected dataset may not be stored in an interpretable or useful format. The network data is often modeled in a power system software depending on the considered application (e.g., OpenDSS, GridLAB-D, CYME, PSCAD, MATLAB/Simulink) which must be appropriately converted for the development stage. If the dataset is not small, an import tool should be implemented and used instead of a manual conversion. Almost always, data cleaning is required to guarantee that there are no missing, inaccurate, invalid or inconsistent values, which can be

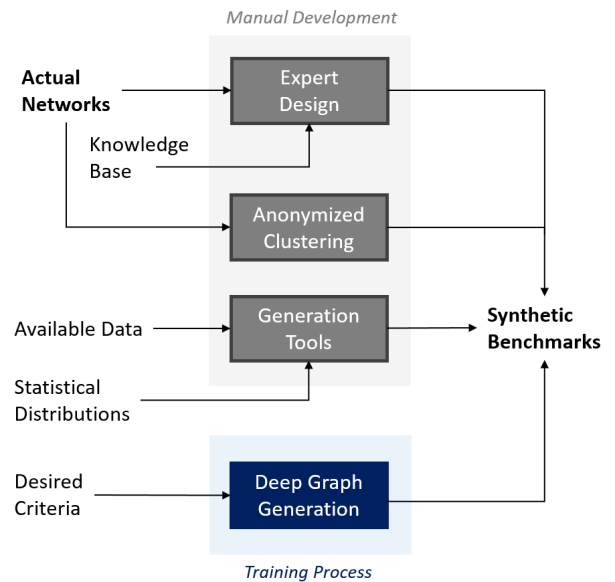


FIGURE 4. Different types of models for generating synthetic benchmarks.

time-consuming due to domain and problem-specific requirements. Data anonymization can help maintain confidentiality and hide sensitive information for minimizing security risks of learning real information from a synthetic grid. After all necessary considerations for data post-processing are complete, the dataset is ready to be used for developing models that generate synthetic benchmarks.

V. MODEL DEVELOPMENT AND NETWORK GENERATION

The first half of the generation process in Sections III–IV focused on specifying the problem, collecting data and pre-processing it for ensuring valid use. Upon defining the application and practical considerations, generative models are developed and generated in Steps 4–5 of Figure 2.

Several groups developed different types of generative models abstracted in Figure 4 to account for different levels of modeling grids, such as transmission or distribution systems, and North America or Europe. These high-level types are (i) expert design, (ii) anonymized clustering, and (iii) generation tools. The required inputs may not necessarily be the same, and each one can be broken down based on the adopted approach, e.g. heuristic algorithm or statistical sampling for generation tools. Developing a generative model may rely on manually handcrafting features or implementing a training process to automatically learn from a dataset. Deep graph generation, an alternative to manual approaches of development, is introduced and discussed in the next section. The following subsections summarize, explain, and compare the various methodologies in the literature for each type.

A. EXPERT DESIGN

Power system benchmarks were originally created and published by active working groups such those at IEEE [6], [10] and CIGRÉ [8], [11] for performing power flow simulations

in the early years of digital computing. Working group members typically comprise academic and industry experts with deep understanding of the structure and operation of power systems. During the development stage, single-line diagrams of actual transmission or distribution systems were used to manually design benchmarks and incorporate features to facilitate subsequent analyses. In general, these standard test cases provided a common basis for researchers and industry professionals to compare results of novel algorithms and ensure validity before field deployment. Some datasets are discussed later in Section VIII-B.

Each of the IEEE distribution test feeders [6], [9] was developed for a particular application and the intended use has evolved to reflect the historical needs of research and development. For example, the 1991 4-node feeder was originally used to evaluate 3-phase transformer models and their connections, while the more recent 8500-node radial system of 2010 was developed to test the scalability of power flow algorithms considering a phase-unbalanced grid in North America. Other power system benchmarks model highly meshed urban areas, include a variety of power system components, evaluate the protection of a distributed wind turbine generator, and account for daily load profiles. Regional differences are also considered when developing these benchmarks as shown in Table 1. For example, the initial purpose and connectivity are significantly different between North American and European distribution networks, resulting in multiple versions of a given benchmark [11]. The use cases, benefits and limitations of standard test systems are well-documented in [12].

B. ANONYMIZED CLUSTERING

A natural extension of expert design is reducing a set of actual networks and selecting a representative one from a common group. This clustering, top-bottom approach does not require to explicitly incorporate domain knowledge. Several networks can be mapped onto a predefined feature space and clustered together based on their characteristics as illustrated in Figure 5. A network can then be selected to fully represent each cluster and anonymized to avoid any confidentiality risks of releasing proprietary or system information.

More than a decade ago, the Pacific Northwest National Laboratory (PNNL) released a taxonomy of 24 radial test feeders representative of the U.S. distribution grid to facilitate analyzing smart grid technologies [17], [18]. While many factors determine a feeder's topology and structure, one model cannot fully represent most feeders. Publishing information of all feeders, however, can be redundant for power system studies, especially when particular regions have similar characteristics. The adopted procedure addressed these concerns through these steps, similar to Figure 5:

- 1) A total of 575 feeder models were collected from 17 different utilities, including public districts, municipalities, and rural areas. These models were pre-processed to ensure data consistency and quality.

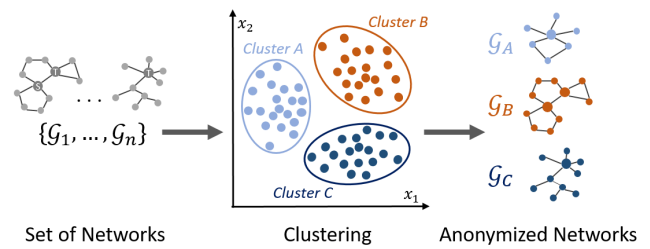


FIGURE 5. Typical methodology for anonymized clustering: n networks are projected onto the 2-D feature space, from which three clusters are found.

- 2) Regional U.S. differences in feeder design and operation were identified. Two largest factors of a distribution feeder were set as its voltage level and location based on a coarse classification, which is why the U.S. map was divided into 5 major climate regions: cold (north), temperate (west), hot/arid (southwest), hot/cold (central east), and hot/humid (southeast).
- 3) Statistical variables were used to hierarchically cluster the 575 distribution feeders based on region and voltage level. These variables include topological, system, and operational properties. The prototypical models were organized by their region, base kV voltage, kVA feeder loading, and service area description.

Another development by PNNL was the Sustainable Data Evolution Technology (SDET) for generating large-scale open-access grid datasets [92], [93]. Existing power system models are fed as inputs to create sub-networks. These ‘fragments’ constitute building blocks of a power grid and are created by removing tie lines and boundary buses between sub-networks. All sensitive information (e.g. bus names/numbers, geographic details) is replaced by random values to preserve anonymity. Equivalent generators and loads are modeled at the fragment boundaries to guarantee power balance and maintain the same operation. Next, different fragments are ranked based on user specifications (e.g. asset statistics and other system properties) before reassembling them into a synthetic grid. During this process, operational constraints are checked and tie-line impedances between fragments are modified until power flows have converged and N–1 contingency is enforced along with the bus voltage limits. The SDET generation tool is available for use at [94].

The Cluster-and-Connect model [31] generates synthetic transmission grids based on topological and system characteristics. First, the main clusters were identified using a Kirk graph [95], which sequentially places nodes along a circle and displays edges as chords connecting two nodes. Since nodes are typically numbered based on their geographical proximity, distinct groups with few crossings within the circle can be identified. Second, the intra- and inter-clusters were synthetically created using the identified clusters; statistical sampling of nodal degrees was mostly used to connect the synthetic nodes. Third, isolated nodes were

reconnected using a distance matrix by adding more edges. Line impedances were finally assigned to edge weights based on various statistical distributions.

Similarly, German LV distribution networks were clustered in [96] by their system characteristics. The analyzed dataset included the total feeder length, underground cable length, overhead line length, rated apparent power of distribution transformers, and number of delivery points (load buses). This dataset was then reduced to two dimensions through principal component analysis before using k -means clustering, resulting in six feeder categories identifiable by their line lengths and supply point densities. All these publications on anonymized clustering are summarized and compared with other approaches later in Table 4 of Section VIII.

C. GENERATION TOOLS

Generation tools are the most common type of generative models and can be divided based on whether they are for transmission or distribution systems. Available data and statistical distributions are inputs to these models, which can be further categorized based on their overall approach: *heuristic algorithms* or *statistical sampling*. The former attempts to approximate solving the generation problem through an iterative algorithm while satisfying constraints; the latter samples values from statistical distributions meant to represent the structure and operation of an actual grid. While some methods employ a hybrid approach, the dominant ones in each reference are highlighted below.

1) TRANSMISSION SYSTEMS

In [33] and [64], synthetic transmission networks were generated using tunable parameters based on North American grids to enhance resilience and efficiency. The proposed procedure focused on statistical sampling, as follows:

- 1) A Gaussian mixture model is fit to spatial node positions of an actual grid and grouped into clusters. Synthetic node locations are then sampled from the spatial distribution.
- 2) Generated nodes are connected using a tunable weight spanning tree algorithm that prioritizes selecting nodes closest to a dense cluster as in Figure 6. The parameter controls the sampling process affecting the topology.
- 3) More edges are added based on distributions of node distances and degrees to increase the grid's robustness and adjust its properties for resembling a real network.
- 4) Synthetic networks are validated through statistical tests of a few topological metrics (node degree, average path length, clustering coefficient) against real grids.

The synthetic and actual grids were visually and statistically compared for their topological properties (node degree and line length distributions). A Julia-based implementation based on U.S. census and generator data is available at [97]. However, the lack of nominal line voltages, reactive power demands, and transformer characteristics is a limitation.

A statistical methodology was also adopted in [59]. Initially, power system data and geographical information are

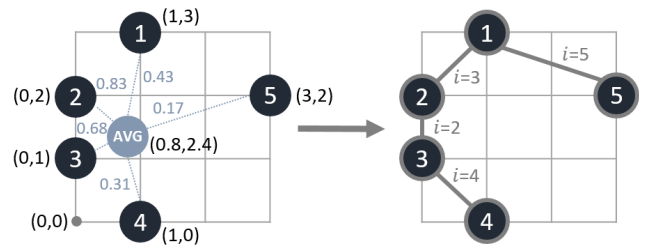


FIGURE 6. Example of the algorithm in [33]: nodes are connected to nearest neighbors using inverse distances (dotted).

analyzed to extract parameters of statistical distributions, e.g. per-unit reactances of lines and transformers. These distributions are categorized based on nominal voltage levels before verifying their relationships using electric circuit laws. Introduced in [21] for synthesizing scalable grids with small-world topologies, the *RT-nested-SmallWorld* model was enhanced to account for voltage-level dependent transformers and transmission lines. Combined with correlated bus assignments (generation, load, connection) [47] and sampled line ratings, this model generated synthetic transmission networks with adjustments made based on DC power flows. Statistical distributions and their results were reported for system and operational characteristics and compared with FERC data.

Moving onto heuristic algorithms, an optimization procedure was developed in [62] to create synthetic power flow cases. The following constraints were considered in the formulation: nodal energy balance, linearized branch flows, node permutations, power flow limits, and minimum losses. To find the optimal system and operational values, the sum of active power generated in all nodes and reactive power flows and losses in all branches were minimized. This problem was approximately solved using an evolutionary algorithm with some modifications to the original formulation due to computational bottlenecks. This optimization approach resulted in synthetic cases comparable to the Polish 2383-bus and *RT-nested-SmallWorld* systems, useful for subsequent power flow, economic, and expansion studies. However, geographical aspects were not considered in the modeling process. An open-source MATLAB package named SynGrid [98] was developed based on this work and [21], [47].

Another heuristic method was developed in [22], [53], and [56] for generating synthetic transmission networks across a geographical area. Specifically, ground resistances and geographical coordinates of substations were modeled, mainly for a geomagnetic disturbance study and reactive power planning. The inputs to this model include (i) statistics of overlaid graphs, (ii) nominal voltage levels, (iii) geographical coordinates of loads and generators, and (iv) load demands and generation capacities. The methodology's steps are described as follows:

- 1) Using public information, synthetic substations with load and generation profiles are clustered and sited. Demand levels are modeled and estimated using

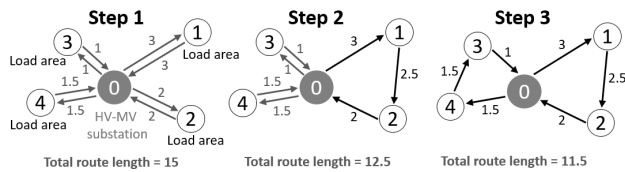


FIGURE 7. Example of the savings heuristics used in DINGO.

population size. Grounding resistances are assigned based on substation size, voltage level, and number of buses.

- 2) Transmission lines are placed to connect substations using grid statistics, i.e. node degree distribution, average shortest path length, and average clustering coefficient. Delaunay triangulation, as in Figure 3, is used to link substation nodes based on nearest neighbors while preserving properties of real transmission systems.
- 3) A synthetic network is built using a Euclidean minimum spanning tree per voltage level. More lines are iteratively added based on DC power flows and the average degree.

Upon creating synthetic transmission networks, their topological, system, and operational metrics were validated against the actual grid. For example, the single-line diagram and voltage magnitude map of a 10k-bus system were compared with the WI's summer operation. This methodology was extended to include transient stability analysis in [68] and synthetic load profiles in [99] based on open-source data. A generative process purely based on topological characteristics is proposed in [35] to create transmission networks comparable to the EI, TI and Polish datasets for different voltage levels.

Complex-network techniques were devised in [20] to generate synthetic transmission networks for European grids. While nodes with electrical and geographical properties are generated and grouped based on spatial distributions, the difference lies in the economic factors considered; assigning a given generator (supply) is prioritized for a higher load (demand). The tradeoff between investment and operation costs over a given time period is considered. After connecting an initial graph, the network is reinforced based on DC power flow and N-1 contingency analysis.

2) DISTRIBUTION SYSTEMS

The Distribution Network Generator (DINGO) [58] was developed to create MV and LV grids based on open or accessible data. Its inputs include (i) spatial representation of demand (areas, districts), and (ii) load and generation levels. The main assumptions are (i) MV grids are connected in a ring topology to resemble 84% of Germany; (ii) all lines and components are 3-phase based on Europe; (iii) a fixed power factor is set for all loads and generators; (iv) the model ignores natural boundaries and transport infrastructure; and (v) only the peak demand values are used.

To generate a synthetic network, DINGO considers a capacitated vehicle routing problem, i.e. finding the optimal set of lines for delivering power to demand areas while considering line capacity constraints. The objective is to minimize material and line installation costs by reducing the total line length. Its overall methodology is summarized below:

- 1) Initial MV lines are built using Clarke and Wright's savings heuristic as in (7), where D_{ij} denotes the distance between nodes i and j , and S_{ij} is the expected savings when looping through both nodes. Line lengths are minimized by computing S_{ij} in each step while checking for constraints (line congestion, voltage violation). An example is shown in Figure 7, where an HV-MV substation is connected to surrounding load areas. Instead of direct routing, new lines are incrementally added between loads to form loops, satisfy N-1 criteria, and reduce line costs. This heuristic repeats until no improvements are left.

$$S_{ij} = D_{0i} + D_{0j} - D_{ij} \quad (7)$$

- 2) Stable solutions in the previous step could be suboptimal, so a local search heuristic iteratively explores neighboring solutions to reduce the total line length.
- 3) Missing assets such as islanded load areas, distribution transformers, and DERs are connected.
- 4) Violations of constraints are checked using power flow. The topology and equipment types are reinforced to resolve any issues with line capacities and overvoltages.

The resulting networks were statistically validated for assets using 3,608 MV grids in Germany. Between the real and synthetic data, the number of HV-MV substations and distribution transformers deviated by 10.3% and 8.3% respectively, while the total cable length (sum of underground and overhead) was off by 2.3%. Other topological and operational criteria, however, were not validated. An open-source implementation of DINGO based on Python is accessible from [100].

A similar work [52] creates distribution networks of various voltage levels using available data. Geographical data is extracted from OpenStreetMap before building the HV grid using lines and substations visible from satellite imagery. Next, a Voronoi diagram is created to separate regions and satisfy a minimum distance with a substation's customers. Distribution transformers are positioned closer to high load densities and clustered using k -means to represent a given load area. As in DINGO, a heuristic algorithm based on the traveling salesman problem is solved to minimize the total MV line length in a ring topology. The LV grid is finally created based on building and street information, i.e. the rated power per load point is scaled based on an average building area. The nodal voltage and line capacity constraints were respected for the case study. Extensions to this methodology with time series load profiles are given in [51] and [101].

Another large-scale planning tool is the Reference Network Model (RNM) used for designing MV and LV

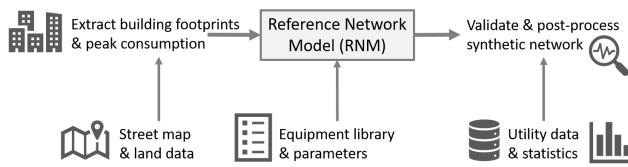


FIGURE 8. General methodology of RNM for generating synthetic networks.

distribution systems comprising substations and feeders [50], [57], [65]. Its main purposes were to (i) build a power grid in a cost-efficient manner, (ii) assess distribution networks costs under incentive regulation, and (iii) investigate the impact of integrating DERs in distribution networks. The RNM tool is developed and maintained by [102], available under research or commercial license. Its inputs include

- Georeference data of customers and transmission substations, including coordinates and contracted power;
- Geographical data and constraints, i.e. street map, orography, lakes, and nature reserves;
- Standardized equipment library of substations, transformers, lines, and protective equipment; and
- Set of technical and economic parameters, e.g. continuity of supply targets, demand increase, and loss factors.

The general methodology is illustrated in Figure 8. Initially, a geographical dataset is pre-processed to extract building footprints and peak consumption. This represents the spatial load representation similar to DINGO, plus considering the street map and land constraints. Upon providing its parameters and standardized equipment library to appropriately size assets, RNM generates a synthetic network using a heuristic optimization algorithm based on the minimum spanning tree, Delaunay triangulation, and a branch-exchange technique [103]. Next, the generated network is validated with reference grids (e.g. through utility statistics) and post-processed to model additional information and assess the planning costs. Future scenarios can be defined to expand and reinforce the initial network and to connect more resources. As in Figure 1, RNM is organized into four modeling layers, each comprising data and algorithms: (i) topological, (ii) geographical, (iii) electrical, and (iv) reliability of supply.

Since RNM's development, different groups have adapted its approach in their research projects, e.g. the Distribution Network Model (DiNeMo) [104] for Europe. Similarly, the National Renewable Energy Laboratory (NREL) collaborated with IIT Comillas to develop the Synthetic Models for Advanced, Realistic Testing: Distribution Systems and Scenarios (SMART-DS) platform for generating synthetic distribution network models and scenarios based in the United States [105]. A variation of this tool which considers phase selection for North American feeders was developed known as RNM-US [106]. Their research efforts resulted in various synthetic distribution networks; the largest represents the Bay Area with more than 2 million customers, 2,236 feeders, and 148 substations [34], [49]. This model was integrated with [22], [53], [56] to generate a

combined transmission-distribution system for the central Texas area [107], ranging from 120 V to 230 kV, with 448 feeders. The distribution system was created using RNM-US before planning the transmission (generators, substations, lines) and interfacing them. Published datasets of synthetic networks are discussed in Section VIII-B.

Aside from heuristics, a statistical approach was developed in [19] to generate synthetic MV and LV feeders. Using a DSO dataset from the Netherlands, statistical distributions were fit to relevant metrics and properties. The following distributions were used to generate radial networks by exploiting trends in the data: (i) hop distance to source, (ii) node degree, (iii) fraction of zero load nodes, hop distance of intermediate nodes, deviation of power injection, and deviation of power consumption, (iv) ratio of estimated and nominal cable current, and (v) cable length.

Upon fitting the statistical models, the feeder nodes were generated using a negative binomial distribution of the hop distance from the HV source to impose the radial topology assumption. Next, the created nodes were connected based on the degree distribution which follows different rates of exponential decay. The properties of intermediate (no load), generation, and load nodes were then assigned before setting the cable types and lengths based on the available equipment library and predefined model distributions.

These synthetic feeders were visualized for a few instances and statistically compared to actual data using operational metrics. Good matches in the metric distributions, including downstream power and voltage drop, suggested that the synthetic feeders resemble the real data well. It should be noted that other metrics of realistic behavior were not tested and may create deviations in a feeder's performance. Also, transformers and capacitors were ignored in the generative process as well as geographical considerations.

VI. DEEP GRAPH GENERATION

The performance of the previous approaches may be limited by their assumptions and methods. They generally assume that certain characteristics are more important for generating networks, e.g. nodal degree. Ignoring the underlying distributions of realistic grids by using a fixed hand-crafted process, however, can limit a model's performance and introduce bias. To address these gaps, machine learning can be used to learn hidden representations from a training dataset and avoid hard-coding specific properties into a generative model.

Recent advances to computational power and training algorithms were instrumental in adopting machine learning across various disciplines. In general, a problem can be solved using machine learning if a sample dataset is representative of the target application and on which a learning algorithm can be trained to predict and generalize for unseen data. Significant improvements are possible in tackling hard problems, since there is less reliance on domain experts to manually develop specific algorithms, and the same model can be reused for other problems with similar structure [108].

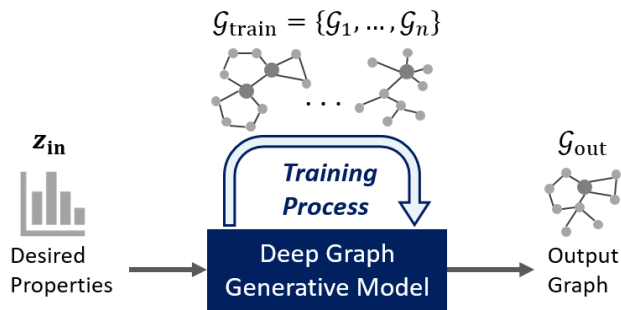


FIGURE 9. General training process for deep graph generative models.

Machine learning models are generally adapted to work with tabular data (e.g. characteristics), structured grids (e.g. images), and sequences (e.g. time series). Existing learning frameworks, however, are not well-suited to train and evaluate high-dimensional, non-Euclidean graph structures. It is challenging to generalize and design learning frameworks to be independent of topological limitations [109]. Graphs can be irregular with a variable size of unordered nodes, and each node may have a different number of neighbors. These complexities compound the difficulty when learning the representation of graph-structured data [110].

Within power systems, machine learning can provide fast and intelligent decision making as well as contribute to increased grid flexibility and DER integration [111], [112]. Some applications for incorporating learning techniques include microgrid operation and control [113], fault diagnosis for protection [114], and load forecasting [115]. Most of these topics can be considered as regression or classification tasks positioned under supervised learning, i.e. when the training dataset consists of both features (inputs) and labels (outputs). Deep graph generation, as explained below, may also have a widely explored role in power systems.

A graph \mathcal{G} defined in (1) is used to learn hidden or latent embeddings (e.g. node, edge or graph) for performing tasks such as node classification or relation prediction [116], [117]. Deep graph generation inverts this problem as displayed in Figure 9, where a given model learns the hidden representations from a training dataset, $\mathcal{G}_{\text{train}}$, without user interference or the creation of hand-engineered features. After training and during evaluation, the desired properties, \mathbf{z}_{in} , are fed to the generative model to create an output graph resembling those in the training dataset. The challenge is to develop tractable methods for generating graphs (for sampling or analyzing them) that have certain properties.

The subsequent paragraphs discuss general approaches of deep graph generation, mainly for creating novel molecules, modeling social networks, and generating new benchmarks. Since this subfield is relatively new with most prominent works published recently, technological advances are expected in the upcoming years with possible applications for generating synthetic benchmarks for power systems. A high-level summary of these approaches is provided in Table 3.

TABLE 3. Different approaches for deep graph generative models.

Approach	Benefits and Limitations
Variational Autoencoder (VAE)	+ Well-defined probabilistic framework + Interpretable latent space – Limited generative performance – Prior distribution assumption – Node ordering dependence
Generative Adversarial (GAN)	+ Implicit, no likelihood computation + Node ordering independence – Training instability (mode collapse) – Lack of convergence metrics
Reinforcement Learning (RL)	+ Desired properties using rewards + Novel graphs through exploration – Sensitive to reward function – Complexity of environment setup
Autoregressive	+ Accounts for edge dependence + Graph-scalable framework – Sequential generation process

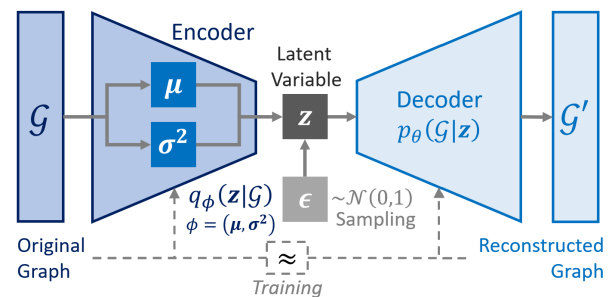


FIGURE 10. Variational autoencoder architecture with multivariate Gaussians.

A. VARIATIONAL AUTOENCODERS

The *variational autoencoder* (VAE) proposed in [118] with architecture shown in Figure 10 is a popular deep generative model that can be applied for graph datasets. In summary, a VAE consists of the following components:

- A probabilistic encoder, $q_{\phi}(\mathbf{z}|\mathcal{G})$, defines a distribution over a latent representation, \mathbf{z} , for an input graph, \mathcal{G} , for a power system. Gaussian random variables are typically used to design this encoder conditioned on \mathcal{G} , where $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}^2$ are neural network outputs for the mean and variance vectors. Both parameters are used to sample the latent vector \mathbf{z} that feeds into the decoder.
- A probability decoder, $p_{\theta}(\mathcal{G}|\mathbf{z})$, takes a latent representation \mathbf{z} to specify a conditional distribution over graphs. In other words, a power system's graph is compressed as a real-valued vector for the decoder to use. This function computes entries of an adjacency matrix through conditional probabilities, $p_{\theta}(A_{ij} = 1|\mathbf{z})$, i.e. existence of an edge between any pair of nodes, in order to generate a reconstructed graph, \mathcal{G}' .
- A prior distribution over the latent space, $p(\mathbf{z})$, is defined as a standard Gaussian, $\mathbf{z} \sim \mathcal{N}(0, 1)$, allowing stochastic behavior for graph generation. Based on the principle of maximum entropy, a Gaussian distribution can be chosen since it has the least prior information of all real-valued functions with a specified variance.

A VAE reconstructs a high-dimensional input by passing it through a bottleneck layer to encode its vital information. The objective function maximizes the decoder's reconstruction ability (i.e. $\mathcal{G} \approx \mathcal{G}'$) and minimizes the difference between the encoder's posterior latent distribution and the assumed prior (e.g. Gaussian). The sampled z encodes sufficient information for a decoder to reconstruct \mathcal{G} , whereas matching z with the prior ensures meaningful graphs are decoded upon sampling. For example, consider a dataset of several power systems. Each one is used to learn VAE parameters: ϕ and θ . Once training is complete, the encoder is dropped and the remaining decoder can be used to generate synthetic power systems by directly sampling z . The sampling space of z can also be interpolated within defined bounds to generate hybrid solutions.

Among graph-based models, VGAE [119] uses node-level embeddings; its encoder generates latent representations for each node, and its decoder computes the dot product per pair of embeddings in order to predict the edge likelihood between two nodes. VGAE was used to learn and visualize the latent space of a citation network in an unsupervised manner. However, it suffers from a rather simplistic and non-parametric decoder. Instead, GraphVAE [120] encodes the graph topology by pooling functions and decodes using a feedforward neural network, i.e. a multilayer perceptron (MLP), to predict a synthetic adjacency matrix. It was trained on a dataset of organic molecules and produced novel, chemically valid molecules by interpolating its latent space. For power systems, these models can help reveal representative features of a dataset by interpreting and analyzing the compressed, latent space.

B. GENERATIVE ADVERSARIAL APPROACHES

As listed in Table 3, VAEs are explicit, likelihood-based models with accessible latent spaces. However, they suffer from serious limitations such as their tendency to produce blurry outputs due to the Gaussian prior assumption. *Generative adversarial networks* (GANs) are implicit generator models that do not need to encode the data or model distributions [121]. Only samples are drawn, whose generation quality is optimized within an adversarial framework as summarized below and in the displayed architecture of Figure 11:

- A generator network g_θ is trained to produce fake, but realistic samples $\tilde{\mathcal{G}}$ from a random seed z (e.g. sampled from a uniform distribution). The discriminator network d_ϕ with trainable parameter ϕ aims to distinguish between the real data samples $\mathcal{G}_i \in \mathcal{G}_{\text{train}}$ and those generated by g_θ , i.e. $\tilde{\mathcal{G}}$. The discriminator outputs the probability of an input being fake, p_{fake} , as in (8).

$$\begin{aligned} \tilde{\mathcal{G}} &= g_\theta(z) \\ p_{\text{fake}} &= d_\phi(\tilde{\mathcal{G}}) \in [0, 1] \end{aligned} \quad (8)$$

- Both networks are trained in an adversarial game by minimax optimization as in (9). Here, $p_{\text{data}}(\mathcal{G})$ and $p_{\text{seed}}(z)$ are the empirical data and random seed

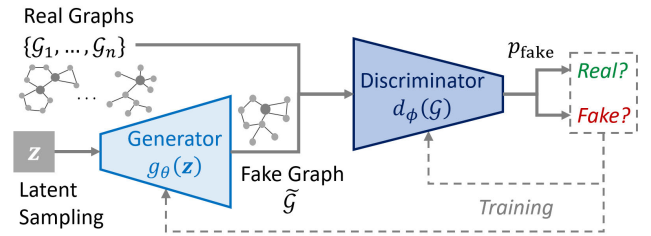


FIGURE 11. Generative adversarial network architecture.

distributions. The generator attempts to minimize the discriminator's power, while the discriminator maximizes its ability to detect fake samples. Both terms resemble the binary cross-entropy loss for the discriminator in a binary classification task, i.e. detecting fakes and originals. A difference is $g_\theta(z)$ in the second term accounts for generating a fake sample $\tilde{\mathcal{G}}$ from a random z .

$$\begin{aligned} \min_{\theta} \max_{\phi} \mathbb{E}_{\mathcal{G} \sim p_{\text{data}}} \left[\log(1 - d_\phi(\mathcal{G})) \right] \\ + \mathbb{E}_{z \sim p_{\text{seed}}} \left[\log d_\phi(g_\theta(z)) \right] \end{aligned} \quad (9)$$

At the Nash equilibrium point, the generator models the real data while the discriminator network classifies the generator's output with a 0.5 probability (unsure if fake or not). Using GAN helps circumvent the explicit modeling of p_{data} , which may be complicated (or intractable to infer) for power grids. Instead, the focus is on sampling from the training dataset, p_{data} , to train both networks against each other.

For implicit methods, MolGAN [122] generates the discrete adjacency matrix, $\tilde{\mathbf{A}}$, given z using an MLP and sampling. The same process repeats for the nodal features, $\tilde{\mathbf{X}}$. The discriminator classifies the labeled graph, $\tilde{\mathcal{G}} = (\tilde{\mathbf{A}}, \tilde{\mathbf{X}})$, that remains permutation invariant and insensitive to node orders. MolGAN was trained on the QM9 dataset to generate 98% valid and 94% novel chemical compounds, an improvement compared to GraphVAE (55% valid, 61% novel).

NetGAN [123] is another model that captures graph topologies by learning the distribution over random walks. The generator models a sequential process using a recurrent neural network (RNN). It was trained on a citation network using Wasserstein algorithm [124], which prevents mode collapse and maintains training stability. Although no topological properties were fed to the model, NetGAN learned them by training and closely matched the nodal degree distribution.

A recent publication applied the Wasserstein algorithm on graph convolutional neural networks to generate synthetic distribution networks. FeederGAN [125] encoded various component attributes, including device length, conductor capacity, distance from the source, tree level, phasing, and load value. Geographical information of feeders was removed to simplify the modeled graph structures. Using a training dataset of 664 real feeder graphs, the generated feeders were validated for their voltage distribution and other attributes through empirical statistics.

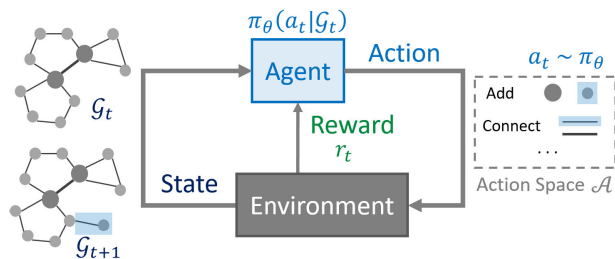


FIGURE 12. Agent and environment interaction in reinforcement learning.

Using GAN as a generative model has several advantages as seen in Table 3: there is no need to explicitly compute a likelihood, nor consider node orderings. It is possible to generate the adjacency and feature matrices, useful for creating a power system's topology and assigning its existing assets to nodes (e.g. transformers) and edges (e.g. lines). However, the main challenge is to solve its minimax objective with available computational resources. Since two networks are trained in an adversarial environment, it is hard for the generator to converge and work well at the start of training.

C. REINFORCEMENT LEARNING

Generating novel and valid graphs by directly optimizing desired properties is challenging since these objectives are complex and non-differentiable. *Reinforcement learning* (RL) is an approach where no explicit dataset or supervision is required. Generating synthetic networks for power systems can benefit from RL, especially since utility data may be limited. This problem is formulated through an agent–environment interaction shown in Figure 12. An agent self learns by interacting with an environment, e.g. a power system simulator. Every action, a_t , changes the environment's state in the next time step, \mathcal{G}_{t+1} . Based on the action taken (e.g. add a new node), the environment provides positive or negative feedback to the agent through a reward signal, r_t , which can be tied to the operational characteristics discussed in Section III-C, e.g. no voltage violations. From state and reward signals, the agent's objective is to learn an optimal policy, $\pi_\theta(a_t|\mathcal{G}_t)$, with parameter θ to take actions given a current state. This policy represents a distribution over all possible actions, \mathcal{A} , which a_t can be sampled from. For example, nodes are added at the start when no network exists. During subsequent steps, the policy can favor connecting islanded nodes to a feeder until the power flow converges and the process terminates. Due to RL's temporal nature, an agent's actions affect subsequent data that it receives [126].

RL has two main advantages as in Table 3. Incorporating desired properties (topological and operational) into the objective function is less complex, since they can be represented through a suitable environment and reward function (e.g. no voltage violations). Also, active exploration of the vast and combinatorial design space can be encouraged through an RL framework that goes beyond predefined samples of a dataset. Allowing generative models to explore

different solutions can promote the generation of new and feasible graphs.

GCPN [127] generates molecules through RL guided by objectives. An agent learns by experience and iteratively adds subgraphs in a chemistry environment within a Markov decision process (MDP). Generating a graph can be described by trajectory $(\mathcal{G}_0, a_0, r_0, \dots, \mathcal{G}_n, a_n, r_n)$, where \mathcal{G}_n is the final generated graph. At each step, the graph is modified by the state transition distribution in (10). Instead of adding nodes and edges based on the complete trajectory, the MDP assumes that the policy network π_θ only needs the state \mathcal{G}_t to select the next action, i.e., memoryless Markov property.

$$p(\mathcal{G}_{t+1}|\mathcal{G}_t, \dots, \mathcal{G}_0) = \sum_{a_t \in \mathcal{A}} p(\mathcal{G}_{t+1}|\mathcal{G}_t, \dots, \mathcal{G}_0, a_t)\pi_\theta \quad (10)$$

In a given time step, \mathcal{G}_t along with the set of subgraphs are observed by the agent to compute the node embeddings using graph convolutional networks. An action is sampled from the policy network, i.e. selecting two nodes, predicting the edge type, and predicting the process's termination. Next, the simulation environment performs action a_t if feasible and obeys domain-specific rules before computing the next state \mathcal{G}_{t+1} and reward r_t . This process continues until a terminating action and the final reward is given. A similar RL process could be used for generating synthetic benchmarks.

Most generative models discussed above assume that edges are generated independently. The likelihood of a graph given a latent representation, $p(\mathcal{G}|z)$, can be efficiently calculated using independent edge likelihoods. However, this i.i.d. assumption is strong and problematic for real-world graphs that demonstrate complex dependencies between edges. For example, connecting more loads to a distribution system depends on several factors, such as the distance from a substation and transformer capacity, affecting the interdependence of edges. One way to address this limiting assumption and maintain tractability is to use an *autoregressive* model.

D. AUTOREGRESSIVE METHODS

In an autoregressive approach, edges are assumed to be generated sequentially and each likelihood is conditioned on previously generated data as described in (11). Here, \mathbf{L} is the lower-triangular submatrix of \mathbf{A} and $\mathbf{L}[v_i, :]$ denotes a row corresponding to node v_i . The overall graph likelihood can be decomposed using (11), meaning that generating $\mathbf{L}[v_i, :]$ for node v_i is conditioned on all the previous rows generated. Hence, the i.i.d. assumption is no longer ignored and edge dependency is considered during the generation process.

$$p(\mathcal{G}|z) = \prod_{i=1}^{|\mathcal{V}|} p(\mathbf{L}[v_i, :] | \mathbf{L}[v_1, :], \dots, \mathbf{L}[v_{i-1}, :], z) \quad (11)$$

GraphRNN [128] is a scalable framework that adds new nodes and edges sequentially and can deal with variable graph sizes, account for edge dependencies, and scale up due to its node ordering algorithm. Its hierarchy consists of (i) a graph-level RNN that maintains the graph's state and generates new

nodes, and (ii) an edge-level RNN that generates edges per new node. Both networks were trained by maximizing the overall likelihood of training graphs using the teacher forcing strategy [129]. While GraphRNN performed relatively well in mimicking the structure of large-sized graphs, its shortcomings include unrealistic artifacts and random node ordering.

Instead of using RNNs, GRAN [130] considers the conditional distribution of $L[v_i, :]$ through a graph neural network given the graph generated so far. A block of rows representing multiple new nodes is simultaneously generated at each step, benefiting from computational efficiency and quality. Large graphs of up to 5,000 nodes were trained on GRAN. Compared to other models, it outperformed for different graph topologies and scaled for large datasets such as radial structures, suggesting their suitability for generating synthetic power grids.

DeepGDL [131] is another recurrent model that captures node/edge distributions and generates synthetic transmission grids using RNNs. To date, it represents one of two deep graph generative models applied to power grids, demonstrating the aforementioned benefits. The WI dataset was used for training, consisting of 14.4k buses, 18.8k lines and 72 communities. A hundred synthetic grids were generated, whose topological properties and power flow statistics were reported to be superior when compared to [64] discussed in Section V-C.

VII. SYNTHETIC NETWORK VALIDATION

Upon generating a synthetic network in Step 5 of Figure 2, it is imperative to validate its structure and performance by comparing with reference grids or available standards. The various characteristics discussed in Section III and summarized in Table 2 can be used for validation, considering that they were initially specified in the generation process. The methods for comparing them can vary depending on the available data, given application, and required simplicity.

Synthetic U.S. distribution networks in [34] and [49] were statistically validated in terms of their topological and system characteristics, checked for their operational integrity, and manually assessed by experts. Statistical and operational validation was performed by computing the percentage overlap of empirical distributions with typical and uncommon ranges obtained from utility data. A higher overlap translated to a positive grade; otherwise, marginal cases were manually checked and reinforced to meet the desired criteria by revisiting previous steps of the generation process.

Other works approached network validation differently. For example, thousands of U.S. feeders were clustered in [61] to identify the most important parameters for PV hosting capacity, including the line voltage, total feeder length, and number of regulators. Distinct feeders from different clusters were evaluated across 0–100% hosting capacity range for various operational issues: overvoltages, voltage deviations, element fault current, sympathetic breaker tripping, and breaker reduction of reach. For each issue, the computed PV penetration levels were binned into three violation classes:

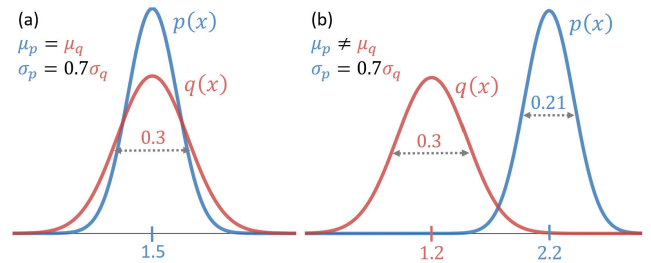


FIGURE 13. Statistical comparison of normal distributions: (a) $D_{KL}(p||q) = 0.1$, $D_H = 0.18$, (b) $D_{KL}(p||q) = 5.7$, $D_H = 0.92$.

(i) none, (ii) at specific locations, and (iii) at all locations. Feeders within the same cluster behaved similarly, and the dependency of key parameters was statistically validated.

Beyond comparing aggregated statistics (e.g. mean and range), statistical tests can be used to quantify the closeness between synthetic and reference grids (e.g. EI, WI and TI as a whole and/or for each voltage level). Capturing the same statistical distribution ensures that a given metric (refer to Section III) has similar trends, thereby validating the generated network through a goodness of fit. This approach represents the automated portion of Step 6 in Figure 2. The following paragraphs summarize and review various statistical tests that were used to validate synthetic benchmarks.

A common statistical test is the Kullback–Leibler (KL) divergence [132], [133] as defined in (12); it measures how much a given probability distribution, $q(x)$, deviates from a reference one, $p(x)$. Its value is non-negative, ranging from zero when $p(x) = q(x)$ to infinity when two distributions are completely different from each other. Figure 13 demonstrates this effect through two examples of normal distributions with different parameters. Even in the case when their mean values are similar, slight differences can still be observed in the measure due to the dissimilar variances. While the KL divergence is not a distance metric due to its asymmetry, i.e. $D_{KL}(p||q) \neq D_{KL}(q||p)$, it does not raise an issue for validation since $p(x)$ is selected as the reference distribution from realistic grids. Several works [19], [33], [48], [59] used the KL divergence to compare statistical distributions for different topological and system characteristics of synthetic networks with actual grids. Small reported values indicated that the distributions matched well with each other.

$$D_{KL}(p||q) = \int_{-\infty}^{\infty} p(x) \log \left(\frac{p(x)}{q(x)} \right) dx \quad (12)$$

Another test is the Hellinger distance [134], which computes the deviation between two distributions in (13) with a bounded value within 0 and 1. The latter term is the Bhattacharyya coefficient which measures the overlap between two distributions. Both D_H and D_{KL} have similar trends, i.e. when one is small so is the other (see Figure 13). Previous works [48], [62] used the Hellinger distance to evaluate how well a statistical distribution was fit to real data and validate

TABLE 4. Comparative summary on generating synthetic benchmark networks.

Publication			Problem Specification & Data Collection								Model Development & Network Generation						Results					
Studied System	First Author (abbr.)	Reference	Region		Characteristics			Application			Datasets		Model Type		Layers		Validation		Developed Tool	Published Dataset		
			North America	Europe / Other	Topological	System	Operational	Power Quality	Oper. & Reliab.	Prot. & Control	IEEE, CIGRÉ *	EI, WI, TI, Others	Anonym. Clustering	Generation Tools	Deep Graph Generation	Geographical	Other	Aggregate Stats.			Empirical Distrib.	Statistical Fitting
Transmission	You Hu	[93] [31]	✓		✓	✓	✓	✓			ut	✓					✓			[94]	[138]	
	Bir Xu Li	[22] [68] [99]	✓		✓	✓	✓	✓			✓	h	✓	tr			✓	✓	✓	[97]	[139]	
	Aks Sch Esp Sol	[35] [62] [20] [64]	✓	✓	✓	✓	✓	✓	✓	✓	✓	h	h	ec	ec		✓	✓	✓	✓	[98]	
	Ath Ely Wan	[59] [47] [21]	✓		✓	✓	✓	✓	✓	✓	✓	s	✓				✓	✓	✓	✓	[97]	[138]
	Kho	[131]	✓		✓		✓	✓	✓	✓	✓	s					✓	✓	✓	✓	[98]	[98]
	Sne Di	[17] [96]	✓	✓	✓	✓	✓	✓			ut	✓				✓	✓					[138]
	Am Kay Mei	[58] [52] [60]		✓	✓	✓	✓	✓	✓	✓	ge	h	h	ts			✓				[100]	
	Pis	[51]		✓	✓	✓	✓	✓	✓	✓	ge	h	h	ts			✓					[101]
	Mat Do	[49] [50]	✓		✓	✓	✓	✓	✓	✓	ut	h	h	ec	ec		✓				[102]	[138]
	Trp Sch	[36] [19]		✓	✓	✓	✓	✓	✓	✓	ut	sg	h				✓	✓	✓		[102]	
	Lia	[125]		✓		✓		✓	✓	✓	ut	ne	s				✓	✓	✓	✓		

* Other standard test cases. *ec*: economic study. *fs*: ENTSO-e & RTE datasets [140], [141]. *ge*: German dataset. *h*: heuristic algorithm. *it*: Italian dataset. *ne*: Netherlands dataset. *s*: statistical sampling. *sg*: Singapore dataset. *tr*: transient. *ts*: time series. *ut*: utility data.

synthetic performance.

$$D_H(p||q) = \sqrt{1 - \int_{-\infty}^{\infty} \sqrt{p(x)q(x)} dx} \quad (13)$$

The Kolmogorov–Smirnov (KS) statistical test measures the maximum deviation between two cumulative distributions, $P(x)$ and $Q(x)$, as defined in (14). Given the non-parametric nature of the KS test, it can effectively compare empirical distributions. Combining the resulting value of D_{KS} with hypothesis testing helps to statistically reject a produced distribution (i.e. synthetic data) if it does not satisfy a certain significance level or threshold. For example, the KS test was used in [33], [35], [48], and [62] to validate various topological, system, and operational characteristics.

$$D_{KS}(p||q) = \max_x |P(x) - Q(x)| \quad (14)$$

Another metric is the graph Relative Hausdorff (RH) distance [135] inspired by its topological variation [136]. It quantifies the closeness of two graphs based on their degree distributions. The benefit is its linear time complexity with respect to the maximum graph degree [137]. In [35], the RH distance was used to compare synthetic and actual transmission grids (EI, TI, Polish) for different voltage levels. As a

guideline, RH values less than 0.30 were interpreted as a ‘good’ match, while less than 0.10 were ‘excellent’.

VIII. APPLICATIONS OF SYNTHETIC NETWORKS

Different stages of the generation process in Figure 2 are discussed earlier. While the various approaches proposed in the literature are explained individually, identifying their differences is challenging. The following subsections comparatively summarize them, discuss research trends or gaps, and present published datasets of synthetic benchmarks.

A. COMPARATIVE SUMMARY OF LITERATURE

Two summary tables are introduced and discussed: Table 4 focuses on generative methodologies (refer to Section V), while Table 5 goes through validating networks (refer to Section VII). Each row corresponds to a publication, while various criteria are given by columns with similar ones clustered together. Studies on transmission and distribution systems are grouped in horizontal blocks.

The category of problem specification and data collection (refer to Sections III–IV) includes the geographical region, considered characteristics, applications, and datasets used. If the studied grid was not situated in North America or Europe, it is referred from the table’s footnote.

TABLE 5. Comparative summary on validating synthetic benchmark networks.

Studied System	Publication		Problem Specification & Data Collection							Network Validation					
	First Author (abbr.)	Reference	Region		Characteristics			Application		Datasets		Aggregate Stats.	Empirical Distrib.	Statistical Fitting	Goodness of Fit
			North America	Europe / Other	Topological	System	Operational	Power Quality	Oper. & Reliab.	Prot. & Control	IEEE, CIGRÉ *				
Transmission	Ros	[25]		✓	✓						✓	✓	✓		
	Pa	[28]	✓	✓	✓					✓	✓	✓		✓	
	You	[37]	✓		✓						<i>ut</i>	✓			
	Bir	[55]	✓		✓	✓	✓	✓	✓		✓	✓	✓		
	Sch	[48]	✓		✓	✓	✓	✓	✓		✓	✓	✓	✓	
	Sch	[63]	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
	Cot	[32]	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
Distribut.	Jai	[61]	✓		✓	✓	✓	✓	✓		<i>ut</i>	✓	✓		
	Kri	[34]	✓		✓	✓	✓	✓	✓		<i>ut</i>	✓	✓		
	Pre	[54]		✓	✓	✓	✓		✓		<i>ut</i>	✓	✓		
	Abe	[30]		✓	✓						<i>ch</i>	✓	✓		

* Other standard test cases. *ch*: China dataset. *ut*: utility data.

Under model development and network generation (refer to Sections V–VII), the model type for generating networks, the modeling layers considered in Figure 1 (aside from the ones under ‘Characteristics’), and the validation aspects are identified. If any reference developed a tool or published a dataset of synthetic benchmarks, they are listed in the last two columns.

In general, there are more references on transmission systems than for distribution due to the availability of real datasets, e.g. EI and WI. It is observed in Tables 4 and 5 that non-standard (and often private due to confidentiality) datasets were used for distribution networks, which can complicate the validation process for subsequent studies. Generalizing distribution networks is not easy as they tend to evolve based on utility and client requirements and must consider more geographical constraints compared to transmission networks. For example, neighborhoods in urban districts have tight restrictions on where distribution lines are placed and various buildings are serviced.

In addition, most transmission publications focused on North America, whereas European datasets are commonly used for distribution due to two modeling reasons: (i) various European initiatives motivated research groups to model LV and MV grids for economic and DER integration studies, and (ii) the consideration of phase unbalance poses a relative difficulty for modeling North American grids. This choice affected the datasets used for distribution networks, such as those from China, Italy or the Netherlands (refer to the table footnote). While some are available for use, e.g. ENTSO-e [140] and RTE [141], utility datasets tend to be confidential and only their statistics can be retrieved from publications, e.g., histograms of various metrics in [17] and [34].

Moving onto the characteristics, almost all references accounted for the graph topology. Those which only focused on it, e.g., [35], [47], have no application entries. Otherwise,

references with system and/or operational characteristics considered some applications. Power quality is the most common for generating or validating synthetic benchmarks (refer to Section III-C) since it relates directly to guidelines followed by utilities, e.g. voltage and power flow limits. While other aspects are considered as constraints in the generation process, e.g. reliability indices in [50], some works prefer appending them on existing benchmarks, e.g. HIL testing for the CIGRÉ MV system [69].

From the model types, heuristic algorithms and statistical sampling are common in generating synthetic networks (refer to Section V-C). Expert design was excluded from Table 4 due to its manual approach. Only two references applied deep graph generation: [131] for transmission and [125] for distribution systems. This research gap suggests deep graph learning could be explored and implemented in the coming years, especially as it continuously learns through examples based on a data-driven process (refer to Section VI). In terms of modeling layers, transmission works typically ignore natural or manmade boundaries that are otherwise core constraints in distribution networks. They instead consider siting synthetic substations and minimizing line lengths based on geometry. Other layers such as transient analysis [51] and economic study [62] are also included in some methodologies.

Synthetic benchmarks can be validated in different ways; most publications employ aggregate statistics, e.g. mean, median or ranges, to compare synthetic data to actual grids. These statistics use topological, system and/or operational metrics introduced in Section III. Beyond this level of abstraction, other works analyze empirical distributions by visual verification or checking their statistical similarity. This latter approach fits a statistical model on generated data and compares the parameter values, e.g. mean and variance. A more rigorous approach computes the goodness of fit through statistical tests, such as the KL divergence used in a few publications (mostly transmission). This quantitative measure does

not suffer from comparing single points, but rather considers entire ranges of two distributions for a more comprehensive judgment. Except for [19], most distribution publications lack the latter two approaches, mainly due to confidentiality concerns of private datasets and the relative simplicity of using aggregate statistics.

Every publication in Table 4 results in a synthetic benchmark dataset and/or a generation tool, which may be open-source or licensed. The most inclusive references in terms of modeling layers such as [49] are widely adopted and used in various research projects, e.g. SMART-DS at NREL [105].

B. PUBLISHED DATASETS

Upon running a generative model, synthetic benchmarks can be created and released for use. Several initiatives pushed to document and share available datasets. For example, the IEEE PES PGLib-OPF Task Force [142] compiled benchmarks for testing optimal power flow algorithms. Through the ARPA-E GRID DATA program, two online repositories, BetterGrids [143] and DR POWER [138], were established to publish and find open-access grid datasets in various file formats. Each representative network has different characteristics that must be considered before a power systems study.

Table 6 summarizes published datasets of synthetic benchmarks available through online repositories. Datasets exclusively reported in publications are not listed since they are not easy to replicate for large networks and components. Some standard datasets (e.g. IEEE, CIGRÉ and EPRI) are used to compare and highlight their differences with synthetic ones. Each row corresponds to a dataset, with common ones grouped, while the columns demonstrate various characteristics explained as follows.

First, a dataset can represent a specific location and either a transmission or distribution system (except for the last one, to be discussed later). For example, ACTIVSg10k and GenWI represent synthetic transmission systems based on the Western U.S. and are validated with the WI dataset.

Second, the number of assets concerns the feeders, buses, lines, transformers, loads, generators, and DERs. The number of feeders is irrelevant for transmission, while there are no conventional generation units in most distribution networks.

Third, applications correspond to potential use cases in a power system study. For example, the ACTIVSg2000 dataset includes sufficient information to simulate steady-state power flows (Power Quality), perform contingency analysis (Operation and Reliability), and consider system dynamics (Protection and Control). By identifying the required application, one can select an appropriate dataset for future studies.

Finally, the modeled aspects in these benchmarks are identified, i.e. geographical mapping of system components, time series load profiles, inclusion of DERs and/or renewable energy sources (RESSs), and evolution scenarios accounting for increasing demand and penetration levels. The last column lists the corresponding model and repository references.

The first two blocks are available at the Electric Grid Test Case Repository [139] and these HV transmission datasets

operate beyond 1 kV, range from 200 to 70,000 buses across different U.S. states, with most of them relevant to all three applications. Their average degrees are between 2–3, matching the typical values discussed in Section III-A. They all include geographical bus siting and bulk renewable generation, while synthetic load profiles were created based on [99]. Other datasets not shown in Table 6 are the Illini 42 Tornado and HEMP, which consider grid disturbances due to natural disasters. GenWI, similar to ACTIVSg10k, accounts for only power quality and geographical aspects but not transformers.

The PNNL SDET datasets mostly consider power quality applications and ignore other aspects; their lack of representative locations is due to the generative method, i.e. creating and assembling anonymized sub-networks based on existing datasets (refer to Section V-B). Non-U.S. transmission datasets are also given, from which SimBench-HV [101] based on the German grid includes the most detail and simulation scenarios; RE-Europe [76] and the Australian egrimod-NEM [144] do not model transformers, yet include time series data at sub-hourly intervals over multiple years.

Addressing distribution systems, the PNNL taxonomy [17] represents five U.S. climate regions and 24 prototypical feeders (one is general and not associated with any region). These were all modeled in GridLAB-D and consider steady-state power flows across urban, rural and suburban environments. Perhaps the most comprehensive dataset are those developed by NREL in collaboration with IIT Comillas. These large-scale distribution networks representing the urban and rural areas of Santa Fe (New Mexico), Greensboro (North Carolina), and San Francisco (California) were generated through the RNM-US tool. Their outcomes were GIS models of system components and OpenDSS models that can simulate each or a combination of feeders. NREL-SFO is the largest modeled distribution network with over 2 million customers and 116,837 kms of power lines while considering phase unbalance in a North American grid. Some datasets also include contingency analysis and load evolution scenarios. However, the lack of time series and DER data requires adapting these datasets for some studies.

SimBench-MV and LV represent German distribution networks at different voltage levels while considering all modeling aspects. The available datasets, generated through the methodology proposed in [101], are available as CSV files and modeled in Python through the pandapower library. The MV networks have ring structures, while the LV ones are radial. Load and RES profiles at 15-minute intervals for an entire year are provided for each scenario.

Some datasets of EPRI [145], IEEE [6], [9], and CIGRÉ [11] are listed to compare with the synthetic ones and are not meant to be comprehensive. For example, the IEEE 9, 14, 30, 39, 57, 118, 300-bus and other test systems are discussed more in [12]. Except for a few novel cases, these benchmarks are not all representative of a given location with a lack of geographical information, evolution scenarios, and modern grid considerations. Their network sizes are relatively smaller than the synthetic ones which come in a

TABLE 6. Published datasets and characteristics of synthetic benchmark networks.

System	Dataset Description		Number of Assets							Application			Modeling Aspects				References
	Name	Location	#Feeders	#Buses	#Lines	#Transformers	#Loads	#Generators	#DERs	Power Quality	Oper. & Reliab.	Prot. & Control	GIS Mapping	Load Profiles	DER / RES	Evol. Scenarios	
Transmission	UIUC150	TN, U.S.	N/A	150	217	60	90	27		✓			✓		✓		[22], [139]
	ACTIVSg200	IL, U.S.	N/A	200	245	66	160	49		✓	✓	✓	✓	✓	✓		[53], [139]
	ACTIVSg500	SC, U.S.	N/A	500	597	131	206	90		✓	✓	✓	✓	✓	✓		[53], [139]
	ACTIVSg2000	TX, U.S.	N/A	2k	3.2k	861	1.3k	544		✓	✓	✓	✓	✓	✓		[53], [139]
	ACTIVSg10k	West U.S.	N/A	10k	12.7k	2.4k	4.8k	2.4k		✓	✓	✓	✓	✓	✓		[53], [139]
	GenWI	West U.S.	N/A	14.4k	18.8k		3k	3k		✓			✓				[33], [138]
	ACTIVSg25k	N.E. U.S.	N/A	25k	32.2k	6.2k	9.4k	4.8k		✓	✓	✓	✓	✓	✓		[53], [139]
	ACTIVSg70k	East U.S.	N/A	70k	88.2k	13.1k	38.1k	10.3k		✓	✓	✓	✓	✓	✓		[53], [139]
	SDET-563	N/A	N/A	563	571	89	442	140		✓				✓			[94], [138]
	SDET-793	N/A	N/A	793	769	143	568	210		✓							[94], [138]
	SDET-2316	N/A	N/A	2.3k	2.1k	860	1.5k	604		✓							[94], [138]
	SDET-4661	N/A	N/A	4.6K	4.6k	1.3k	3k	1.1k		✓							[94], [138]
	SimBench-HV	Germany	N/A	527	1k	217	527	345		✓	✓		✓	✓	✓	✓	[101], [146]
	egrimod-NEM	Australia	N/A	912	1.4k		264	203		✓	✓		✓	✓	✓		[144], [147]
	RE-Europe	Europe	N/A	1.4k	2.1k		1.5k	969		✓	✓		✓	✓	✓		[76], [148]
Distribution	PNNL-R1	West U.S.	5	1.6k	1.9k	879	156			✓							[17], [138]
	PNNL-R2	North U.S.	5	2.9k	3.5k	1.7k	620			✓							[17], [138]
	PNNL-R3	S.W. U.S.	3	2.9k	4.6k	2.2k	234			✓							[17], [138]
	PNNL-R4	C.E. U.S.	3	1k	1.7k	889	97			✓							[17], [138]
	PNNL-R5	S.E. U.S.	7	5.1k	6.9k	3k	397			✓							[17], [138]
	NREL-SAF	NM, U.S.	28	89k	77.7k	11.3k	38.5k			✓	✓		✓				[49], [143]
	NREL-GSO	NC, U.S.	98	182k	156k	25.9k	70.5k			✓			✓			✓	[49], [143]
	NREL-SFO	CA, U.S.	2k	5M	4.4M	560k	2M			✓			✓				[49], [143]
	SimBench-LV	Germany	33	174	175	6	486		44	✓			✓	✓	✓	✓	[101], [146]
	SimBench-MV	Germany	40	465	476	8	448		450	✓			✓	✓	✓	✓	[101], [146]
	EPRI-M1	N/A	1	505	440	159	1.5k			✓			✓	✓			[13], [145]
	EPRI-K1	S.E. U.S.	1	650	646	323	321			✓			✓	✓			[13], [145]
	EPRI-J1	N.E. U.S.	1	1.2k	1.2k	819	1.3k		4	✓			✓	✓	✓		[13], [145]
	IEEE-342 LV	N/A	1	390	390	70	624			✓							[9], [10]
	European LV	U.K.	1	906	905	1	55			✓			✓	✓			[9], [10]
	IEEE-8500	N/A	1	4.8k	2.5	1.1k	3.5k			✓			✓				[9], [10]
	CIGRÉ-LV	U.S. / EU	3	32	26	3	19			✓		✓		✓	✓		[11], [69]
CIGRÉ-MV	U.S. / EU	2	87	81	2	86			✓		✓		✓	✓		[11]	
TDgrid-v03	TX, U.S.	448	856k	739k	132k	307k		39	✓			✓		✓	✓	[107], [139]	

C.E.: Central East. N.E.: Northeast. S.E.: Southeast. S.W.: Southwest. k: thousand. M: million.

variety of ranges. Extensions were created to address different potential uses, e.g. HIL benchmark for DERs [69] based on CIGRÉ-MV.

Last but not the least, a coupled transmission–distribution dataset was recently created and released [107]. It combines the works of [49], [53], [107] and represents 448 feeders in the Austin, Texas metropolitan area. While the transmission system was modeled using PowerWorld and the distribution through OpenDSS, the two entities operating on different voltage levels were linked through the Hierarchical Engine for Large-scale Infrastructure Co-Simulation (HELICS) platform [149]. Future scenarios including economic and renewable growth were also considered.

IX. CONCLUSION

The process of creating a synthetic benchmark must adhere to three attributes, namely: (i) representativeness, (ii) confidentiality, and (iii) modeling. Being representative is considered in the problem specification and network validation stages, confidentiality is accounted for in the data collection and pre-processing, and the latter is manifested in the

model development and network generation. These attributes, however, are intertwined and can appear in other stages as well.

The general procedure for creating synthetic benchmarks is systematic rather than linear, where the starting steps can be revisited. Initially, this problem is shaped by the given specifications, which are defined with a two-dimensional set of metrics. The first dimension is the class of power system characteristics: (i) topological, (ii) system, and (iii) operational, while the second dimension is the referenced use of characteristics: (i) input, (ii) constraint, (iii) output, and (iv) validation. Only a subset of these metrics is typically selected based on the intended application of the generated synthetic network. Next, the collection and pre-processing of relevant datasets are defined as requirements, i.e. input or constraint, in the problem specification and are paramount for the quality of a generated benchmark.

Four types of generative models were also presented: (i) expert design, (ii) anonymized clustering, (iii) statistics-heuristic hybrid generation tools, and (iv) deep graph generation. This survey revealed that while, historically, power

systems benchmarks relied on expert design, current research trends tend to favor generation tools. Nevertheless, the lack of autonomous generative models was highlighted as a research gap. Deep graph generative models, which consist of four different learning approaches, were suggested as strong candidates to fill this gap based on their performance in other disciplines, similar to the power system problem.

Using the metrics defined in the specifications stage, validation approaches for benchmarks were explained. A comparative summary of the various works on generating and validating synthetic benchmarks was presented, along with a list of their published datasets and characteristics.

ACKNOWLEDGMENT

The authors would like to thank Steven Wong for his technical and linguistic review of this survey article.

REFERENCES

- [1] IEA, "World energy outlook 2020," Int. Energy Agency, Paris, France, Tech. Rep., 2020.
- [2] IEA, "Introduction to system integration of renewables," Int. Energy Agency, Paris, France, Tech. Rep., 2020.
- [3] IRENA, "Towards 100% renewable energy: Utilities in transition," Int. Renew. Energy Agency, Abu Dhabi, United Arab Emirates, Tech. Rep., 2020.
- [4] IRENA, "Market integration of distributed energy resources: Innovation landscape brief," Int. Renew. Energy Agency, Abu Dhabi, United Arab Emirates, Tech. Rep., 2019.
- [5] NERC, "Distributed energy resources: Connection modeling and reliability considerations," Distributed Energy Resour. Task Force Rep., North American Electr. Rel. Corp., Atlanta, GA, USA, Tech. Rep., Feb. 2017.
- [6] IEEE PES. (2020). *IEEE Test Cases*. PSACE Test Case Coordination Working Group. [Online]. Available: <https://site.ieee.org/pestccwg/links-to-test-cases/>
- [7] C. Canizares, T. Fernandes, E. Galdi, Jr., L. Gérin-Lajoie, M. Gibbard, I. Hiskens, J. Kersulis, R. Kuiava, L. Lima, F. Marco, and N. Martins, "Benchmark systems for small-signal stability analysis and control," IEEE PES Task Force Benchmark Syst. Stability Controls, Power Syst. Dyn. Perform. Committee, New York, NY, USA, Tech. Rep. PES-TR18, Aug. 2015.
- [8] "Benchmarking of power quality performance in transmission systems," CIGRÉ Working Group C4.27, CIGRÉ, Paris, France, Tech. Rep., 2018.
- [9] K. P. Schneider, B. A. Mather, B. C. Pal, C. W. Ten, G. J. Shirek, H. Zhu, J. C. Fuller, J. L. R. Pereira, L. F. Ochoa, L. R. de Araujo, R. C. Dugan, S. Matthias, S. Paudyal, T. E. McDermott, and W. Kersting, "Analytic considerations and design basis for the IEEE distribution test feeders," *IEEE Trans. Power Syst.*, vol. 33, no. 3, pp. 3181–3188, May 2018.
- [10] CIGRÉ, "Benchmarking of power quality performance in transmission systems," CIGRÉ Working Group C4.27, Paris, France, Tech. Rep., 2018.
- [11] CIGRÉ, "Benchmark systems for network integration of renewable and distributed energy resources," CIGRE Task Force C6.04.02, Paris, France, Tech. Rep., Apr. 2014.
- [12] S. Peyghami, P. Davari, M. Fotuhi-Firuzabad, and F. Blaabjerg, "Standard test systems for modern power system analysis: An overview," *IEEE Ind. Electron. Mag.*, vol. 13, no. 4, pp. 86–105, Dec. 2019.
- [13] F. P. Marcos, C. M. Domingo, T. G. S. Román, B. Palmintier, B.-M. Hodge, V. Krishnan, F. de Cuadra García, and B. Mather, "A review of power distribution test feeders in the United States and the need for synthetic representative networks," *Energies*, vol. 10, no. 11, p. 1896, Nov. 2017.
- [14] M. N. Alam, S. Chakrabarti, and X. Liang, "A benchmark test system for networked microgrids," *IEEE Trans. Ind. Inform.*, vol. 16, no. 10, pp. 6217–6230, Oct. 2020.
- [15] S. A. Raza and J. Jiang, "A benchmark distribution system for investigation of residential microgrids with multiple local generation and storage devices," *IEEE Open Access J. Power Energy*, vol. 7, pp. 41–50, 2020.
- [16] H. Farhangi and G. Joos, "Microgrid benchmarks," in *Microgrid Planning and Design: A Concise Guide*. Hoboken, NJ, USA: Wiley, Apr. 2019, ch. 2, pp. 25–36.
- [17] K. P. Schneider, Y. Chen, D. P. Chassin, R. G. Pratt, D. W. Engel, and S. E. Thompson, "Modern grid initiative distribution taxonomy final report," Pacific Northwest Nat. Lab., Richland, WA, USA, Tech. Rep. PNNL-18035, Nov. 2008, Art. no. 1040684.
- [18] K. P. Schneider, Y. Chen, D. Engle, and D. Chassin, "A taxonomy of north American radial distribution feeders," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, Jul. 2009, pp. 1–6, doi: [10.1109/PES.2009.5275900](https://doi.org/10.1109/PES.2009.5275900).
- [19] E. Schweitzer, A. Scaglione, A. Monti, and G. A. Pagani, "Automated generation algorithm for synthetic medium voltage radial distribution systems," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 7, no. 2, pp. 271–284, Jun. 2017.
- [20] R. Espejo, S. Lumberas, and A. Ramos, "A complex-network approach to the generation of synthetic power transmission networks," *IEEE Syst. J.*, vol. 13, no. 3, pp. 3050–3058, Sep. 2019.
- [21] Z. Wang, A. Scaglione, and R. J. Thomas, "Generating statistically correct random topologies for testing smart grid communication and control networks," *IEEE Trans. Smart Grid*, vol. 1, no. 1, pp. 28–39, Jun. 2010.
- [22] A. B. Birchfield, K. M. Gegner, T. Xu, K. S. Shetye, and T. J. Overbye, "Statistical considerations in the creation of realistic synthetic power grids for geomagnetic disturbance studies," *IEEE Trans. Power Syst.*, vol. 32, no. 2, pp. 1502–1510, Mar. 2017.
- [23] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393, no. 6684, pp. 440–442, 1998.
- [24] R. Albert and A. Barabási, "Statistical mechanics of complex networks," *Rev. Modern Phys.*, vol. 74, no. 1, pp. 47–97, Jan. 2002.
- [25] V. Rosato, S. Bologna, and F. Tiriticco, "Topological properties of high-voltage electrical transmission networks," *Electr. Power Syst. Res.*, vol. 77, no. 2, pp. 99–105, 2007.
- [26] Z. Wang, A. Scaglione, and R. J. Thomas, "The node degree distribution in power grid and its topology robustness under random and selective node removals," in *Proc. IEEE Int. Conf. Commun. Workshops*, May 2010, pp. 1–5, doi: [10.1109/ICCW.2010.5503926](https://doi.org/10.1109/ICCW.2010.5503926).
- [27] P. A. Rikvold, I. A. Hamad, B. Israels, and S. V. Poroseva, "Modeling power grids," *Phys. Proc.*, vol. 34, pp. 119–123, Jan. 2012.
- [28] G. A. Pagani and M. Aiello, "The power grid as a complex network: A survey," *Phys. A, Statist. Mech. Appl.*, vol. 392, no. 11, pp. 2688–2700, 2013.
- [29] B. Cloteaux, "Limits in modeling power grid topology," in *Proc. IEEE 2nd Netw. Sci. Workshop (NSW)*, Apr. 2013, pp. 16–22.
- [30] S. Abeysinghe, J. Wu, M. Sooriyabandara, M. Abeyskera, T. Xu, and C. Wang, "Topological properties of medium voltage electricity distribution networks," *Appl. Energy*, vol. 210, pp. 1101–1112, Jan. 2018.
- [31] J. Hu, L. Sankar, and D. J. Mir, "Cluster-and-connect: An algorithmic approach to generating synthetic electric power network graphs," in *Proc. 53rd Annu. Allerton Conf. Commun., Control, Comput. (Allerton)*, Sep. 2015, pp. 223–230, doi: [10.1109/ALLERTON.2015.7447008](https://doi.org/10.1109/ALLERTON.2015.7447008).
- [32] E. Cotilla-Sanchez, P. D. H. Hines, C. Barrows, and S. Blumsack, "Comparing the topological and electrical structure of the North American electric power infrastructure," *IEEE Syst. J.*, vol. 6, no. 4, pp. 616–626, Dec. 2012.
- [33] S. Soltan and G. Zussman, "Generation of synthetic spatially embedded power grid networks," in *Proc. IEEE Power Energy Soc. Gen. Meeting (PESGM)*, Jul. 2016, pp. 1–5, doi: [10.1109/PESGM.2016.7741383](https://doi.org/10.1109/PESGM.2016.7741383).
- [34] V. Krishnan, B. Bugbee, T. Elgindy, C. Mateo, P. Duenas, F. Postigo, J.-S. Lacroix, T. G. S. Roman, and B. Palmintier, "Validation of synthetic U.S. electric power distribution system data sets," *IEEE Trans. Smart Grid*, vol. 11, no. 5, pp. 4477–4489, Sep. 2020.
- [35] S. G. Aksoy, E. Purvine, E. Cotilla-Sanchez, and M. Halappanavar, "A generative graph model for electrical infrastructure networks," *J. Complex Netw.*, vol. 7, no. 1, pp. 128–162, Aug. 2018.
- [36] A. Trpovski, D. Recalde, and T. Hamacher, "Synthetic distribution grid generation using power system planning: Case study of Singapore," in *Proc. 53rd Int. Universities Power Eng. Conf. (UPEC)*, Sep. 2018, pp. 1–6, doi: [10.1109/UPEC.2018.8542054](https://doi.org/10.1109/UPEC.2018.8542054).
- [37] S. J. Young, Y. Makarov, R. Diao, M. Halappanavar, M. Vallem, R. Fan, R. Huang, J. OrBrien, and Z. H. Huang, "Topological power grid statistics from a network-of-networks perspective," in *Proc. IEEE Power Energy Soc. Gen. Meeting (PESGM)*, Aug. 2018, pp. 1–5, doi: [10.1109/PESGM.2018.8586475](https://doi.org/10.1109/PESGM.2018.8586475).
- [38] M. Ouyang, Z. Pan, L. Hong, and L. Zhao, "Correlation analysis of different vulnerability metrics on power grids," *Phys. A, Stat. Mech. Appl.*, vol. 396, pp. 204–211, Feb. 2014.

- [39] E. Schweitzer, "Creating, validating, and using synthetic power flow cases: A statistical approach to power system analysis," Ph.D. dissertation, School Electr., Comput. Energy Eng., Arizona State Univ., Tempe, AZ, USA, 2019.
- [40] G. A. Pagani and M. Aiello, "Towards decentralization: A topological investigation of the medium and low voltage grids," *IEEE Trans. Smart Grid*, vol. 2, no. 3, pp. 538–547, Sep. 2011.
- [41] P. Hines, E. Cotilla-Sanchez, and S. Blumsack, "Do topological models provide good information about electricity infrastructure vulnerability?" *Chaos, Interdiscipl. J. Nonlinear Sci.*, vol. 20, no. 3, Sep. 2010, Art. no. 033122.
- [42] Z. Wang, M. H. Athari, and S. Hamid Elyas, "Statistically analyzing power system network," in *Proc. IEEE Power Energy Soc. Gen. Meeting (PESGM)*, Aug. 2018, pp. 1–5, doi: [10.1109/PESGM.2018.8586110](https://doi.org/10.1109/PESGM.2018.8586110).
- [43] E. Bompard, E. Pons, and D. Wu, "Extended topological metrics for the analysis of power grid vulnerability," *IEEE Syst. J.*, vol. 6, no. 3, pp. 481–487, Sep. 2012.
- [44] P. Hines and S. Blumsack, "A centrality measure for electrical networks," in *Proc. 41st Annu. Hawaii Int. Conf. Syst. Sci.*, Jan. 2008, p. 185.
- [45] Z. Wang, A. Scaglione, and R. J. Thomas, "Electrical centrality measures for electric power grid vulnerability analysis," in *Proc. 49th IEEE Conf. Decis. Control (CDC)*, Atlanta, GA, USA, Dec. 2010, pp. 5792–5797, doi: [10.1109/CDC.2010.5717964](https://doi.org/10.1109/CDC.2010.5717964).
- [46] Z. Wang, S. H. Elyas, and R. J. Thomas, "A novel measure to characterize bus type assignments of realistic power grids," in *Proc. IEEE Eindhoven PowerTech*, Jun. 2015, pp. 5792–5797, doi: [10.1109/PTC.2015.7232436](https://doi.org/10.1109/PTC.2015.7232436).
- [47] S. H. Elyas and Z. Wang, "Improved synthetic power grid modeling with correlated bus type assignments," *IEEE Trans. Power Syst.*, vol. 32, no. 5, pp. 3391–3402, Sep. 2017.
- [48] E. Schweitzer, A. Scaglione, and R. Thomas, "The validation of synthetic power system cases," in *Proc. IREP Symp.*, Sep. 2017, pp. 1–7.
- [49] C. Mateo, F. Postigo, F. de Cuadra, T. G. S. Roman, T. Elgindy, P. Duenas, B.-M. Hodge, V. Krishnan, and B. Palmintier, "Building large-scale U.S. synthetic electric distribution system models," *IEEE Trans. Smart Grid*, vol. 11, no. 6, pp. 5301–5313, Nov. 2020.
- [50] C. M. Domingo, T. G. S. Roman, Á Sanchez-Mirallas, J. P. P. Gonzalez, and A. C. Martinez, "A reference network model for large-scale distribution planning with automatic street map generation," *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 190–197, Feb. 2011.
- [51] G. Pisano, N. Chowdhury, M. Coppo, N. Natale, G. Petretto, G. G. Soma, R. Turri, and F. Pilo, "Synthetic models of distribution networks based on open data and georeferenced information," *Energies*, vol. 12, no. 23, p. 4500, Nov. 2019.
- [52] J. Kays, A. Seack, T. Smirek, F. Westkamp, and C. Rehtanz, "The generation of distribution grid models on the basis of public available data," *IEEE Trans. Power Syst.*, vol. 32, no. 3, pp. 2346–2353, May 2017.
- [53] A. B. Birchfield, T. Xu, K. M. Gegner, K. S. Shetye, and T. J. Overbye, "Grid structural characteristics as validation criteria for synthetic networks," *IEEE Trans. Power Syst.*, vol. 32, no. 4, pp. 3258–3265, Jul. 2017.
- [54] G. Prettico, F. Gangale, A. Mengolini, A. Lucas, and G. Fulli, "Distribution system operators observatory 2018," Publ. Office Eur. Union, Luxembourg, Luxembourg, Tech. Rep. EUR 29615 EN, 2019.
- [55] A. Birchfield, E. Schweitzer, M. Athari, T. Xu, T. Overbye, A. Scaglione, and Z. Wang, "A metric-based validation process to assess the realism of synthetic power grids," *Energies*, vol. 10, no. 8, p. 1233, Aug. 2017.
- [56] A. B. Birchfield, T. Xu, and T. J. Overbye, "Power flow convergence and reactive power planning in the creation of large synthetic grids," *IEEE Trans. Power Syst.*, vol. 33, no. 6, pp. 6667–6674, Nov. 2018.
- [57] C. Mateo, G. Prettico, T. Gómez, R. Cossent, F. Gangale, P. Frías, and G. Fulli, "European representative electricity distribution networks," *Int. J. Electr. Power Energy Syst.*, vol. 99, pp. 273–280, Jul. 2018.
- [58] J. Amme, G. Pleßmann, J. Bühler, L. Hülk, E. Kötter, and P. Schwaegerl, "The eGo grid model: An open-source and open-data based synthetic medium-voltage grid model for distribution power supply systems," *J. Phys., Conf. Ser.*, vol. 977, Feb. 2018, Art. no. 012007.
- [59] M. H. Athari and Z. Wang, "Introducing voltage-level dependent parameters to synthetic grid electrical topology," *IEEE Trans. Smart Grid*, vol. 10, no. 4, pp. 4048–4056, Jul. 2019.
- [60] S. Meinecke, N. Bornhorst, and M. Braun, "Power system benchmark generation methodology," in *Proc. Conf. Sustain. Energy Supply Energy Storage Syst. (NEIS)*, Hamburg, Germany: VDE, Sep. 2018, pp. 1–6.
- [61] A. K. Jain and B. Mather, "Clustering methods and validation of representative distribution feeders," in *Proc. IEEE/PES Transmiss. Distrib. Conf. Expo. (T&D)*, Apr. 2018, pp. 1–9.
- [62] E. Schweitzer and A. Scaglione, "A mathematical programming solution for automatic generation of synthetic power flow cases," *IEEE Trans. Power Syst.*, vol. 34, no. 1, pp. 729–741, Jan. 2019.
- [63] E. Schweitzer, T. Xu, A. B. Birchfield, A. Scaglione, T. J. Overbye, R. J. Thomas, and Z. Wang, "Towards operational validation: Mapping power system inputs to operating conditions," in *Proc. Power Syst. Comput. Conf. (PSCC)*, Jun. 2018, pp. 1–7, doi: [10.23919/PSCC.2018.8442880](https://doi.org/10.23919/PSCC.2018.8442880).
- [64] S. Soltan, A. Loh, and G. Zussman, "A learning-based method for generating synthetic power grids," *IEEE Syst. J.*, vol. 13, no. 1, pp. 625–634, Mar. 2019.
- [65] T. Gomez, C. Mateo, A. Sanchez, P. Frias, and R. Cossent, "Reference network models: A computational tool for planning and designing large-scale smart electricity distribution grids," in *High Performance Computing in Power and Energy Systems (Power Systems)*. Berlin, Germany: Springer, 2013, pp. 247–279.
- [66] F. Blaabjerg, R. Teodorescu, M. Liserre, and A. V. Timbus, "Overview of control and grid synchronization for distributed power generation systems," *IEEE Trans. Ind. Electron.*, vol. 53, no. 5, pp. 1398–1409, Oct. 2006.
- [67] P. Mackin, R. Daschmans, B. Williams, B. Haney, R. Hunt, and J. Ellis, "Dynamic simulation studies of the frequency response of the three U.S. interconnections with increased wind generation," Ernest Orlando Lawrence Berkeley Nat. Lab., Berkeley, CA, Tech. Rep. LBNL-4146E, Dec. 2010.
- [68] T. Xu, A. B. Birchfield, and T. J. Overbye, "Modeling, tuning, and validating system dynamics in synthetic electric grids," *IEEE Trans. Power Syst.*, vol. 33, no. 6, pp. 6501–6509, Nov. 2018.
- [69] P. Kotsampopoulos, D. Lagos, N. Hatziaargyriou, M. O. Faruque, G. Lauss, O. Nzimako, P. Forsyth, M. Steurer, F. Ponci, A. Monti, V. Dinavahi, and K. Strunz, "A benchmark system for hardware-in-the-loop testing of distributed energy resources," *IEEE Power Energy Technol. Syst. J.*, vol. 5, no. 3, pp. 94–103, Sep. 2018.
- [70] M. Newman, *Networks*, 2nd ed. London, U.K.: Oxford Univ. Press, 2018.
- [71] D. B. West, *Introduction to Graph Theory*, 2nd ed. Upper Saddle River, NJ: Prentice-Hall, 1996.
- [72] F. Berger, P. Gritzmann, and S. de Vries, "Minimum cycle bases for network graphs," *Algorithmica*, vol. 40, no. 1, pp. 51–62, May 2004.
- [73] M. Fiedler, "Laplacian of graphs and algebraic connectivity," in *Combinatorics Graph Theory*, vol. 25. Warsaw, Poland: Banach Center Publ., 1989, pp. 57–70.
- [74] W. H. Kersting, *Distribution System Modeling and Analysis*, 3rd ed. Boca Raton, FL, USA: CRC Press, 2012.
- [75] U.S. Energy Information Association. (Sep. 2020). *Form EIA-860 Detailed Data With Previous Form Data (EIA-860A/860B)*. [Online]. Available: <https://www.eia.gov/electricity/data/eia860/>
- [76] T. V. Jensen and P. Pinson, "RE-Europe, a large-scale dataset for modeling a highly renewable European electricity system," *Sci. Data*, vol. 4, no. 1, Dec. 2017, Art. no. 170175.
- [77] J. D. Glover, T. Overbye, and M. S. Sarma, *Power System Analysis & Design*, 6th ed. Boston, MA, USA: Cengage Learning, 2016.
- [78] AEMO, "Power system requirements," Australian Energy Market Operator, Melbourne, VIC, Australia, Tech. Rep., Jul. 2020.
- [79] "Procedures for validation of powerflow and dynamics cases," NERC Model Validation Work. Group, North Amer. Electr. Reliability Corp., Atlanta, GA, USA, Tech. Rep., Dec. 2011.
- [80] *Electric Power Systems and Equipment Voltage Ratings (60 Hz)*, document ANSI C84.1-2020, 2020.
- [81] T. A. Short, *Electric Power Distribution Handbook*, 2nd ed. Boca Raton, FL, USA: CRC Press, 2014.
- [82] R. Dugan, M. F. McGranaghan, S. Santoso, and H. W. Beaty, *Electrical Power Systems Quality*, 3rd ed. New York, NY, USA: McGraw-Hill, 2012.
- [83] *IEEE Recommended Practice and Requirements for Harmonic Control in Electric Power Systems*, Standard Std 519-2014, 2014.
- [84] *Photovoltaic (PV) Systems Characteristics of the Utility Interface*, document IEC 61727, 2004.
- [85] A. J. Wood, B. F. Wollenberg, and G. B. Sheblé, *Power Generation, Operation, and Control*, 3rd ed. Hoboken, NJ, USA: Wiley, 2013.
- [86] *IEEE Guide for Electric Power Distribution Reliability Indices*, Standard 1366-2012, 2012.

- [87] A. Gómez-Expósito, A. J. Conejo, and C. Cañizares, *Electric Energy Systems: Analysis and Operation*, 2nd ed. Boca Raton, FL, USA: CRC Press, 2018.
- [88] *IEEE Standard for Interconnection and Interoperability of Distributed Energy Resources with Associated Electric Power Systems Interfaces*, Standard 1547-2018, 2018.
- [89] *Recommended Practice for Calculating AC Short-Circuit Currents in Industrial and Commercial Power Systems*, Standard 551-2006, 2006.
- [90] *Power Transformers—Part 5: Ability to Withstand Short Circuit*, document IEC 60076-5, 2006.
- [91] C. D. Vournas, P. W. Sauer, and M. A. Pai, "Relationships between voltage and angle stability of power systems," *Int. J. Electr. Power Energy Syst.*, vol. 18, no. 8, pp. 493–500, Nov. 1996.
- [92] R. Huang, Y. V. Makarov, Z. Huang, S. J. Young, R. Fan, A. Tbaileh, Z. Hou, J. O'Brien, J. C. Fuller, J. Hansen, and L. D. Marinovici, "Final scientific/technical report sustainable data evolution technology (SDET) for power grid optimization," Pacific Northwest Nat. Lab., Richland, WA, USA, Tech. Rep. PNNL-28338, 1524091, Dec. 2018.
- [93] S. Young, Y. Makarov, R. Diaio, R. Fan, R. Huang, J. O'Brien, M. Halappanavar, M. Vallem, and Z. Huang, "Synthetic power grids from real world models," in *Proc. IEEE Power Energy Soc. Gen. Meet. (PESGM)*, Piscataway, NJ, USA, Aug. 2018, doi: 10.1109/PESGM.2018.8585792.
- [94] R. Huang. (2019). *SDET Topology Generation Tool*. [Online]. Available: <https://eGRIDdata.org/dataset/sdet-topology-generation-tool>
- [95] E. J. Kirk. (2012). *Count Loops in a Graph*. [Online]. Available: <https://www.mathworks.com/matlabcentral/fileexchange/10722>
- [96] J. Dickert, M. Domagk, and P. Schegner, "Benchmark low voltage distribution networks based on cluster analysis of actual grid properties," in *Proc. IEEE Grenoble Conf.*, Jun. 2013, pp. 1–6, doi: 10.1109/PTC.2013.6652250.
- [97] E. Perim. (2021). *SyntheticGrids.Jl*. [Online]. Available: <https://invenia.github.io/SyntheticGrids.jl/latest/>
- [98] Z. Wang, H. Sadeghian, S. H. Elyas, R. D. Zimmerman, E. Schweitzer, and A. Scaglione. (Jun. 2019). *SynGrid User's Manual*. [Online]. Available: <https://matpower.org/docs/SynGrid-manual.pdf>
- [99] H. Li, J. H. Yeo, A. L. Bornsheuer, and T. J. Overbye, "The creation and validation of load time series for synthetic electric power systems," *IEEE Trans. Power Syst.*, vol. 36, no. 2, pp. 961–969, Mar. 2021.
- [100] J. Amme. (Sep. 2019). *Distribution Network Generator (DINGO)*. [Online]. Available: <https://dingo.readthedocs.io/en/dev/>
- [101] S. Meinecke, D. Sarajlić, S. R. Drauz, A. Klettke, L.-P. Lauven, C. Rehtanz, A. Moser, and M. Braun, "SimBench—A benchmark dataset of electric power systems to compare innovative solutions based on power flow analysis," *Energies*, vol. 13, no. 12, p. 3290, Jun. 2020.
- [102] C. M. Domingo. (2021). *RNM: Reference Network Model*. [Online]. Available: <http://www.iit.comillas.edu/technology-offer/rnm>
- [103] E. Miguez, J. Cidras, E. Diaz-Dorado, and J. L. Garcia-Dornelas, "An improved branch-exchange algorithm for large-scale distribution network planning," *IEEE Trans. Power Syst.*, vol. 17, no. 4, pp. 931–936, Nov. 2002.
- [104] M. Grzanic, M. G. Flammini, and G. Pretticco, "Distribution network model platform: A first case study," *Energies*, vol. 12, no. 21, p. 4079, Oct. 2019.
- [105] B. Palmintier, T. Elgindy, C. Mateo, F. Postigo, T. Gómez, F. de Cuadra, and P. D. Martinez, "Experiences developing large-scale synthetic U.S.-style distribution test systems," *Electr. Power Syst. Res.*, vol. 190, Jan. 2021, Art. no. 106665.
- [106] F. Postigo, C. Mateo, T. Gómez, F. de Cuadra, P. Dueñas, T. Elgindy, B.-M. Hodge, B. Palmintier, and V. Krishnan, "Phase-selection algorithms to minimize cost and imbalance in U.S. synthetic distribution systems," *Int. J. Electr. Power Energy Syst.*, vol. 120, Sep. 2020, Art. no. 106042.
- [107] H. Li, J. L. Wert, A. B. Birchfield, T. J. Overbye, T. G. S. Roman, C. M. Domingo, F. E. P. Marcos, P. D. Martinez, T. Elgindy, and B. Palmintier, "Building highly detailed synthetic electric grid data sets for combined transmission and distribution systems," *IEEE Open Access J. Power Energy*, vol. 7, pp. 478–488, 2020.
- [108] I. Goodfellow, Y. Bengio, A. Courville, and Y. Bengio, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.
- [109] W. L. Hamilton, R. Ying, and J. Leskovec, "Representation learning on graphs: Methods and applications," *IEEE Data Eng. Bull.*, vol. 40, no. 3, pp. 1–23, Sep. 2017.
- [110] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and P. S. Yu, "A comprehensive survey on graph neural networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 1, pp. 4–24, Jan. 2021.
- [111] IRENA, "Innovation landscape brief: Artificial intelligence and big data," Int. Renew. Energy Agency, Abu Dhabi, United Arab Emirates, Tech. Rep., 2019.
- [112] M. Khodayar, G. Liu, J. Wang, and M. E. Khodayar, "Deep learning in power systems research: A review," *CSEE J. Power Energy Syst.*, vol. 7, no. 2, pp. 209–220, Mar. 2021.
- [113] T. Wu and J. Wang, "Artificial intelligence for operation and control: The case of microgrids," *Electr. J.*, vol. 34, no. 1, Jan. 2021, Art. no. 106890.
- [114] H. Yang, X. Liu, D. Zhang, T. Chen, C. Li, and W. Huang, "Machine learning for power system protection and control," *Electr. J.*, vol. 34, no. 1, Jan. 2021, Art. no. 106881.
- [115] W. Guo, L. Che, M. Shahidehpour, and X. Wan, "Machine-learning based methods in short-term load forecasting," *Electr. J.*, vol. 34, no. 1, Jan. 2021, Art. no. 106884.
- [116] W. L. Hamilton, "Graph representation learning," *Synth. Lectures Artif. Intell. Mach. Learn.*, vol. 14, no. 3, pp. 1–159, 2020.
- [117] D. Bacciu, F. Errica, A. Micheli, and M. Podda, "A gentle introduction to deep learning for graphs," *Neural Netw.*, vol. 129, pp. 203–221, Sep. 2020.
- [118] D. P. Kingma and M. Welling, "Auto-encoding variational Bayes," in *Proc. Int. Conf. Learn. Represent.*, Calgary, AB, Canada, Apr. 2014, pp. 1–15.
- [119] T. N. Kipf and M. Welling, "Variational graph auto-encoders," in *Proc. NIPS Workshop Bayesian Deep Learn.*, Barcelona, Spain, Dec. 2016, pp. 1–3.
- [120] M. Simonovsky and N. Komodakis, "GraphVAE: Towards generation of small graphs using variational autoencoders," in *Proc. Int. Conf. Artif. Neural Netw.* Cham, Switzerland: Springer, Feb. 2018, pp. 412–422.
- [121] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Proc. Neural Inf. Process. Syst.*, 2014, pp. 2672–2680.
- [122] N. De Cao and T. Kipf, "MolGAN: An implicit generative model for small molecular graphs," 2018, *arXiv:1805.11973*.
- [123] A. Bojchevski, O. Shchur, D. Zügner, and S. Günnemann, "NetGAN: Generating graphs via random walks," in *Intl. Conf. Mach. Learn.*, Montreal, QC, Dec. 2018, pp. 610–619.
- [124] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein generative adversarial networks," in *Proc. Int. Conf. Mach. Learn.*, vol. 70, Aug. 2017, pp. 214–223.
- [125] M. Liang, Y. Meng, J. Wang, D. L. Lubkeman, and N. Lu, "FeederGAN: Synthetic feeder generation via deep graph adversarial nets," *IEEE Trans. Smart Grid*, vol. 12, no. 2, pp. 1163–1173, Mar. 2021.
- [126] R. S. Sutton and A. G. Barto, *Introduction to Reinforcement Learning*, 2nd ed. Cambridge, MA, USA: MIT Press, 2018.
- [127] J. You, B. Liu, Z. Ying, V. Pande, and J. Leskovec, "Graph convolutional policy network for goal-directed molecular graph generation," in *Proc. NeuralIPS*, Montreal, QC, Canada, Dec. 2018, pp. 6410–6421.
- [128] J. You, R. Ying, X. Ren, W. Hamilton, and J. Leskovec, "GraphRNN: Generating realistic graphs with deep auto-regressive models," in *Proc. Int. Conf. Mach. Learn.*, Montreal, QC, Canada, Dec. 2018, pp. 5708–5717.
- [129] R. J. Williams and D. Zipser, "A learning algorithm for continually running fully recurrent neural networks," *Neural Comput.*, vol. 1, no. 2, pp. 270–280, 1989.
- [130] R. Liao, Y. Li, Y. Song, S. Wang, W. Hamilton, D. K. Duvenaud, R. Urtasun, and R. Zemel, "Efficient graph generation with graph recurrent attention networks," in *Proc. Adv. Neural Inf. Process. Syst.*, Vancouver, BC, Canada, Dec. 2019, pp. 4255–4265.
- [131] M. Khodayar, J. Wang, and Z. Wang, "Deep generative graph distribution learning for synthetic power grids," 2019, *arXiv:1901.09674*.
- [132] S. Kullback and R. A. Leibler, "On information and sufficiency," *Ann. Math. Statist.*, vol. 22, no. 1, pp. 79–86, 1951.
- [133] D. MacKay, *Information Theory, Inference and Learning Algorithms*. Cambridge, U.K.: Cambridge Univ. Press, 2003.
- [134] E. Hellinger, "Neue Begründung der Theorie quadratischer Formen von unendlichvielen Veränderlichen," *J. Für Die Reine Und Angewandte Mathematik*, vol. 1909, no. 136, pp. 210–271, 1909.
- [135] O. Simpson, C. Seshadhri, and A. McGregor, "Catching the head, tail, and everything in between: A streaming algorithm for the degree distribution," in *Proc. IEEE Int. Conf. Data Mining*, Nov. 2015, pp. 979–984.

- [136] F. Hausdorff, *Grundzüge der Mengenlehre*. Leipzig, Germany: Veit, 1914.
- [137] S. G. Aksoy, K. E. Nowak, E. Purvine, and S. J. Young, "Relative Hausdorff distance for network analysis," *Appl. New. Sci.*, vol. 4, no. 1, pp. 1–25, Oct. 2019.
- [138] GridBright. (2020). *Data Repository for Power System Open Models With Evolving Resources (DR POWER)*. [Online]. Available: <https://eGRIDdata.org/>
- [139] Texas A&M University. (2020). *Electric Grid Test Case Repository*. [Online]. Available: <https://electricgrids.engr.tamu.edu/>
- [140] European Network for Transmission System Operators Electricity. (2016). *Pan-European Power Transmission Grid Dataset (TYNDP 2016)*. [Online]. Available: <https://tyndp.entsoe.eu/>
- [141] Réseau de Transport d'Électricité. (2017). *Open Data Réseaux Énergies*. [Online]. Available: <https://opendata.reseaux-energies.fr/>
- [142] S. Babaeinejadsarookolae et al., "The power grid library for benchmarking AC optimal power flow algorithms," 2019, *arXiv:1908.02788*.
- [143] BetterGrids. (2020). *GRID DATA Repository*. [Online]. Available: [Online]. Available: <https://db.bettergrids.org/>
- [144] A. Xenophon and D. Hill, "Open grid model of Australia's national electricity market allowing backtesting against historic data," *Sci. Data*, vol. 5, no. 1, Oct. 2018, Art. no. 180203.
- [145] Electric Power Research Institute. (2020). *Distributed PV Monitoring and Feeder Analysis*. EPRI Feeders. [Online]. Available: <https://dpv.epri.com/>
- [146] S. Meinecke, *SimBench—Benchmark*. Kassel, Germany: Univ. Kassel, 2019. [Online]. Available: <https://simbench.de/en/>
- [147] A. Xenophon and D. Hill. (Aug. 2018). *Geospatial Modelling of Australia's National Electricity Market Dataset*. Zenodo. [Online]. Available: <https://zenodo.org/record/1326942>
- [148] T. V. Jensen, H. de Sevin, M. Greiner, and P. Pinson. (Sep. 2017). *The RE-Europe Data Set*. [Online]. Available: <https://zenodo.org/record/999150>
- [149] B. Palmintier, D. Krishnamurthy, P. Top, S. Smith, J. Daily, and J. Fuller, "Design of the HELICS high-performance transmission-distribution-communication-market co-simulation framework," in *Proc. Workshop Modeling Simulation Cyber-Phys. Energy Syst. (MSCPES)*, Apr. 2017, pp. 1–6, doi: [10.1109/MSCPES.2017.8064542](https://doi.org/10.1109/MSCPES.2017.8064542).



M. HOSSAIN MOHAMMADI (Member, IEEE) received the B.S. degree in electrical engineering from American University in Dubai, United Arab Emirates, in 2013, with the Valedictorian Award, and the M.Eng. (Thesis) and Ph.D. degrees in electrical engineering from McGill University, Montréal, Canada, in 2015 and 2019, respectively. He is currently a Postdoctoral Research Scientist with the Renewable Energy Integration Group, CanmetENERGY, Natural Resources Canada, leading the development of synthetic benchmarks representative of the Canadian grid. His other research interests include modeling, design, optimization, and control of low-frequency electromagnetic devices, applied machine learning, and statistical analysis. He has been an Active Volunteer in the IEEE Montreal Section, since 2014.



KHALED SALEH (Member, IEEE) received the B.Sc. degree in electrical power engineering from the Petroleum Institute, Abu Dhabi, United Arab Emirates, in 2012, the M.Sc. degree in electrical power engineering from the Masdar Institute of Science and Technology, Abu Dhabi, in 2014, and the Ph.D. degree in electrical engineering from the University of Waterloo, Waterloo, ON, Canada, in 2017. He was a Postdoctoral Fellow with the Department of Electrical and Computer Engineering, University of Waterloo, till 2018. He is currently a Research Scientist with the Renewable Energy Integration Group, CanmetENERGY, Natural Resources Canada. His research interests include the control and protection of renewable energy systems, microgrids, and dc grids. He is an Associate Editor of the IEEE TRANSACTIONS ON SMART GRID.

• • •