Statistical Considerations in the Creation of Realistic Synthetic Power Grids for Geomagnetic Disturbance Studies

Adam B. Birchfield, *Student Member, IEEE*, Kathleen M. Gegner, *Student Member, IEEE*, Ti Xu, *Student Member, IEEE*, Komal S. Shetye, *Member, IEEE*, and Thomas J. Overbye, *Fellow, IEEE*

***Abstract*—Studies to evaluate the power system impacts of geomagnetic disturbances (GMDs) can benefit from geographically realistic public test cases to validate methodologies and analysis tools. Presently very few GMD test cases exist that are not restricted by data confidentiality. In this paper, we outline a method to generate completely synthetic transmission system networks suitable for GMD studies. Public energy and census data form the basis for generation, load, and geographic substation placement. The transmission line topology of the synthetic network is designed to match statistical characteristics observed on the Eastern Interconnect in North America: average nodal degree, average shortest path length, and average clustering coefficient. We apply the Delaunay triangulation to transmission network synthesis, showing it provides an excellent starting place for generating realistic topologies. A 150-bus case is developed and released with benchmark GMD results, for using in testing GMD models and methods.**

***Index Terms*—Geomagnetic disturbance (GMD), geomagnetically induced currents (GICs), power grid topology, synthetic networks.**

I. Introduction

SIMULATION and analysis of geomagnetic disturbances (GMDs) are an important part of building more resilient electric power transmission systems. GMDs occur when solar activity such as earth-directed coronal mass ejections or solar coronal holes causes variations in the earth's magnetic field. Resulting geomagnetically induced currents (GICs) tend to flow through low resistance paths such as transmission lines, connected to earth at grounded substations through transformer neutrals. From a 60 Hz grid perspective, GICs can be deemed as "quasi-dc" (less than 1 Hz), superimposed on the predominantly ac grid. Sufficiently large GICs may cause half-cycle saturation in transformers and damage them; more likely, GICs may lead to relay mis-operation, the tripping of reactive power support devices, and voltage collapse in the grid [1].

Manuscript received February 12, 2016; revised May 9, 2016; accepted June 18, 2016. Date of publication July 18, 2016; date of current version February 16, 2017. This work was supported in part by the National Science Foundation under Award NSF 15-20864, and in part by the U. S. Department of Energy Consortium for Electric Reliability Technology Solutions. Paper no. TPWRS-00235-2016.

The authors are with the University of Illinois, Urbana, IL 61801 USA (e-mail: birchfi2@illinois.edu; gegner2@illinois.edu; txu20@illinois.edu; shetye1@illinois.edu; overbye@illinois.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TPWRS.2016.2586460

To evaluate the threats a major GMD would pose to the power system and the mitigation measures needed to operate the grid safely, the Federal Energy Regulatory Commission in Order 797 directed the North American Electric Reliability Corporation (NERC) to develop standards for planning and operating the grid under GMDs [2]. System studies proposed in NERC's planning standards [3] will require specialized planning tools and modeling capabilities. Methodologies for modeling GMDs for large power system studies are given in [4]. GMD studies need particular inputs such as substation grounding resistances and geographic coordinates for electric field calculations, so the typical IEEE transmission system test cases are not sufficient. The 17-bus case of [5] is one test case designed for GIC calculations; it models a version of the Finnish 400-kV grid with dc resistances given for substation grounding and transmission lines. The 20-bus dc case of [6] models two voltage levels and adds transformer dc resistance values. These cases do not include the ac power flow parameters necessary to couple GIC calculations into a steady-state voltage stability analysis of the power system under a GMD, as [4] describes and the NERC standards [3] propose. Existing larger and more detailed cases of the actual grid contain critical infrastructure information and are subject to data confidentiality, restricting their use for public validation of GMD analysis methodologies. The development of larger public test cases with more detail, including both dc and ac parameters will aid research on this topic and allow for cross-validating software and analysis methods.

This paper presents a two-part method to generate realistic synthetic power system test cases across a geographic area. In the first step, fictitious substations with realistic load and generation profiles are sited with geographic coordinates, using publicly available information. Next, transmission lines are placed to connect these substations. The approach of this paper is to match statistical characteristics of the actual transmission line topologies of the Eastern Interconnect (EI) in North America. Recent studies examining the topological structure of transmission systems are applied, along with our own observations, to build large synthetic power systems which, though not based on the real grid, match topological properties found in real grids. The systematic methods, based on graph and geometric calculations, are useful for automated creation of realistic GMD test cases, larger and more complex than existing cases. The process described in this paper is illustrated on North American

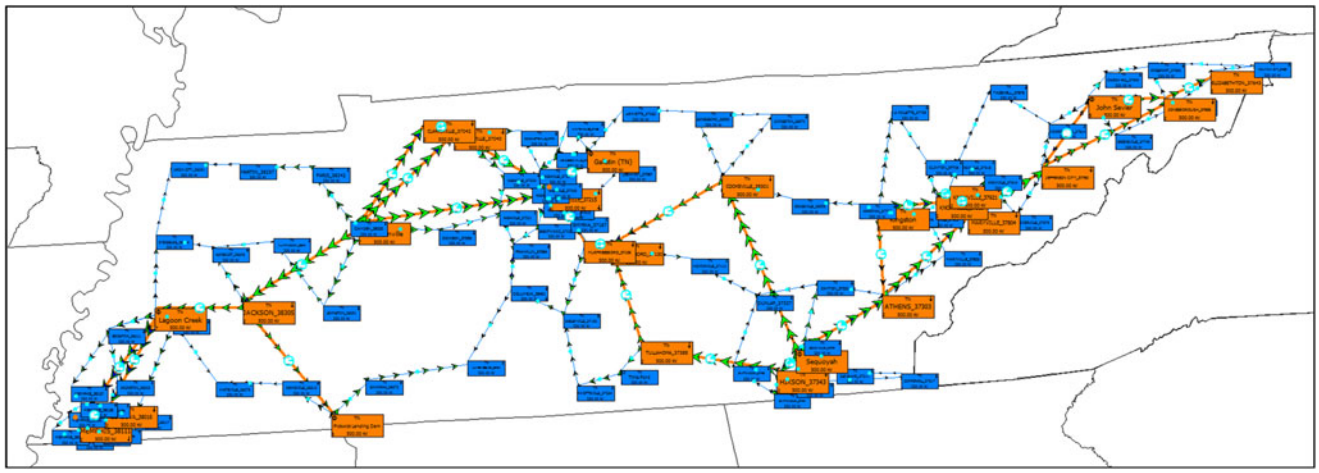

Fig. 1. Geographic single-line diagram of 150-bus GMD benchmark test case. This case is synthetic, created from a statistical analysis of the EI, and does not correspond to the actual grid in this location.

geography by the creation of a synthetic 150-bus test case, with appropriate parameters for dc GIC calculations as well as those necessary for ac power flow solutions under GMD scenarios.

Section II reviews the relevant literature on analyzing power systems as complex graphs, background information which will inform the automated line placement method proposed. Section III introduces our 150-bus ac case. Section IV describes the process for geographically placing substations, and Section V describes the statistics-based method for creating transmission line topologies. Finally, Section VI presents the benchmark GIC and steady-state voltage stability results, and Section VII concludes the paper.

TABLE I
150-BUS CASE SUMMARY

Substations	98
Load substations	90
Generating substations	8
Generating units	27
Generator step-up transformers	27
500 kV–230 kV transformers	33
230 kV transmission lines	127
500 kV transmission lines	30
Shunt capacitor banks	3

II. BACKGROUND: THE GRID AS A COMPLEX NETWORK

In the last twenty years, attention has grown for analyzing the circuit structure of electric power transmission systems as a complex graph. Viewing electric buses as vertices connected by electric branches as edges, this paradigm allows the power grid to be compared with a variety of real-world technological, biological, and social networks. A host of properties and statistics can be observed to characterize these graphs, and recently these statistics have been applied to the synthesis of fictitious networks whose properties match actual grids.

Two salient properties of many real networks are highlighted in [7]: high clustering and a low average path length between any two nodes. Networks with these properties are called Small World networks, and they can be modeled as a regular lattice with some edges rewired at random. Several researchers have declared power systems to be Small World networks [8]; others, such as [9], have noted differences between electric grids and Small World networks as introduced by [7]. Node degrees, or the number of edges connected to each node, are examined in [10]. In contrast to many other real-world networks, electric power grids have an exponential node degree distribution. This property reflects the grid's geographic constraints and contributes to grid robustness. Numerous studies have presented various statistics on power transmission system topology, showing that the grid's network structure strongly affects grid physical security, communication system design, stability, and optimal control [8]–[13]. A method to synthesize some of these properties in automatically-generated power system topologies was proposed in [13], using a modified Small World algorithm.

An important distinctive of power system graph analysis that has not been adequately addressed in previous studies is the geographic layout of system buses. Building on previous work, we group buses into electric substations, with each substation having a particular geographic location. Individual substations can contain multiple nominal voltage levels and multiple buses. Transmission lines then connect these buses in overlaid, interconnected graphs at each nominal voltage level. Though not always included in standard power system models, geographic coordinates are critical for GMD analysis, and they also allow for the consideration of geometric graph properties in building synthetic topologies. This paper builds on the synthetic model generation method of [13] by using statistics of the overlaid graphs, considering nominal voltage levels, geographic relations, and the load and generation capacities.

III. OVERVIEW: 150-BUS CASE

The 150-bus case shown in Fig. 1 and summarized in Table I is presented as an example of our method. It contains 98 substations placed on the U.S. State of Tennessee, a geographic footprint covering approximately 35°N to 36.5°N latitude and 90°W to 82°W longitude. The size was chosen because it is about an order of magnitude larger than the existing GIC test case from [6] but is still small enough to make for easy GMD analysis

TABLE II
GENERATOR PARAMETERS

Governor type	Max Mvar as fraction of MW capacity	Min Mvar as fraction of MW capacity
Steam	0.466	−0.122
Gas	0.509	−0.111
Gas Turbine	0.560	−0.164
Hydro	0.384	−0.049
Nuclear	0.368	−0.082
Wind	0.213	−0.144

benchmarking. Tennessee spans enough geographic area to make the GIC analysis significant and has a balanced population distribution. 500 kV and 230 kV were chosen as the nominal voltage levels for this case; networks at these high voltage levels have primary applicability to GIC studies because of their low resistance and long geographic distance [14]. The objective is not to recreate the actual Tennessee grid, but to build a synthetic case with representative characteristics of a typical power system. The model is publicly available online [15], with full specifications and benchmark results. The next two sections present the methodology used in creating this case and the details of how the case was made. Section VI explains the GIC analysis used to produce the benchmark GMD study results.

IV. GEOGRAPHIC SUBSTATION PLACEMENT

The approach to creating a synthetic grid begins by assigning geographic coordinates to substations. In statistical analysis of the EI substations, about 90% include either load or generator buses. Hence for the synthetic network method here we assume every substation contains buses with either generation or load. Of course since an individual substation can contain multiple buses, the synthetic models will still include many buses with neither load nor generation, matching the results from prior studies [16]. The substations with generation and load will be based on public information about the underlying geographic region. This section explains the method applied to geography in the United States, a process more thoroughly described in our previous work [17].

The United States Energy Information Administration maintains a yearly survey of the nation's generators. The 2014 survey data, which is released to the public, gives the fuel type, generating capacity, and geographic coordinates of 9250 American power plants [18]. In building fictitious cases, we place generation substations at the reported geographic locations of these actual generators, with generation limits as given. Based on the fuel type and generation capacity, we estimate the generator's capability curve and add corresponding reactive power limits with a linear approximation, as shown in Table II. Hierarchical clustering is used to combine smaller generators, depending on the level of detail desired for a specific case [19].

Although the location and parameters of load substations are not fully public information, load is highly correlated with population. We use the public U.S. census database as the fundamental basis for load substations, starting with the geographic coordinates and population at each U.S. postal code [20].

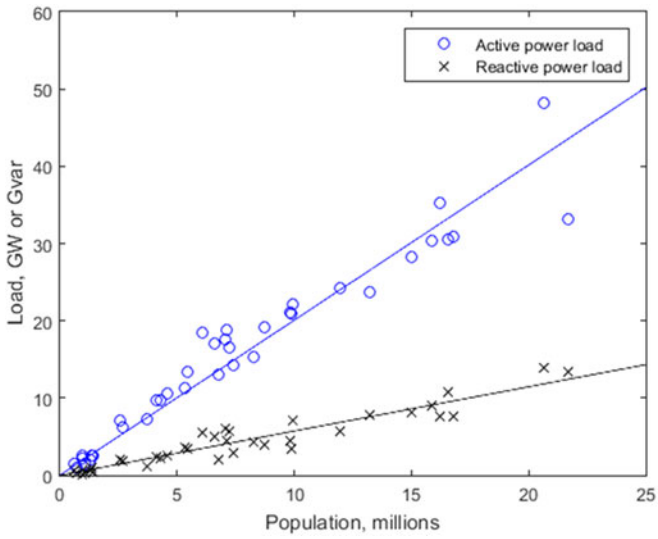

Fig. 2. Active and reactive power load vs population in the 37 U.S. states contained in the EI. Trend lines show 2.00 kW and 0.57 kvar per capita, resulting from regression analyses with $R^2 = 0.935$ and 0.903, respectively.

Depending on the voltage level and system detail desired, combining the postal codes may be necessary using hierarchical clustering. Based on a study of state-level load and population regression (see Fig. 2), a peak summer load of 2.00 kW per capita and 0.57 kvar per capita are assigned.

Substation grounding resistance is another important component of GIC calculations. Current planning models and IEEE test cases do not contain substation grounding data, since they are not used in standard power flow studies [4]. The typically observed range of this parameter is 0.015 to 1.5 Ohms [21], [22]; previous work has shown that GIC results are quite sensitive to these values and that actual values should be used wherever possible [21]. In this method, each substation has a grounding resistance assigned based on the assumed size of the substation, considering the nominal voltage level and the number of buses in a substation.

For the 150-bus test case presented by this paper, we cluster the 276 generation units publicly reported for Tennessee into 27 equivalent units located at 8 substations. Each generating unit will then have a total maximum MW generating capacity and fuel type, both of which are public information. These are used to assign reactive power limits according to Table II. We also add generator step-up transformers, which are important to GIC calculations, from an assumed generation voltage of 13.8 kV to the substation main bus. For load, the clustering process groups the state's 989 zip code tabulation areas into 90 clusters. Each cluster becomes a load substation with a load value based on the total population of the postal codes in the cluster. Substation grounding resistance is assigned for each substation. Summary data on this is given in Table III; the complete data is available at [15]. The result of this step is a set of 90 load substations and 8 generating substations, with geographic coordinates, buses, load, generation parameters, and grounding resistance already assigned.

TABLE III
SUBSTATION DATA SUMMARY

Maximum nominal voltage (kV) ↓	Generating		Non-generating	
	Number of substations	Grounding resistance (Ohm)	Number of substations	Grounding resistance (Ohm)
500	7	0.11–0.15	18	0.18
230	1	0.38	72	0.47

V. BUILDING TRANSMISSION LINE TOPOLOGY

The next step in building synthetic grids is to connect the buses of these synthetic substations with transmission line networks at multiple nominal voltage levels. In this section, we present some basic statistical observations from 94 single-voltage transmission networks within the EI. These networks are of various sizes at voltage levels ranging from 115 kV to 500 kV. The collected topological and geometric statistics are then used in building similar synthetic topologies for the 150-bus case.

A. Statistical Analysis of the EI

For a general transmission network at a single nominal voltage level, *n* substations are connected with *m* transmission lines. The topology of these transmission lines can be expressed with the symmetric adjacency matrix A, where

$$A(i,j) = \begin{cases} 1, & \text{if a exists from } i \text{ to } j \\ 0, & \text{else} \end{cases}$$
$$i,\ j \in \{1,\ 2,\ \ldots,\ n\}. \tag{1}$$

Node degree k_i for node *i* indicates the number of edges incident on that node.

$$k_i = \sum_{j=1}^{n} A\,(i,\ j) \tag{2}$$

The average nodal degree is

$$\langle k \rangle = \frac{1}{n} \sum_{i\,=\,1}^{n} k_i = \frac{2m}{n}\ . \tag{3}$$

The EI networks show a highly linear relationship between $2m$ and n; the correlation coefficient is 0.9991. This indicates an average nodal degree of 2.43, consistent across networks of various sizes and voltage levels. This number fits the range of results found in other studies [8], [13]. From the average nodal degree, the number of transmission lines in a network m will be $1.22n$, according to (3).

The transmission networks of a single voltage level are, in general, totally connected, which means that they will contain a spanning tree with $n-1$ edges. Therefore, this paper proposes constructing synthetic networks as a minimum spanning tree with $0.22n$ additional lines added, to ensure $\langle k \rangle$ will be preserved.

The distribution of node degrees across a system is another common indicator of network structure. The internet, social networks, and some biological networks have a power-law scale-free node degree distribution [10], [23]. In contrast, power system networks lack central hubs with a large degree, and have consistently been shown to have an exponential node degree distribution, defined by

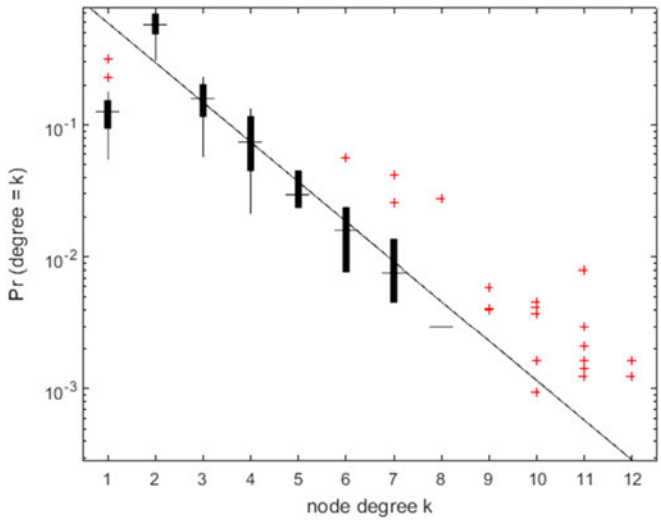

Fig. 3. Node degree distribution for 28 of the single-voltage transmission networks on the EI with $n > 100$, and exponential regression for $k \in [3,\ 7]$. Thin vertical lines are range; thicker lines are inner quartiles; horizontal dash shows the median; and the crosses show outliers.

$$\Pr(k) = \alpha e^{-\beta k}, \tag{4}$$

for some constants α and β, which we confirm in Fig. 3 for the 28 networks of our study with $n > 100$, using $\alpha = 1.19$ and $\beta = 0.69$. The exponential node degree distribution reflects the geographic constraints which dominate the careful planning of power networks. We observe that transmission networks consistently deviate from the exponential distribution with $k = 1$ and $k = 2$. This property has been noted by previous studies [8], [13], and reflects the planning priority to avoid substations relying on a single line for service.

The average shortest path length, $\langle l \rangle$, denotes how many hops along the edges of a network separate two nodes, on average. It can be calculated from a network's adjacency matrix A using Dijkastra's algorithm [24]. A purely random network, defined by a set of nodes in which each pair is connected with probability *p*, will have $\langle l \rangle$ which scales logarithmically with *n* [10]

$$\langle l \rangle_{\text{rand}} \approx \frac{\ln(n)}{\ln(\langle k \rangle)}. \tag{5}$$

This is the small world property: even in very large networks, most nodes are only a few hops apart. On the other hand, regular lattice networks have much larger average path lengths, scaling with a $1/d$ power law, where *d* is the dimension of the lattice. So two-dimensional lattices will have

$$\langle l \rangle_{2D} \propto \sqrt{n}. \tag{6}$$

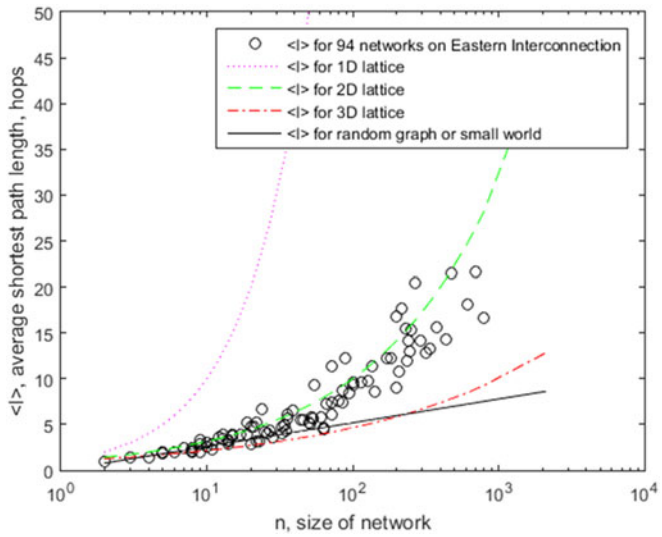

Fig. 4. Average shortest path length $\langle l \rangle$ for 94 single-voltage transmission networks on the EI, compared to expected scaling properties for random graph and 1D, 2D, and 3D regular lattices.

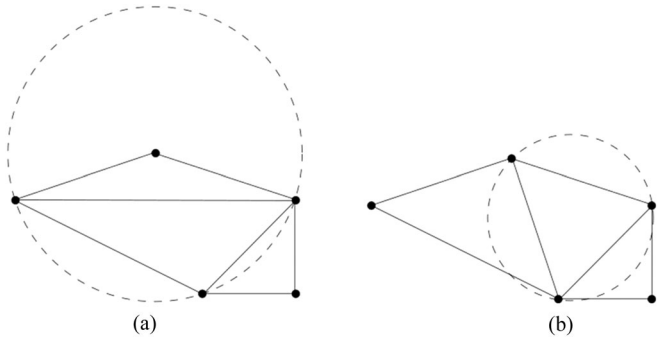

Fig. 5. Triangulation of a set of 5 points. (a) is not the Delaunay triangulation, because at least one triangle's circumcircle contains another point. (b) is the Delaunay triangulation of these points, because no triangle's circumcircle contains another point. Note that there are both larger and smaller angles in (a) than in (b).

Small World networks, as presented in [7] and compared to the power system by many [8], [9], [11]–[13], retain the logarithmic scaling property of random networks.

Fig. 4 shows the average shortest path length calculations for the 94 transmission networks surveyed from the EI. It can be seen that when transmission networks are considered using each voltage level independently, the logarithmic scaling property for average shortest path length is not seen, and the network approximates a regular two-dimensional lattice. Reference [9] also observes a difference between the scaling properties of power systems' average shortest path length and that of Small World networks. The shortcuts mentioned in other previous work [13] reflect the superposition of multiple voltage levels, but these levels individually appear to be more accurately approximated by lattices.

The average clustering coefficient, $\langle c \rangle$, is the second primary indicator of a small world. An individual node with node degree k_i has τ_i possible edges interconnecting its neighbors, where

$$\tau_i = \frac{k_i\,(k_i - 1)}{2}\,. \tag{7}$$

The clustering coefficient for a node c_i is defined by

$$c_i = \frac{e_i}{\tau_i}\,, \tag{8}$$

where e_i is the number of edges actually interconnecting node i's neighbors. $\langle c \rangle$ is the average of c_i over all nodes in the network. For completely random networks the value $\langle c \rangle$ is given by

$$\langle c \rangle_{\text{rand}} \approx \frac{\langle k \rangle}{n}. \tag{9}$$

Clustering in truly random networks therefore decreases rapidly with system size [10]. Regular lattices, in contrast, have much larger clustering coefficients that do not scale with system size [7]. Small World networks approximate the clustering coefficient properties of a lattice, while maintaining the average shortest path length of a random network [7], [10]. In general, the 94 transmission networks studied show a large clustering coefficient uncorrelated to network size, as expected for a Small World network or regular lattice. The average $\langle c \rangle$ among the networks is 7.5%, but the value varies significantly. Many small networks have so few edges that either no clustering appears or the clustering is quite high. For networks with more than 100 substation nodes, $\langle c \rangle$ falls consistently in the range 1%–15%, much higher than for a random network.

For GMD studies, the geographic layout of transmission lines is important, not only the lines' topology. Because our approach begins by giving each substation node latitude and longitude coordinates, this additional information is available for building the synthetic topologies. From computational geometry, the Delaunay triangulation for a set of points is a simple planar mesh that connects n points with non-overlapping line segments, dividing the region into triangles. The Delaunay triangulation is distinguished from a general triangulation by the property that no triangle has a point inside its circumcircle, and the smallest angle is maximized [25], [26]. The triangles are therefore nicely shaped; when possible they have all acute angles. A node's neighbors on the Delaunay triangulation will also be the nodes nearest to it by geometric distance. Fig. 5 illustrates the Delaunay triangulation.

From graph theory, Euler's formula states for a planar graph such as the Delaunay triangulation, with n points, m edges, and f faces, [24]

$$n - m + f \;=\; 2. \tag{10}$$

Since in the Delaunay triangulation each face is a triangle except the region outside the graph,

$$3\,(f - 1) + p \;=\; 2m, \tag{11}$$

where p is the of edges on the outer boundary of the graph Substituting for f and solving for m we get

$$m \;=\; 3n - p - 3. \tag{12}$$

So a large Delaunay triangulation will have about $3n$ edges, since p will be small with respect to n. Therefore the average

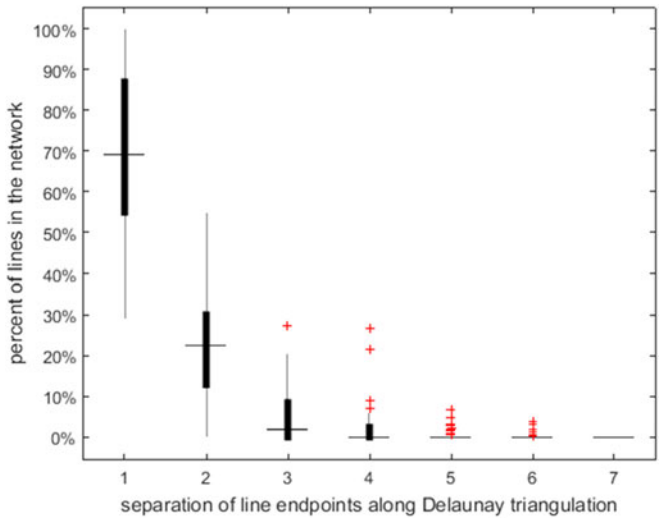

Fig. 6. Separation of transmission line endpoints on the Delaunay triangulation, in hops, for 94 single-voltage transmission networks on the EI. Thin vertical lines are the range; thicker lines are inner quartiles; horizontal dash shows the median; and the crosses show outliers.

nodal degree $\langle k \rangle$ will be 6 regardless of system size [25]. A Delaunay triangulation on points uniformly distributed approximates a two-dimensional regular lattice, with a constant clustering coefficient of 40%, and an average shortest path length of $\sqrt{n}$, where n is the number of points [26].

Because Delaunay matches these basic properties found in real power systems and connects nodes with geometric proximity, it provides a reasonable subject for direct comparison to the 94 single-voltage transmission networks of our study. Each actual transmission line in these networks has a Delaunay separation, indicating the shortest distance between the line's endpoints along the segments of the substations' Delaunay triangulation. A separation of 1 means the line is itself one of the segments of the Delaunay triangulation; a separation of 2 means a path of two Delaunay segments connects the line's endpoints. The results are plotted in Fig. 6. The Delaunay triangulation contains 70% of an average network. This is significant since there are $n(n-1)$ possible transmission lines, $1.22n$ actual transmission lines, and $3n$ segments on the Delaunay triangulation. The actual lines that are not on the Delaunay triangulation are nearly all only a few hops separated. 98% of lines in the average network have a Delaunay separation of 3 or less. A substation's Delaunay neighbors are also its closest neighbors by geometric distance, so this property reflects the geographic constraints of transmission system planning.

B. Method for Adding Branches to the 150-Bus Test Case

In the 150-bus case, we give all 98 substations a 230 kV bus, and the largest 18 load and 7 generation substations a 500 kV bus. The transmission line topologies at each nominal voltage level are built by adding $1.22n$ transmission lines, which are all selected from the $3n$ lines of the Delaunay triangulation. We begin with the Euclidian minimum spanning tree at each voltage level, guaranteed to be a subset of the Delaunay triangulation.

TABLE IV
LINE PARAMETERS

kV	ACSR Conductor	Tower spacing	Per-unit impedance, per 100 miles			MVA limit
			x	r	b	
500	Falcon, 3 bundle	42 ft.	0.00193	0.000912	2.126	3585
230	Cardinal, 1 bundle	24 ft.	0.15186	0.021323	0.278	402

This ensures a connected graph, and reflects the geographic constraints of real power systems. Then about $0.22n$ lines will be added in an iterative fashion, all taken from the $3n$ segments of the substations' Delaunay triangulation. The expected properties $\langle l \rangle$ and $\langle c \rangle$ follow the scaling behaviors for the regular lattice of which they are a subset. These additional lines will be selected from a dc power-flow based metric, described below in this section.

Each transmission line added is given a per-mile series and shunt impedance according to typical values for a line at that voltage level, as shown in Table IV. The length of the transmission line is approximated from the geographic distance between the two substations it connects. The substations which overlap the networks of the two voltage levels are given transformers with 1.6% impedance and x/r ratio of 80, rated for 1000 MVA. For GIC calculations, these transformers will be considered auto-transformers. Following the calculation procedure of [4], the ac resistance of 0.0002 per-unit will be converted into GIC dc resistance values of 0.25 Ohms in the series winding and 0.167 Ohms in the common winding.

Planners use iterative power flow analysis to determine where new lines are needed; likewise, this algorithm finds good locations for new lines from the voltage angle difference between two adjacent points on the Delaunay triangulation. The minimum spanning tree ensures the entire network will be connected, making the system bus susceptance matrix $\bar{B}$ nonsingular, and allowing dc power flow solutions. Each bus's real power injection $\bar{P}_{\text{inj}}$ is known from the load definitions and a proportional generation dispatch, where each generator is scaled down from maximum output to match the total load. The largest generator is selected as the slack bus, and the standard dc power flow solution process can now calculate system voltage angles $\bar{\theta}$ as

$$\bar{\theta} = \bar{B}^{-1}\ \bar{P}_{\text{inj}}. \tag{13}$$

The additional $0.22n$ segments are selected in 10 iterations of this dc power flow, the $\bar{B}$ matrix being updated at each iteration to include newly added $0.022\ n$ transmission lines at each voltage level. In each iteration, each unused Delaunay segment is evaluated based on the expected power P_{exp} it would carry as a transmission line, estimated from the voltage angles θ_1 and θ_2, the geographic distance d_{21} in miles, and the known per-mile reactance x_l of lines used at that voltage level.

$$P_{\text{exp}} = \ x_l \cdot d_{21} \cdot (\theta_2 - \theta_1) \tag{14}$$

The $0.022n$ segments at each voltage level with the largest P_{exp} that do not share nodes are then added to the network. For the

TABLE V
BASIC STATISTICS OF THE 150-BUS CASE

Property	230-kV network	500-kV network
n	98	25
m	121	30
$\langle k \rangle$	2.47	2.40
$\langle l \rangle$	10.83	5.42
$\langle c \rangle$	0.06	0.25

TABLE VI
TRANSFORMER DATA SUMMARY

Nominal voltage (kV)	Number of transformers	Type	Winding config.	Winding resistance, Ω		K
				High (Series)	Low (Common)	
500-230	33	Auto	GWye- GWye-	0.2500	0.1841	1.8
500–13.8	26	GSU	GWye-Delta	0.153–1.3554	N/A	1.8
230–13.8	1	GSU	GWye-Delta	1.4469	N/A	1.5

230 kV network, an additional constraint is enforced to reduce radial loads.

The result is a set of 121 synthetic transmission lines at 230 kV and 30 lines at 500 kV, with per-distance parameters according to Table IV, connecting the 98 synthetic substations across the geographic area. The statistical data from this 150-bus case is shown in Table V. Because the number of transmission lines is directly controlled, the average nodal degree $\langle k \rangle$ is very close to the target 2.43 for both networks. With the segments chosen from the Delaunay triangulation, the average shortest path length $\langle l \rangle$ turns out remarkably close to the trend $\sqrt{n}$ for both networks as well. The clustering coefficient $\langle c \rangle$ is within the wide range typically seen, much higher than for a random network. For the larger 230 kV network, the 6% clustering is nearly centered in the 1%–15% bounds observed from the EI. Therefore, on several important metrics, the resulting case contains a statistically realistic transmission line topology.

Six heavily-loaded transmission lines and eight transformers are upgraded to double circuits by placing an identical device in parallel. Three 150 Mvar capacitor banks are added for additional reactive power support to lower voltage buses. The ac power flow solutions then show no voltage violations or line overloads, and the geographic coordinates and dc element resistances are in place for GMD studies.

VI. GEOMAGNETIC DISTURBANCE CASE RESULTS

As a benchmark, representative GMD study results for the 150-bus case are provided in this section, following the GIC calculation methodology of [4].

Several transformer parameters not used in power flow analysis, which are essential to GMD studies, are included in the model. Table VI summarizes these parameters for the 60 transformers in the case; the full data is available at [15]. The winding resistances were estimated from the ac series resistance values of the transformers using the methodology described in [4]. Note that the low side winding resistances for the GSUs have not been given, due to the delta connection on the low side, which does not provide a path to ground for the GICs. In the last column, the K values relate the transformer effective GICs to reactive power losses according to the following expression derived from [27],

$$Q_{\text{loss, pu}} = V_{\text{pu}}\, K I_{\text{GIC,pu}}, \tag{15}$$

where $Q_{\text{loss, pu}}$ is the total three-phase reactive losses in a transformer in per unit, V_{pu} is the per unit terminal voltage of the high side winding of the transformer, and $I_{\text{GIC,pu}}$ is the "effective" GIC in the transformer windings. The transformer effective GIC depends on the winding configuration, and whether the transformer is an autotransformer or not; its calculation methodology is described in [4]. These losses are mapped to the power flow model in order to perform ac power flow studies to study system voltages.

TABLE VII
TRANSFORMER GIC RESULTS

HV Bus #	LV Bus #	I_n (Amps)	I_{GIC} (Amps)	HV Bus #	LV Bus #	I_n (Amps)	I_{GIC} (Amps)
92	63	19.536	8.111	139	123	-19.038	6.346
93	54	7.111	1.609	139	138c1	-7.824	3.411
94	40	1.274	5.169	139	138c2	-7.824	3.411
95	86	-42.335	13.332	140	135	-8.982	2.994
96	23	-11.798	5.805	142	110	19.558	6.519
97	43	97.751	32.504	142	111	13.004	4.335
98	11	-24.757	3.661	142	112	36.911	12.304
99	79ck1	-67.529	20.723	142	141c1	21.610	6.933
99	79ck2	-67.529	20.723	142	141c2	21.610	6.933
100	64	13.283	3.928	144	1	-42.354	14.118
101	22ck1	18.977	5.189	144	116	-14.083	4.694
101	22ck2	18.977	5.189	144	117	-14.083	4.694
102	8	-13.320	5.415	144	118	-14.083	4.694
103	45	148.574	42.687	144	119	-42.354	14.118
104	71	-69.125	18.184	144	143ck1	-10.624	4.907
105	90	38.977	15.403	144	143ck2	-10.624	4.907
106	27	-39.950	19.080	146	130	3.266	1.089
107	33	16.024	0.093	146	131	3.266	1.089
108	14	4.424	7.313	146	132	3.266	1.089
109	49	27.234	8.327	146	133	7.604	2.535
137	124	-9.658	3.219	146	134	7.604	2.535
137	125	-9.658	3.219	146	145c1	8.225	1.109
137	126	-9.658	3.219	146	145c2	8.225	1.109
137	127	-9.658	3.219	148	129	18.192	6.064
137	128	-14.468	4.823	148	147	32.803	12.641
137	136ck1	-17.414	3.603	150	113	6.695	2.232
137	136ck2	-17.414	3.603	150	114	6.695	2.232
139	120	-6.376	2.125	150	115	2.909	0.970
139	121	-6.376	2.125	150	149c1	7.174	2.541
139	122	-19.038	6.346	150	149c2	7.174	2.541

For a 1 V/km, uniform, eastward electric field (E), Table VII shows the neutral GIC I_n in Amperes and the effective GICs I_{GIC} in Amperes in all of the 60 transformers in the case. The neutral current I_n is the sum of the neutral current for all three phases, and is oriented with positive indicating current into the ground. Transformers highlighted yellow in Table VII show the highest effective GICs for this scenario.

The ac power flow data in the case can be used to perform a steady-state voltage stability study in the presence of GICs. For voltage stability analysis, the applied electric field, still eastward, is gradually increased in steps of 1 V/km. At each step,

TABLE VIII
GIC STEADY-STATE VOLTAGE STABILITY RESULTS

Electric field (V/km)	Bus 99 voltage (pu)	Bus 100 voltage (pu)	Bus 101 voltage (pu)	Bus 103 voltage (pu)	Bus 144 voltage (pu)	Total reactive losses (Mvar)
0	1.02603	1.03946	1.03985	1.03733	1.05404	0
1	1.02003	1.03613	1.03667	1.02811	1.05217	475.21
2	1.01401	1.03279	1.03347	1.01887	1.0503	944.35
3	1.00797	1.02945	1.03027	1.00962	1.04844	1407.41
4	1.00192	1.0261	1.02707	1.00034	1.04656	1864.38
5	0.99585	1.02253	1.02385	0.99041	1.04469	2314.84
6	0.98972	1.01894	1.02062	0.98037	1.0428	2758.91
7	0.98335	1.01505	1.01731	0.96943	1.04081	3195.4
8	0.97694	1.00792	1.01386	0.94816	1.0388	3615.63
9	0.97049	1.0007	1.01016	0.92667	1.03673	4026.2
10	0.96304	0.99308	1.00636	0.90461	1.03456	4425.38
11	0.95528	0.98466	1.00185	0.88151	1.03222	4811.84
12	0.92292	0.96307	0.99336	0.84402	1.02706	5119.83
13	0.88157	0.93449	0.98103	0.79799	1.02038	5368.95

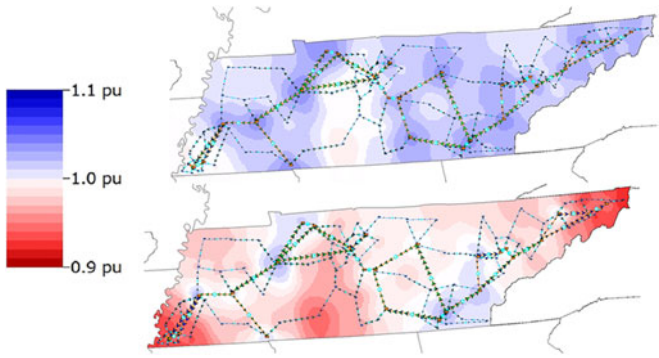

Fig. 7. 150-bus case bus voltage contours. (Above) Base case. (Below) Voltages at power flow solved with GICs and reactive losses at electric field of 7 V/km.

the power flow is solved, and the bus voltages and GIC-induced reactive losses of all transformers are noted. This is done till the power flow solution does not converge. Here the point of non-convergence was reached at $E = 14$ V/km. Table VIII shows bus voltages at select five, 500 kV buses until the last valid power flow solution. The detailed numbers in Tables VII and VIII have been provided to help in benchmark comparisons. These five buses belong to three major metro areas in Tennessee, a key junction substation in the case, and one substation near the edge of the network. Fig. 7 shows the voltages contoured at different electric field strengths to show the impact of the losses on the system voltages. It shows that the voltages start to fall first near the northeast and southwest edges of the network. Note that this in no way implies that this response is expected in this footprint for an actual GMD, as this is a synthetic case. Finally, Fig. 8 plots the bus voltages and losses from Table VIII, showing the gradual voltage collapse in the system. At $E = 13$ V/km there were 49 buses below 0.9 pu voltage.

VII. CONCLUSION

This paper presents a process to generate synthetic power system test cases, focusing on building cases appropriate for

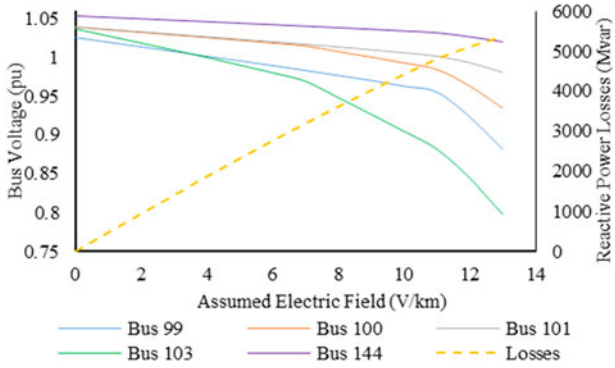

Fig. 8. Graphical representation of Table VII, showing the gradual voltage collapse, and linearly increasing reactive losses in the system.

GMD studies. Because substations are based from the beginning in geographic locations, the transmission line topology is able to reflect actual grids' geographic constraints through the use of the Delaunay triangulation. The line placement method considers both topological properties of actual grids and a power flow metric to connect load and generation. The method can be automated, allowing larger cases to be built.

Cases created using this method are synthetic, with no relation to the actual grid in their geographic location; therefore, by nature they pose no security concern and are public for comparing results among researchers. This framework may be extendable in future work to build test cases for other areas of power systems research and for education.

This example 150-bus test case builds on previous GMD test cases with increased size and complexity. Network parameters for an ac power flow solution allow for voltage stability studies coupled to the dc GIC calculations, to evaluate grid performance under a GMD. Benchmark GIC results are given for the case, as well as a system voltage stability analysis under a GMD scenario.

REFERENCES

[1] "Special reliability assessment interim report: effects of geomagnetic disturbances on the bulk power system," North Amer. Electr. Rel. Corp., Atlanta, GA, USA, Feb. 2012.
[2] *Reliability Standard for Geomagnetic Disturbance Operation*, Federal Energy Regulatory Commission, Docket No. RM14-1-000, Order No. 797, Jun. 2014.
[3] *Transmission System Planned Performance for Geomagnetic Disturbance Events*, NERC Std. TPL-007-1, Jun. 2014.
[4] T. J. Overbye, T. R. Hutchins, K. Shetye, J. Weber, and S. Dahman, "Integration of geomagnetic disturbance modeling into the power flow: A methodology for large-scale system studies," in *Proc. North Amer. Power Symp.*, Champaign, IL, USA, Sep. 2012, pp. 1–7.
[5] R. Pirjola, "Properties of matrices included in the calculation of geomagnetically induced currents (GICs) in power systems and introduction of a test model for GIC computation algorithms," *Earth Planet Space*, vol. 61, no. 2, pp. 263–272, Feb. 2009.
[6] R. Horton, D. H. Boteler, T. J. Overbye, R. Pirjola, and R. C. Dugan, "A test case for the calculation of geomagnetically induced currents," *IEEE Trans. Power Del.*, vol. 27, no. 4, pp. 2368–2373, Oct. 2012.
[7] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393, no. 6684, pp. 440–442, Jun. 1998.
[8] 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, Jun. 2013.

[9] 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.
[10] R. Albert and A.-L. Barabási, "Statistical mechanics of complex networks," *Rev. Mod. Phys.*, vol. 74, no. 1, pp. 47–97, Jan. 2002.
[11] R. Albert, I. Albert, and G. L. Nakarado, "Structural vulnerability of the North American power grid," *Phys. Rev. E*, vol. 69, no. 2, pp. 1–3, Feb. 2004.
[12] P. Hines, S. Blumsack, E. Cotilla Sanchez, and C. Barrows, "The topological and electrical structure of power grids," in *Proc. 43rd Hawaii Int. Conf. Syst. Sci.*, Koloa, HI, USA, Jan. 2010, pp. 1–10.
[13] 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.
[14] White Paper Supporting Network Applicability of EOP-010-1. North Amer. Electr. Rel. Corp., 2013. [Online]. Available: http://www.nerc.com/pa/Stand/Project201303GeomagneticDisturbanceMitigaion/ApplicableNetwork_clean.pdf
[15] (2016). [Online]. Available: http://icseg.iti.illinois.edu/synthetic-power-cases/uiuc-150-bus-system/
[16] Z. Wang and R. J. Thomas, "On bus type assignments in random topology power grid models," in *Proc. 48th Hawaii Int. Conf. Syst. Sci.*, Koloa, HI, USA, Jan. 2015, pp. 2671–2679.
[17] K. M. Gegner, A. B. Birchfield, T. Xu, K. S. Shetye, and T. J. Overbye, "A methodology for the creation of geographically realistic synthetic power flow models," in *Proc. IEEE Power & Energy Conf.*, Champaign, IL, USA, Feb. 2016.
[18] U.S. Energy Information Association. Form EIA-860, 2014. [Online]. Available: http://www.eia.gov/electricity/data/eia860/index.html
[19] R. Xu and D. C. Wunsch, "Hierarchical clustering," in *Clustering*. Hoboken, NJ, USA: Wiley, 2009, pp. 67–72.
[20] U.S. Census Bureau. 2010 Census Gazetteer Files: ZIP Code Tabulation Areas. (2010). [Online]. Available: https://www.census.gov/geo/maps-data/data/gazetteer2010.html
[21] U. Bui, T. J. Overbye, K. Shetye, H. Zhu, and J. Weber, "Geomagnetically induced current sensitivity to assumed substation grounding resistance," in *Proc. North Amer. Power Symp.*, Manhattan, KS, USA, Sep. 2013, pp. 1–6.
[22] *IEEE Guide for Safety in AC Substation Grounding*, IEEE Std. 80-2013, pp. 1–226, May 2015.
[23] M. Newman, "The structure and function of complex networks," *SIAM Rev.*, vol. 45, no. 2, pp. 167–256, Jan. 2003.
[24] D. B.West, *Introduction to Graph Theory*. Upper Saddle River, NJ, USA: Prentice Hall, 1996.
[25] F. M. Preparata and M. I. Shamos. *Computational Geometry: An Introduction*. New York, NY, USA: Springer-Verlag, 1985.
[26] M. S. Smit, "Epidemic Delaunay," (2009). [Online]. Available: http://graphics.tudelft.nl/∼matthijss/epidemic_delaunay/paper.pdf
[27] K. S. Shetye and T. J. Overbye, "Parametric steady-state voltage stability assessment of power systems using benchmark scenarios of geomagnetic disturbances," in *Proc. Power & Energy Conf.*, Champaign, IL, Feb. 2015, pp. 1–7.

Adam B. Birchfield (S'13) received the B.E.E. degree from Auburn University, Auburn, AL, USA, in 2014. He is currently working toward the M.S. degree in electrical and computer engineering at the University of Illinois at Urbana-Champaign, Urbana, IL, USA.

Kathleen M. Gegner (S'12) received the B.S. degree in electrical engineering from the University of Nebraska-Lincoln, Lincoln, NE, USA. She is currently working toward the M.S. degree in electrical and computer engineering at the University of Illinois at Urbana-Champaign, Urbana, IL, USA.

Ti Xu (S'13) received the B.S. degree in 2011 from Tsinghua University, Beijing, China, and the M.S. degree from the University of Illinois at Urbana-Champaign, Urbana, IL, USA, in 2014, where he is currently working toward the Ph.D. degree in electrical and computer engineering.

Komal S. Shetye (S'10–M'11) received the B.Tech. degree from the University of Mumbai, Mumbai, India, in 2009, and the M.S. degree in electrical engineering from the University of Illinois at Urbana-Champaign, Urbana, IL, USA, in 2011. She is currently a Senior Research Engineer with the Information Trust Institute, University of Illinois at Urbana-Champaign.

Thomas J. Overbye (S'87–M'92–SM'96–F'05) received the B.S., M.S., and Ph.D. degrees in electrical engineering from the University of Wisconsin-Madison, Madison, WI, USA. He is currently the Fox Family Professor of Electrical and Computer Engineering at the University of Illinois at Urbana-Champaign, Urbana, IL, USA.