TEXAS A&M UNIVERSITY
Engineering

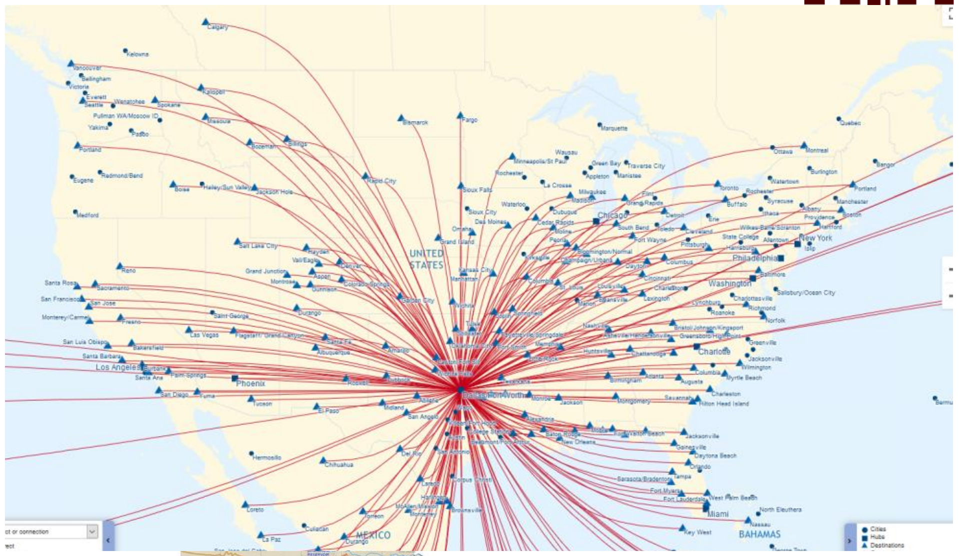
Synthetic Datasets for Energy Infrastructure Modeling

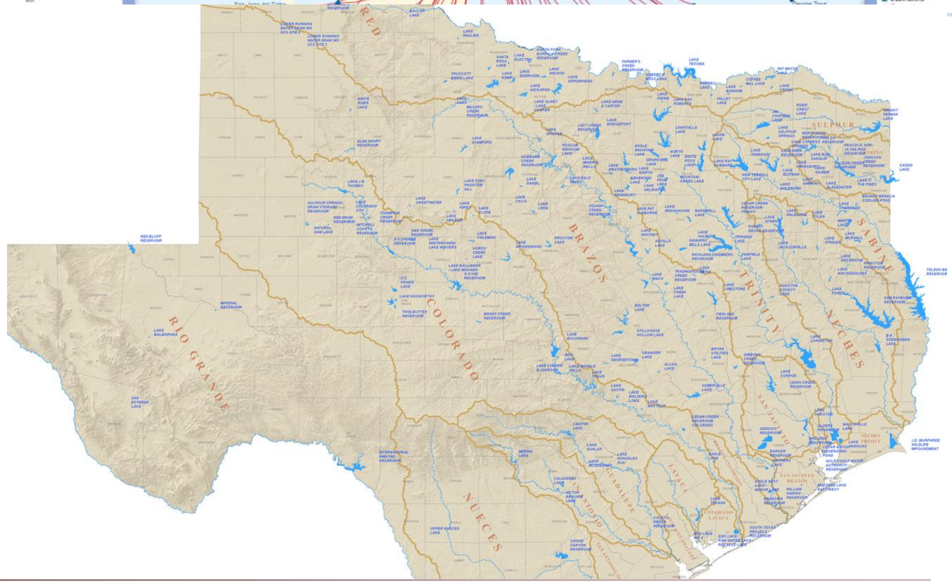
Adam B. Birchfield
Assistant Professor
Texas A&M University

PSERC Webinar
January 25, 2023

Real-World Complex Networks

- A complex network is a large graph: collection of nodes (vertices) connected by links (edges) that represents some real-world interactions
- Examples: Interstate highways, movie co-stars, brain neurons, water drainage, Facebook, airline routes, the internet, the power grid
- Aspects of network structure provide insight for assessing system resilience and designing for greater robustness and efficiency
- For each type of network, what network characteristics or metrics do they have in common with other types, and what characteristics are unique? What sort of network is the power grid?





Sources: American Airlines, Texas Water Development Board

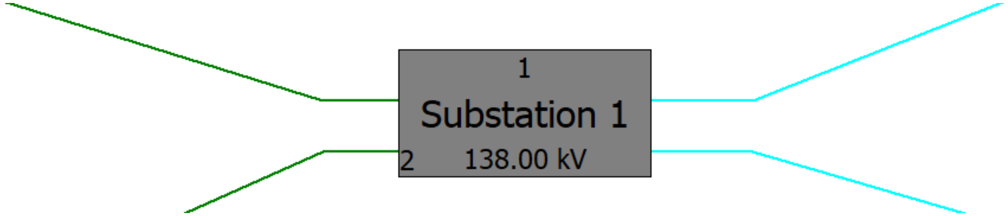
The U.S. Electric Grid Network

- Vertices: electric substations or circuit nodes/buses
- Edges: transmission lines, transformers
- Intuitive characteristics:
  - Carefully designed
  - Interconnected
  - Geographically constrained
  - Multi-level (MV, HV and EHV)
  - Robust to contingent outages

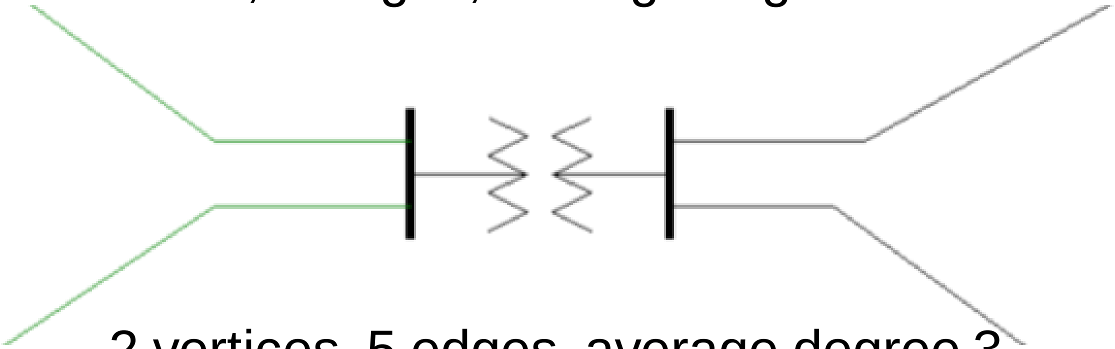
Source: EIA

Modeling Considerations

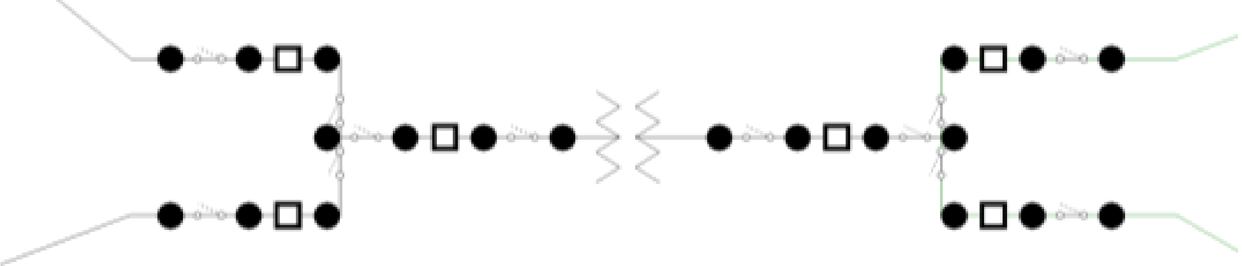
- Power grid statistics can be contradictory in the complex network literature
- Some explanations for this:
  - Different data sources and the fact that electric grids in different places have different design practices—not all electric grids are alike
  - Modeling decisions: substation-line, bus-branch, node-breaker
- To avoid these issues, in this presentation we look primarily at substation-line topologies



1 vertex, 4 edges, average degree 4



2 vertices, 5 edges, average degree 3



20 vertices, 23 edges, average degree 2.1

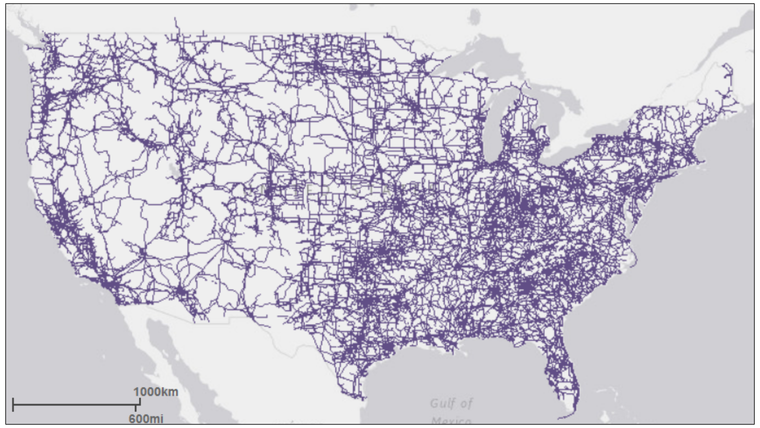
A. B. Birchfield and T. J. Overbye, "Planning sensitivities for building contingency robustness and graph properties into large synthetic grids," *HICSS*, Jan. 2020.

Some Complex Network Properties

Symbol	Description	EI Value	WECC Value	ERCOT Value	Scaling Pattern
n	Number of nodes	36,187	9398	3827	
$\bar{d}$	Average node degree (how many edges on each node)	2.61	2.58	2.61	Independent of system size
$\bar{c}$	Average clustering coefficient (how likely are neighbors to be mutually interconnected)	0.044	0.058	0.032	Independent of system size
$\bar{\ell}$	Average shortest path length between any two points	29.2	18.9	14.2	Approximately quadratic
$\bar{b}$	Average betweenness centrality (what fraction of shortest paths pass through a node)	0.083	0.21	0.4	Approximately inverse quadratic

Geometric Properties – Crossings

- Location-aware metrics can give additional insight
- One of these is graph crossings, which measures the number of times two edges cross without intersecting.
  - For some networks (waterways) this never occurs (planar)
  - For others (airplane routes) it is extremely common.
  - Electric grids are somewhere in-between
- The crossings analysis given on the next slide utilizes data from the U.S. Energy Information Administration (EIA) which gives right-of-way paths for each high-voltage transmission line in the U.S., plus information such as the nominal voltage class.
- This analysis focuses on each voltage class independently.



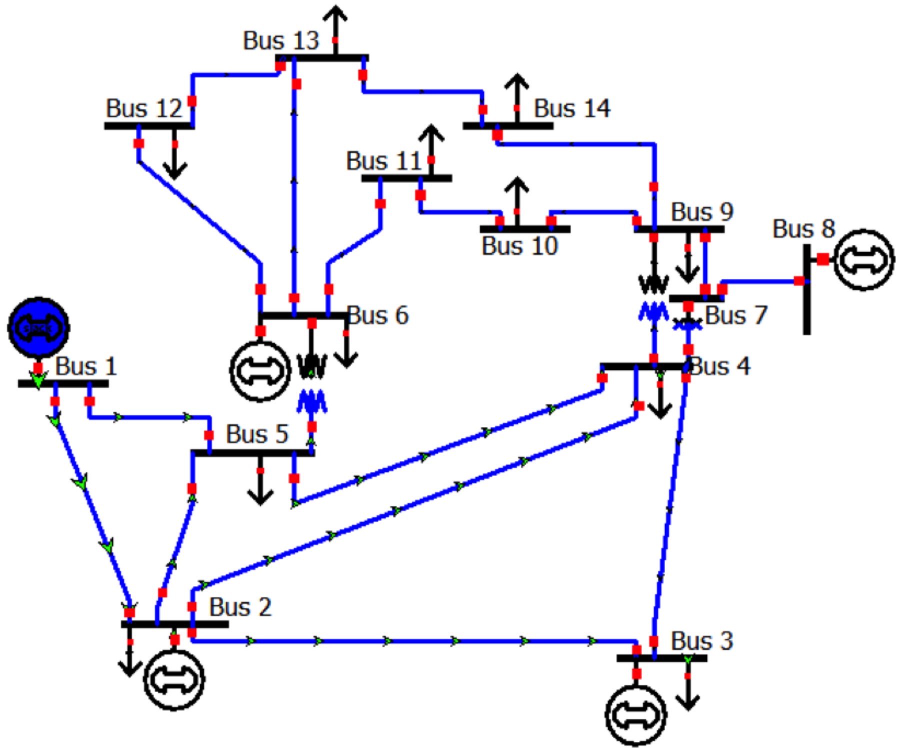
Graph Crossing Statistics

- Most high-voltage networks have crossing numbers around 30% of the number of lines. The very highest and lowest voltage levels tend to be lower. When only considering the straight-line path between the substations (as opposed to the reported actual right-of-way) the crossing numbers reduce to 10-15%.

Voltage class	Number of substations	Number of lines	Number of lines, no parallel	Crossings, straight-line		Crossings, right-of-way		Crossings, no parallel	
				Number	% of lines	Number	% of lines	Number	% of lines
765 kV	40	42	42	1	2.4	1	2.4	1	2.4
500 kV	529	732	596	67	9.2	219	29.9	162	27.2
345 kV	1526	2171	1778	297	13.7	628	28.9	563	31.7
230 kV	4648	6233	5109	935	15.1	2266	36.4	1977	38.7
161 kV	2633	3172	2858	405	13.0	952	30.0	845	29.6
138 kV	8611	10684	9129	1617	15.3	2951	27.6	2658	29.1
115 kV	12826	15031	13161	1485	10.2	3137	20.9	2734	20.8
100 kV	894	1595	1002	118	7.5	282	17.7	215	21.5
69 kV	8022	8022	7271	289	3.7	618	7.7	582	8.0

Application to Building Synthetic Grids

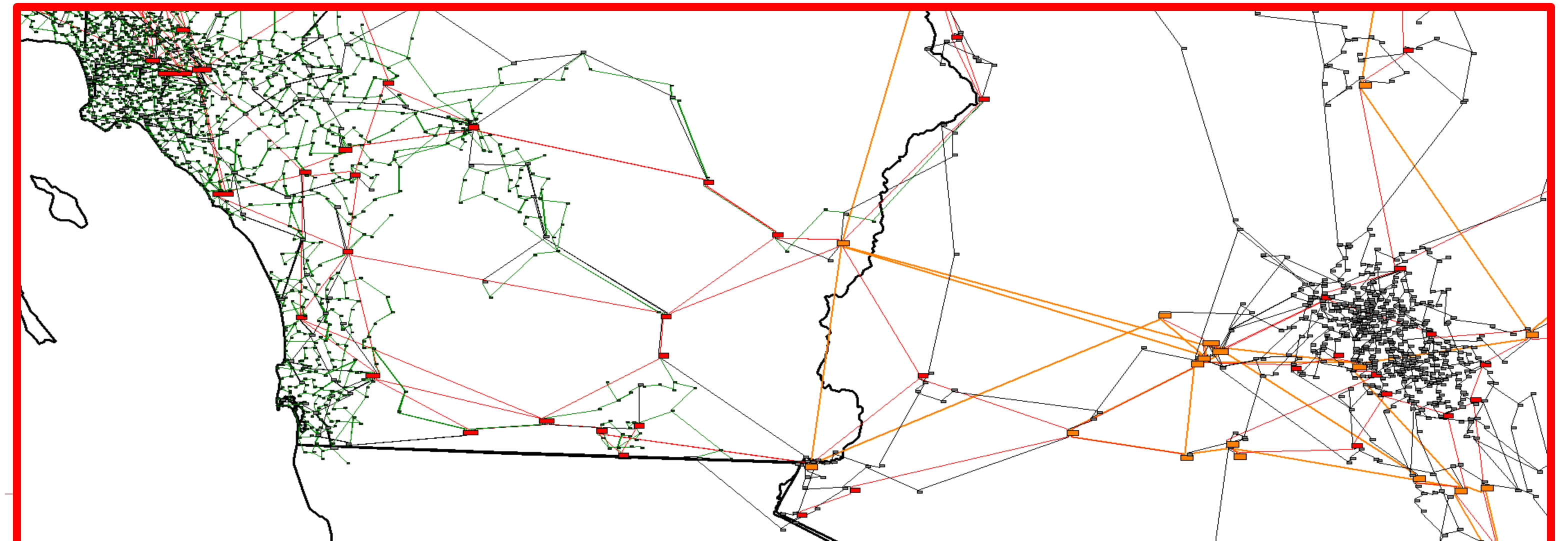
- One application of collecting complex network characteristics of electric grids is to reproduce these properties in synthetic grids.
- Power grid data is critical energy infrastructure information (CEII)
- Existing test cases—prior to synthetic grids—are small, simple, and out of date
- Goal of building synthetic power grids is to drive innovation by providing test cases that are large, complex, realistic, and fully public.

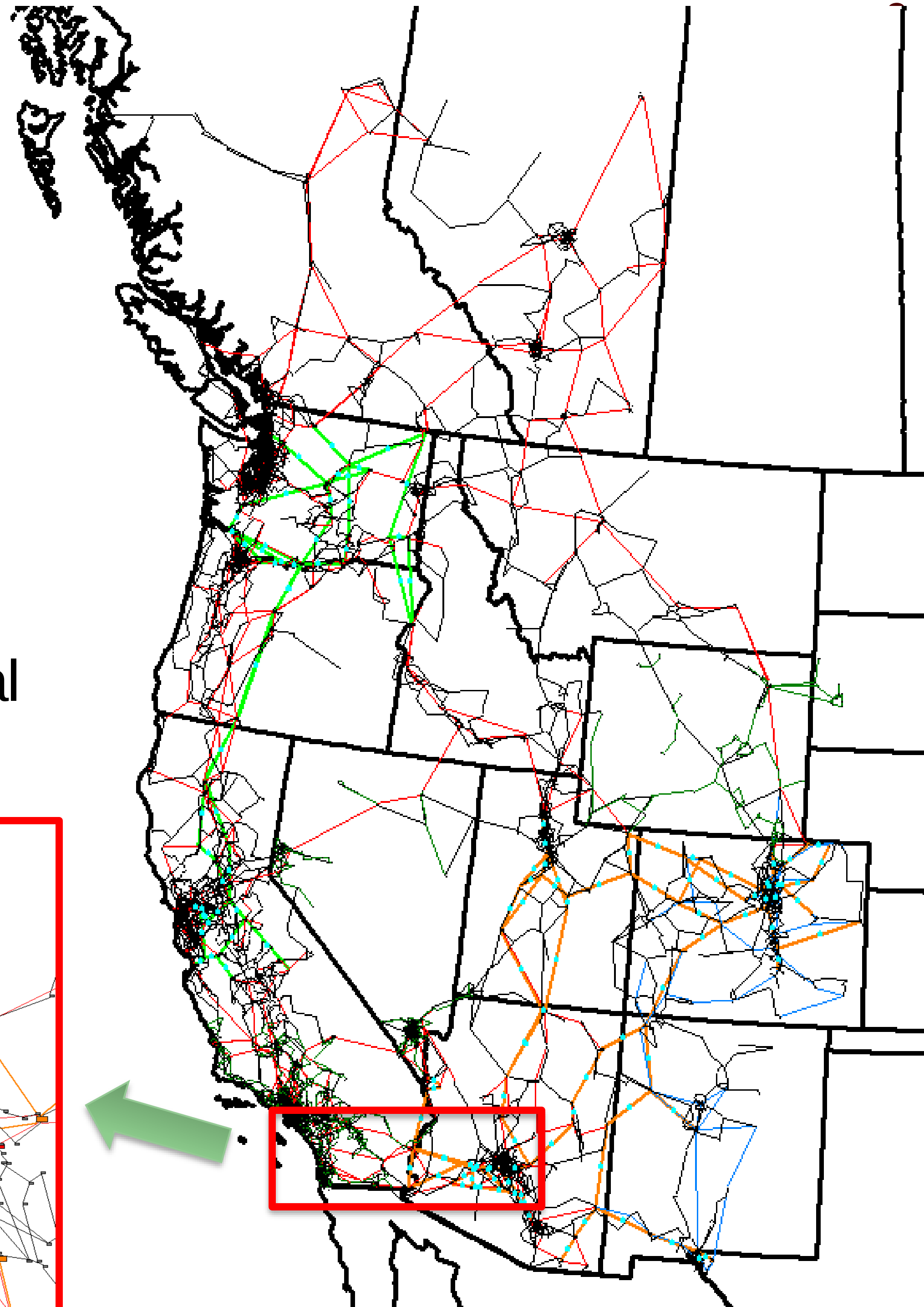


Some existing test cases, such as the IEEE 14-bus (pictured) and 118-bus case, despite their popularity, are known to vary significantly from actual grids.

Recent Progress in Building Synthetic Electric Grids

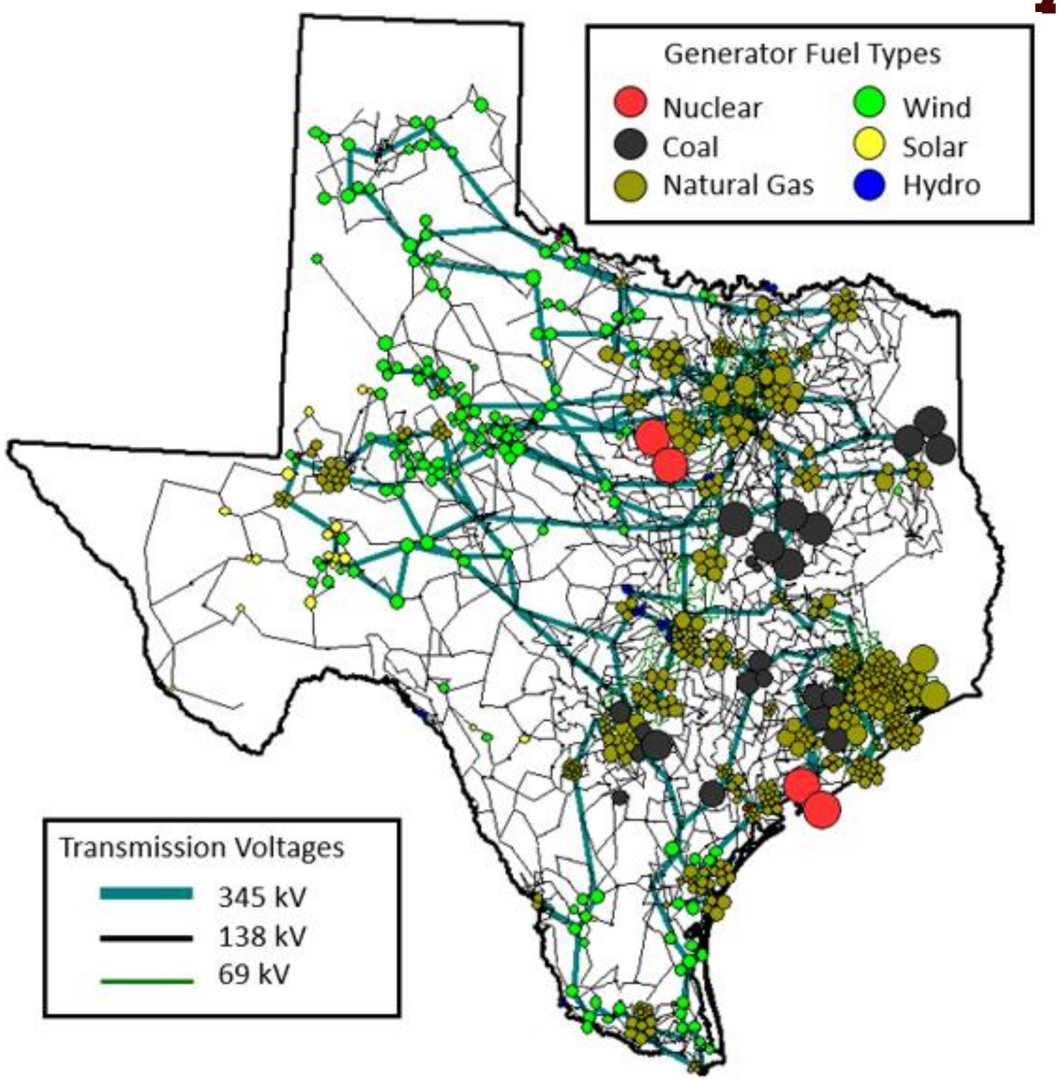
- **Large**: This case is 20,000 buses, similar to the actual Western Interconnect (WECC)
- **Complex**: Multiple kV levels, capacitors, taps
- **Realistic**: Matching a large suite of validation metrics against actual systems
- **Fully public**: It does not correspond to any actual grid or contain any CEII





Creation of Synthetic Grids Overview

- Substation Planning
  - Start with public data for generation, load
  - Cluster substations, add buses, transformers
- Transmission Planning
  - Place lines and transformers
  - Iterative dc power flow algorithm
  - Match topological, geographic metrics
  - Contingency overload sensitivity
- Reactive Power Planning
- Extensions



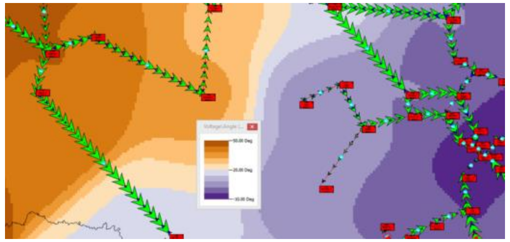
7000-bus synthetic grid on the Texas footprint

Transmission Planning Approach

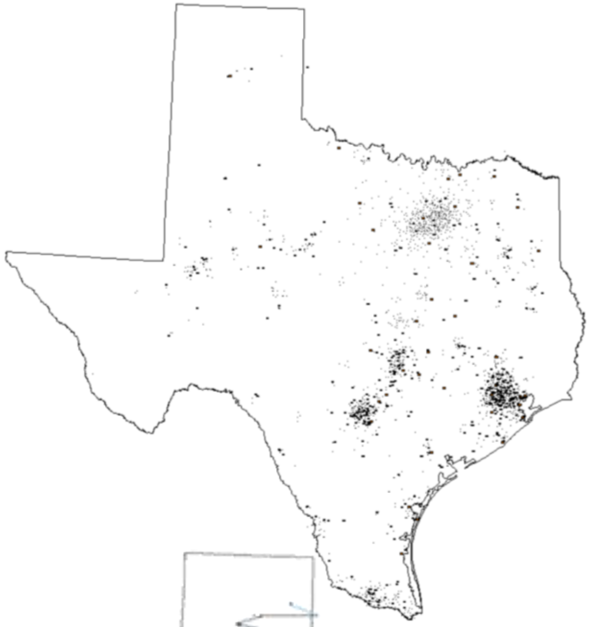
- Key Considerations
  - Geography drives transmission planning
  - Network topology parameters
  - Power flow feasibility in base and N-1 contingency conditions
  - Intractability: possible branches is n^2, possible combinations of branches is intractable
  - Many competing metrics to meet
  - Large grids have many overlapping voltage networks that connect at substations
  - Consideration of contingency conditions increases computation even more
  - Manual adjustments grow with system size

- Outline of Approach
  - Reduce search space from n^2 to $21n$ with Delaunay triangulation (up to 3rd neighbors = 99% of lines)
  - Geographic constraints by voltage level
  - Depth first search to check connectivity
  - Dc Power flow base case and N-1 contingency analysis, determine sensitivity of candidate lines to contingency overloads
  - Iterative process of random removal, analysis, targeted addition for each same-voltage subnet

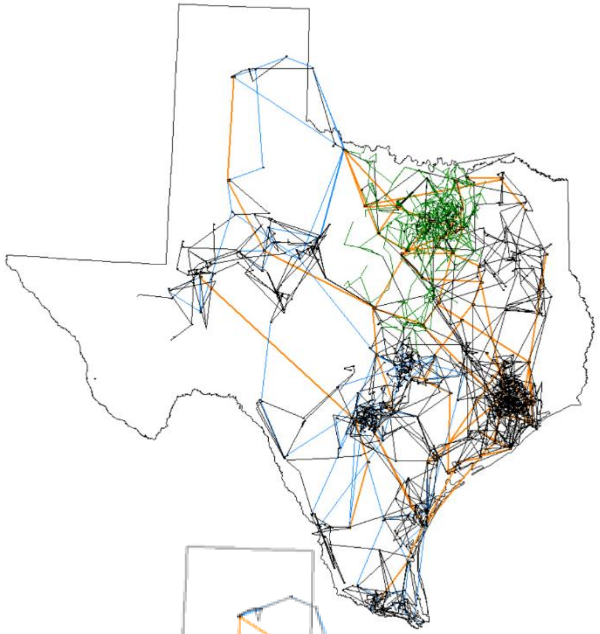
A. Birchfield and T. Overbye, "Planning sensitivities for building contingency robustness and graph properties into large synthetic grids," in *Proceedings of the 53rd Hawaii International Conference on System Sciences*, 2020.



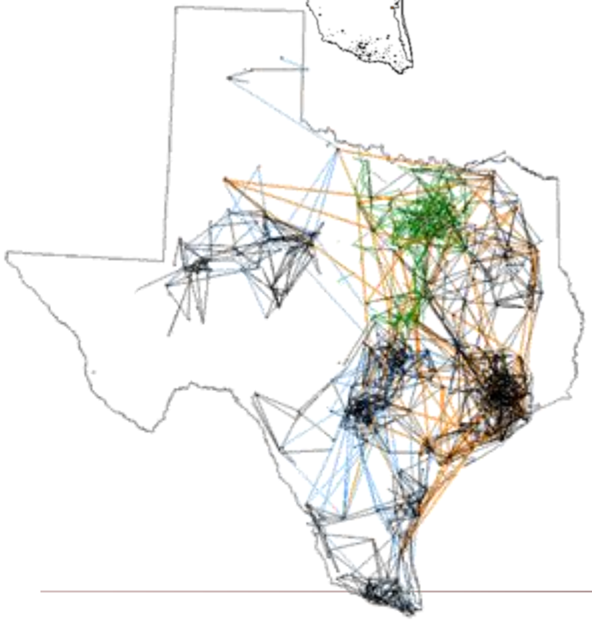
Stages of Transmission Planning Process



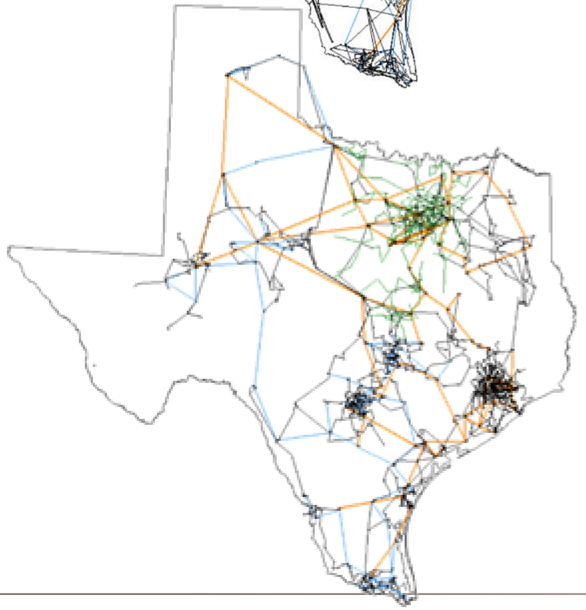
(1) The starting point is the geographic placement of substations



(3) After 100 iterations of random removal followed by targeted addition, the grid begins to match more geographic and reliability constraints



(2) The grid is initialized with a random subset of 1.2n of the 21n candidate transmission lines



(4) After 10,000 iterations, nearly all reliability and geographic constraints are met together.

N-1 Security Planning Sensitivity Calculations

- We have a sensitivity metric that is used for ranking candidate lines, based on assumptions of the dc power flow formulation with the B-bus matrix (approximately the imaginary component of the Y-bus matrix):

$$\bar{P} = \boldsymbol{B} \cdot \bar{\theta}$$

- Now, differentiate this equation with respect to B (small change in impedance)

$$0 = \boldsymbol{B} \cdot \left(\frac{d\bar{\theta}}{\boldsymbol{dB}} \right) + \bar{\theta}$$

$$d\bar{\theta} = -\boldsymbol{B}^{-1} \cdot (\boldsymbol{dB}) \cdot \bar{\theta}$$

- Now, assume the change in B is just adding some admittance (dB_i) to a right-of-way i between two buses indicated by the vector e_i (vector is zeros except 1 at from bus and -1 at to bus). So, the matrix changes in four places,

$$\boldsymbol{dB} = \bar{e}_i \bar{e}_i^T dB_i$$

- Substituting,

$$d\bar{\theta} = -\boldsymbol{B}^{-1} \bar{e}_i \bar{e}_i^T \bar{\theta} dB_i$$

- Now we have the sensitivity of the angle vector to a small admittance change in right-of-way i.

N-1 Security Planning Sensitivity Calculations, cont.

$$d\bar{\theta} = -\boldsymbol{B}^{-1}\bar{e}_i\bar{e}_i^T\bar{\theta}dB_i$$

- Now, the derivative of the phase angle difference across some other right-of-way k is

$$d\phi_k = \bar{e}_k^T d\bar{\theta}$$

$$\frac{\partial\phi_k}{\partial B_i} = -\phi_i\bar{e}_i^T\boldsymbol{B}^{-1}\bar{e}_k$$

- (Note that the rearranging to switch i and k is allowed because B is assumed to be symmetric.)
- The result is a computationally efficient method for calculating the sensitivity of phase angle difference and hence the power flow in line k to the addition of some admittance to the right-of-way i:
- Calculate sensitivity to overload (e_{mon} has 1 at from bus and -1 at to bus of monitored)

$$\bar{s} = \boldsymbol{B}^{-1}e_{mon}$$

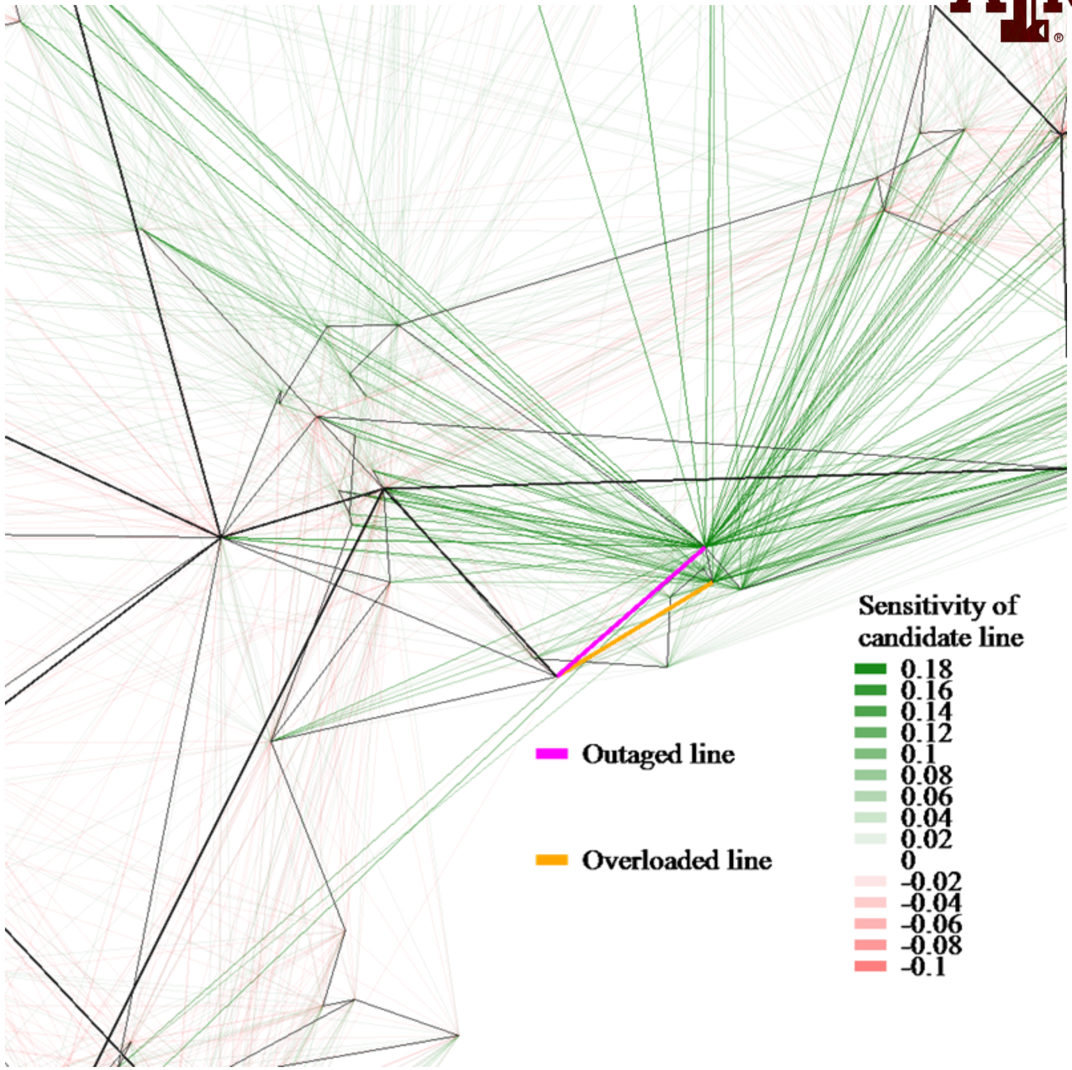
- Now each candidate line sensitivity is

$$s_{mon\to add} = (\bar{s}_{add1} - \bar{s}_{add2}) \cdot (\theta_{add1} - \theta_{add2})$$

- Computational order is linear with number of candidate lines, so we can check all our potential lines for which ones would contribute best to N-1 security

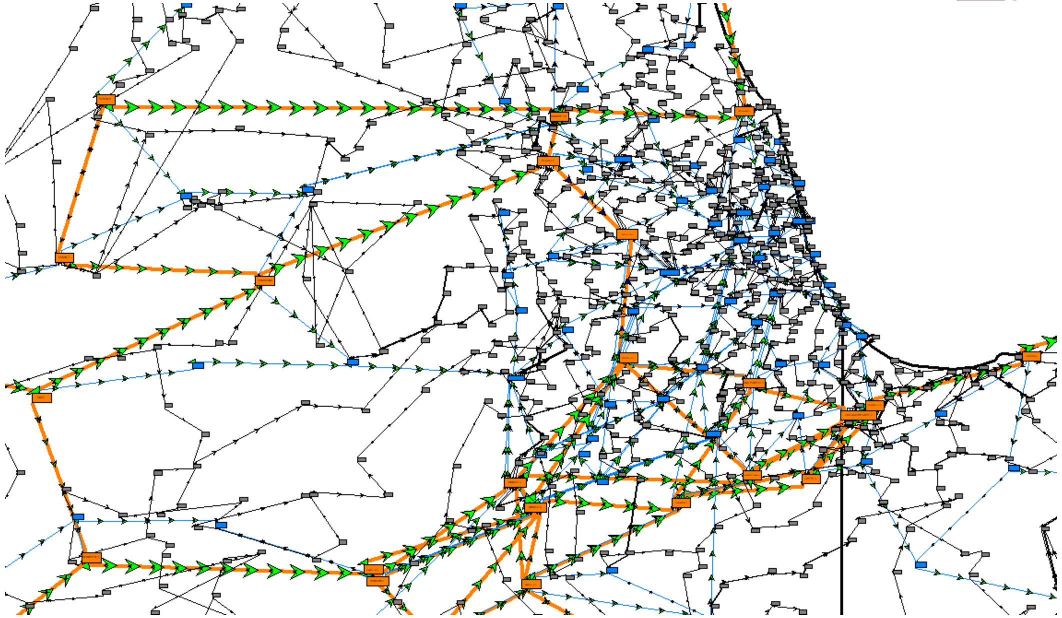
Line Sensitivity Results

- This is part-way through the transmission grid creation process
- Black lines are currently in the network
- We're focusing on a contingency of the magenta line causing an overload in the orange line
- Possible candidates are shown in red, gray, and green
- The darkest green lines are the most helpful to improving this contingency outage
- This metric is balanced against other considerations, such as line length



Synthetic Grid Validation

- Statistics collected from 3 actual interconnects and 12 subset cases, from FERC 715 data
- Because of the variety in engineering design and practice, actual grids are quite diverse
  - Challenge is to capture distribution and realistic ranges
  - Thresholds are set that should not be violated unless justified with an engineering design choice. Thus this is a screening process.
- Categories of metrics
  - Size and structure: ratio and proportions of elements
  - Parameter distributions and correlation
  - Topological network structure
  - Power flow metrics and voltage control devices



Zoomed-in view of Chicago metro area in 70,000 bus case. Constraints include Lake Michigan and the Chicago metro area. Orange is 500 kV, blue is 230 kV, and gray is 115 kV.

Validation Example: Complex Network Properties

Metric	Actual Systems			Synthetic Systems		
	EI	WECC	ERCOT	70K	20K	5000
n	36,187	9398	3827	34,999	9524	2941
$\bar{d}$	2.61	2.58	2.61	2.74	2.67	2.71
$\bar{c}$	0.044	0.058	0.032	0.048	0.034	0.031
$\bar{\ell}$	29.2	18.9	14.2	36.7	20.3	13.8
$\bar{b}$	0.083	0.21	0.40	0.11	0.22	0.50

n : number of substations

$\bar{d}$: average node degree

$\bar{c}$: clustering coefficient

$\bar{\ell}$: average shortest path

$\bar{b}$: average betweenness centrality measure (%)

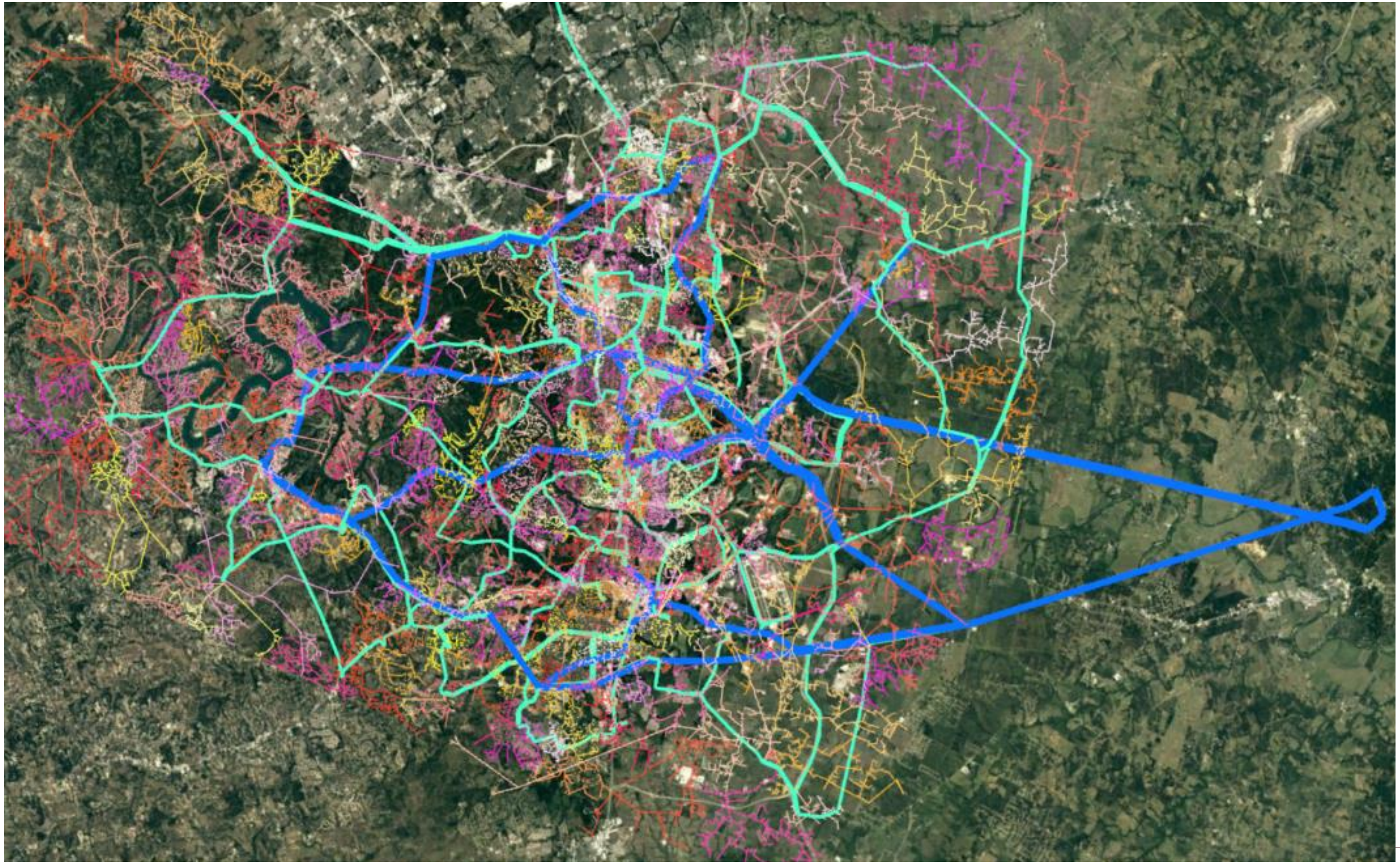
Recent Applications and Extensions of Synthetic Grids

1. Transmission & Distribution Co-Simulation

2. Electric Grid & Natural Gas Pipeline Co-Simulation

3. Dynamics and Stability

Combined Transmission and Distribution Grids

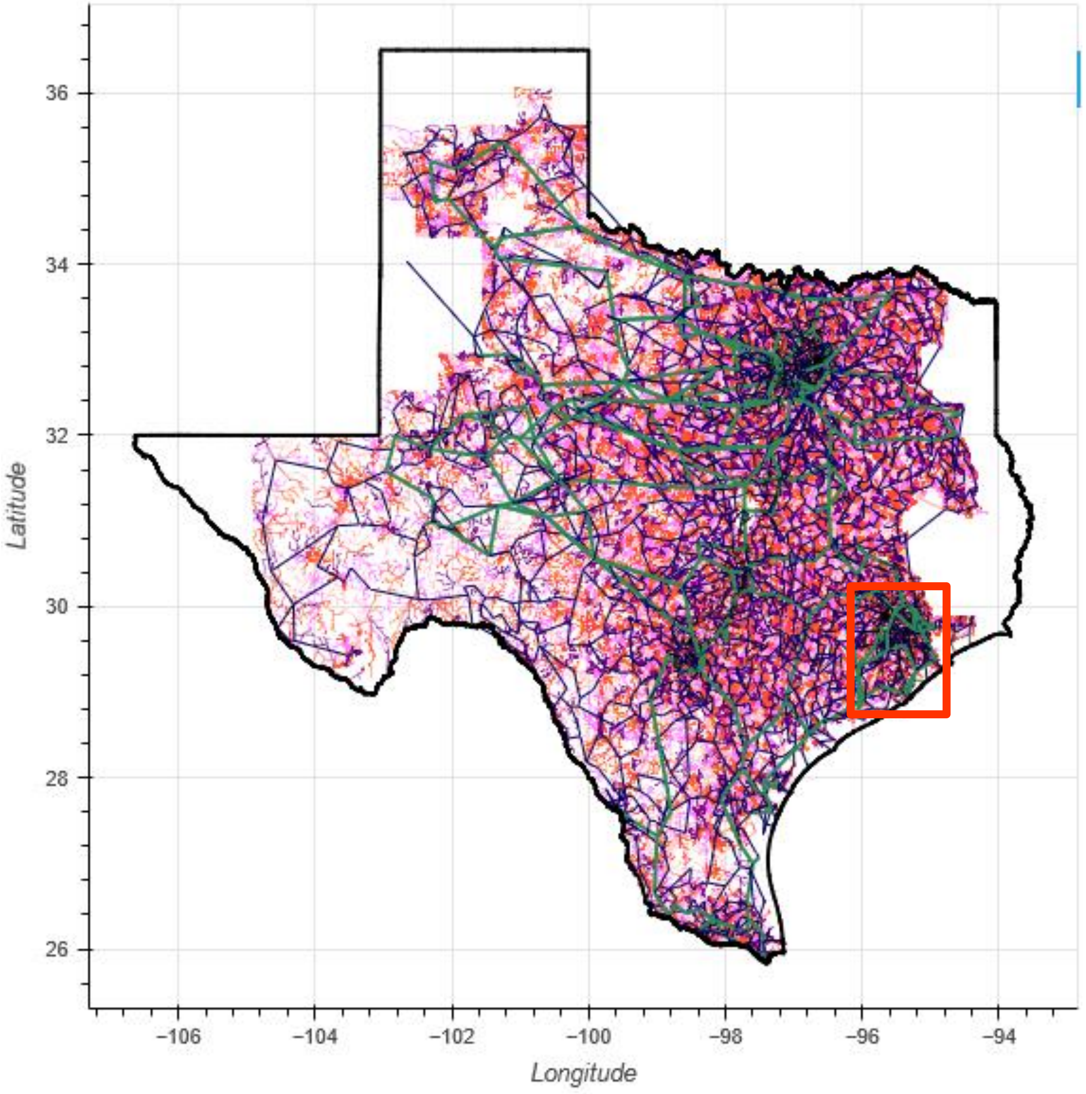


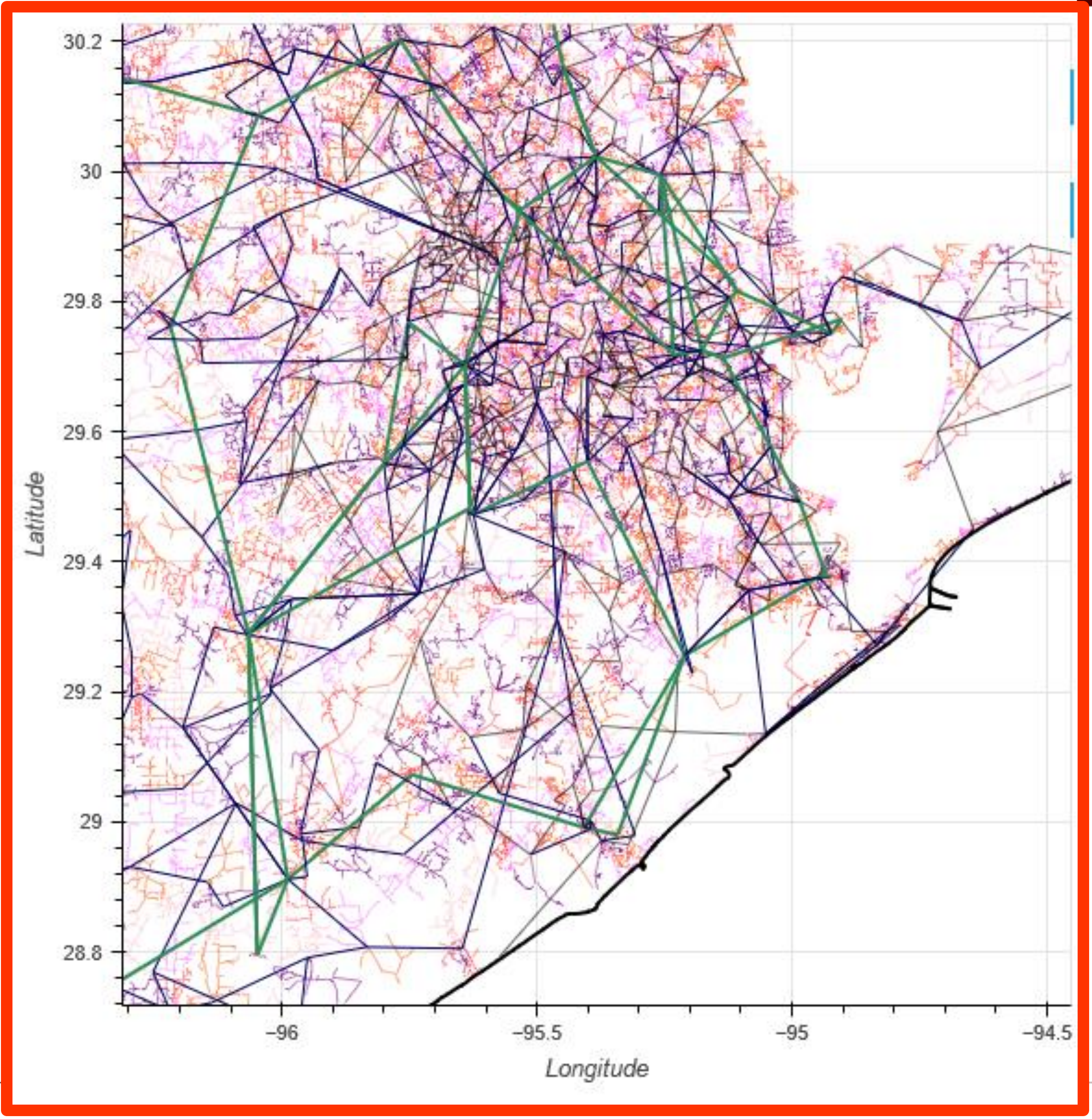
TAMU + NREL work has developed down-to-the-meter synthetic grids

Parcel data is used to seed the algorithms for distribution.

The figure shows the transmission system (blue is 230 kV and cyan 69 kV) and the distribution system modeled down to 307,000 meters. The distribution data is in the OpenDSS format.

We Now Have All of Texas Modeled – Several Million Nodes!





What We Can Do with T&D Models

- Resilience studies to natural disasters
  - Hurricanes, thunderstorms, tornados, ice storms all affect both transmission and distribution networks
  - This enables us to model the combined effect on both
- Studying distribution-connected generation
  - Customer-connected solar panels and other generation is affected by the weather at the particular locations
  - Allows us to know in more detail where geographically each transmission load serves
- Electric vehicle charging
  - We place residential EVs on the distribution grid, but in aggregate they may affect transmission

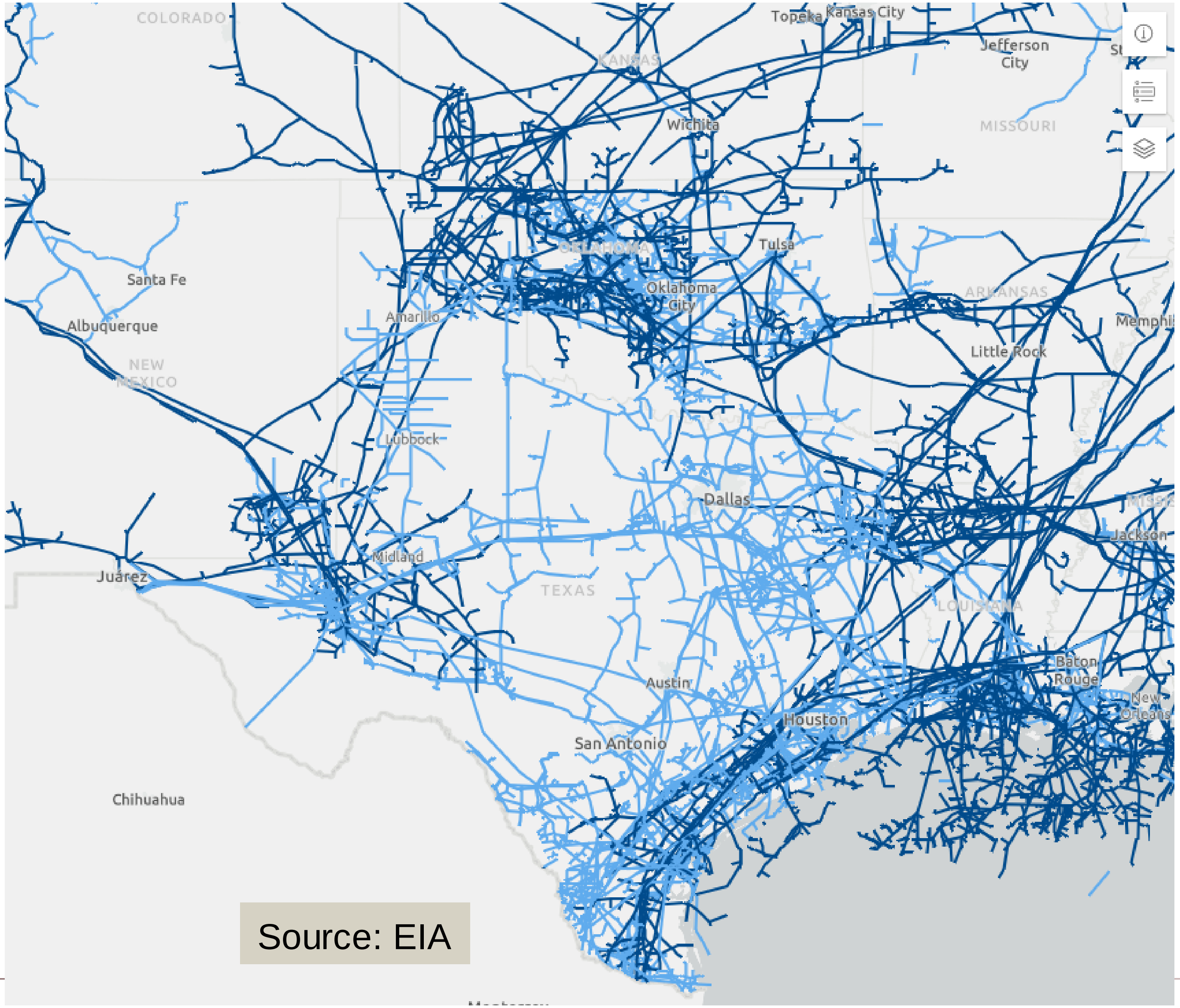
Natural Gas Pipeline Networks and Their Interaction with the Electric Grid

- Natural gas currently makes up 40% of electric generation in the US. (TX: 50%)
- Projected to increase to 2,000 billion kWh per year by 2050.
- Some natural gas equipment requires electric power to operate.
- In 2021's Winter Storm Uri in Texas, significant generation was lost due to challenges in the natural gas supply.
- Challenge in modeling two usually separate networks: different companies, different modeling paradigms.

Y. A. Abu-Khalifa and A. B. Birchfield, "Techniques for Creating Synthetic Combined Electric and Natural Gas Transmission Grids," *2022 IEEE Texas Power and Energy Conference (TPEC)*, College Station, TX, USA, 2022, pp. 1-6.

Characteristics of Pipeline Networks

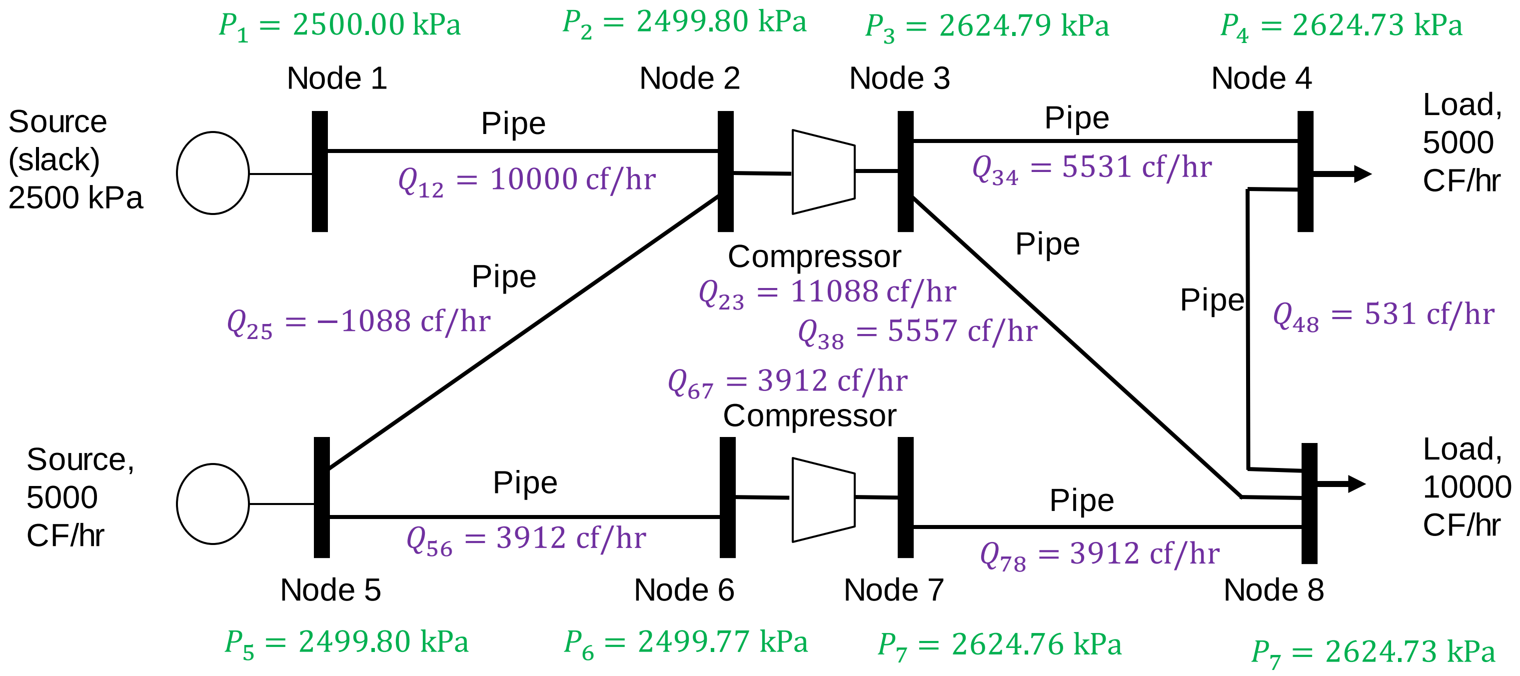
- Similar to power grids:
  - Geographically constrained
  - Carefully designed but not fully centrally planned
  - Developed over many years
- Distinct from power grids:
  - Distinction between voltage levels (EHV/HV) not as pronounced
  - Interstate-intrastate distinction
  - Generally, gas travels farther than electric power



Electric-Gas Co-Simulation Environment

- We now have the capability to solve steady-state gas flows in coordination with electric grid studies.
- Our gas flow solution involves a scalable, Python-based iterative algorithm, coupled in to PowerWorld studies of the electric grid for modeling interdependencies.

8-Bus Test Case

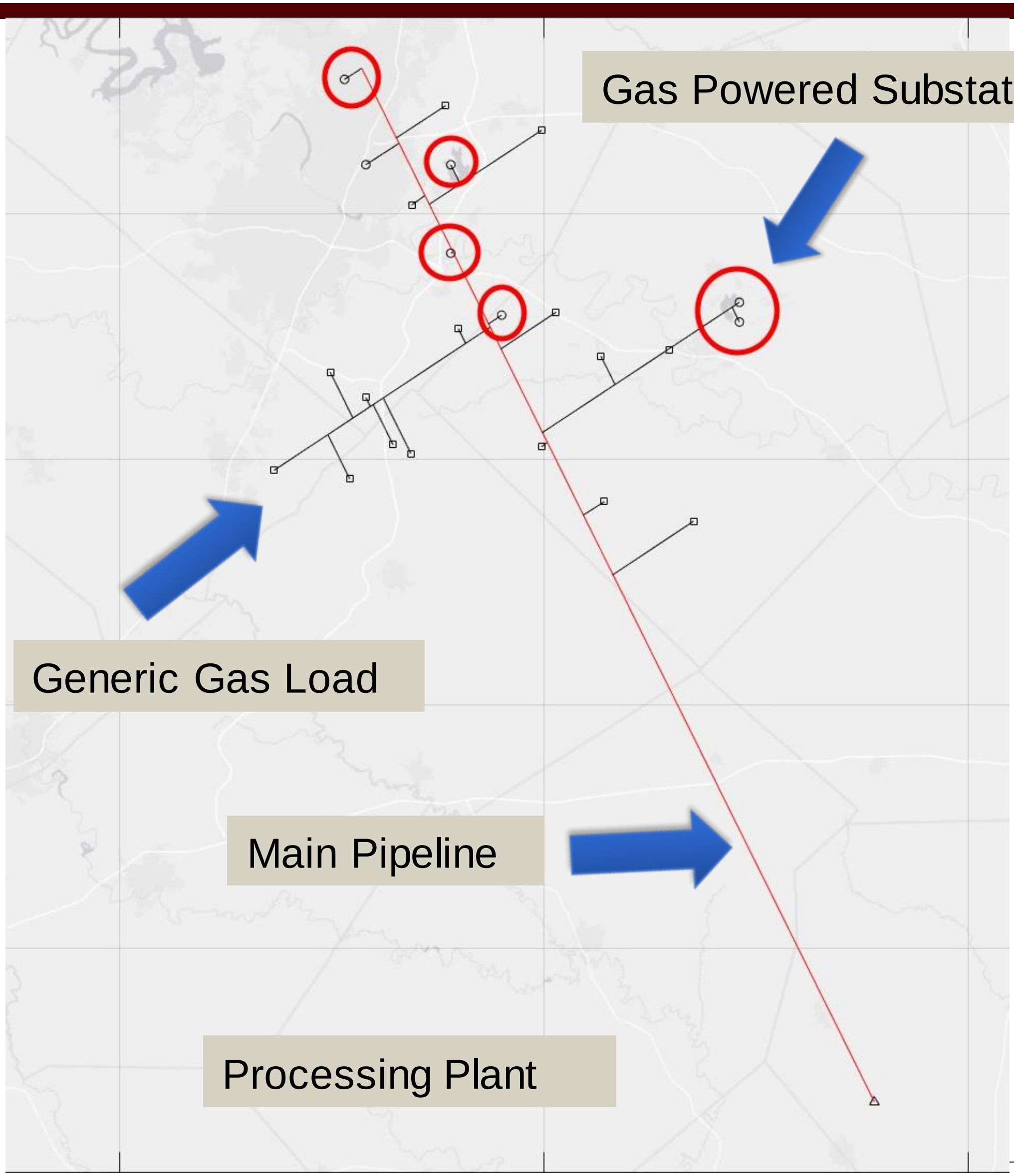


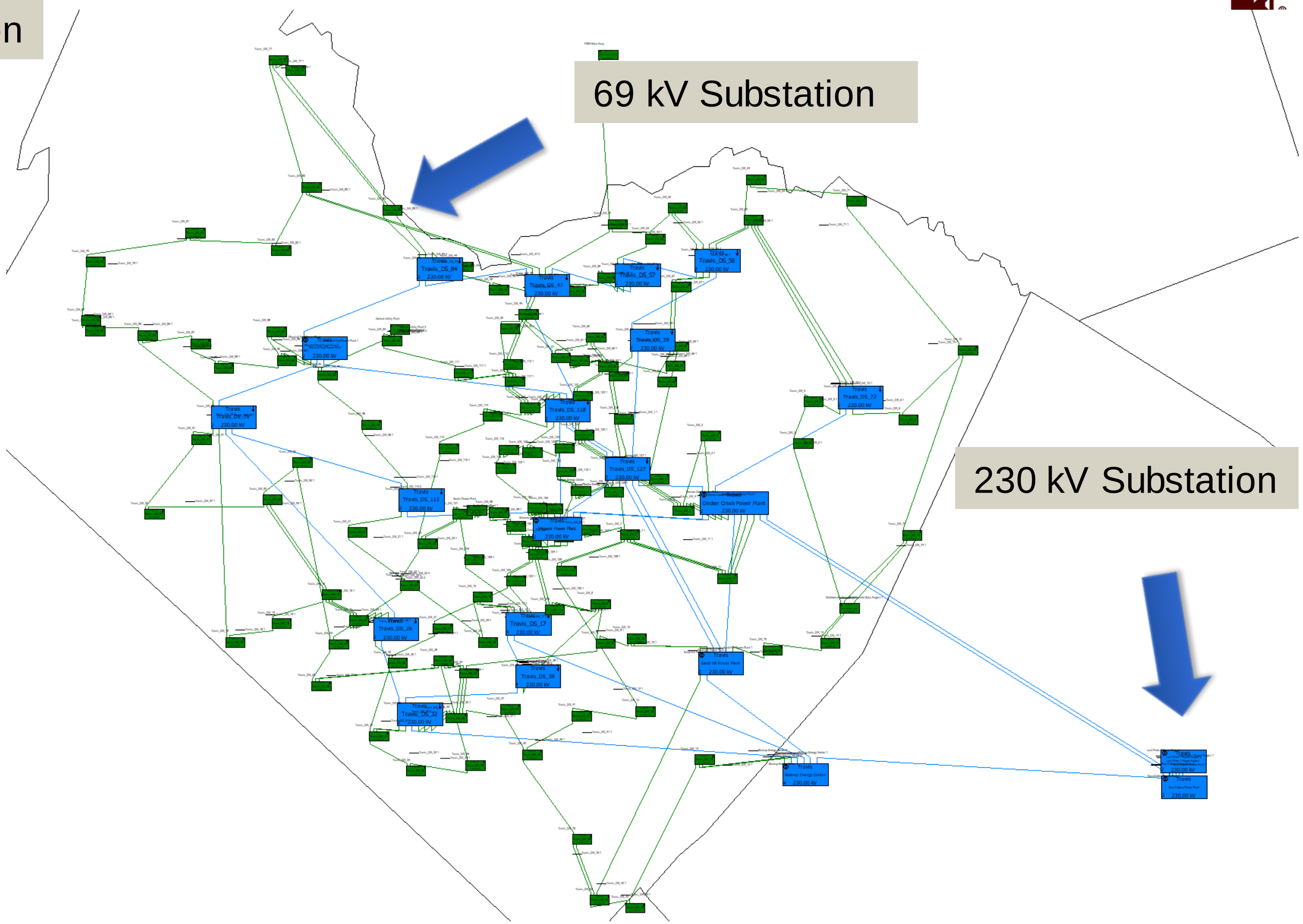
A. J. Osiadacz and K. Pienkosz, “Methods of steady-state simulation for gas networks,” System Science, vol. 19, no. 7, pp. 1311-132, 1988.

Building Synthetic Natural Gas Pipeline Networks

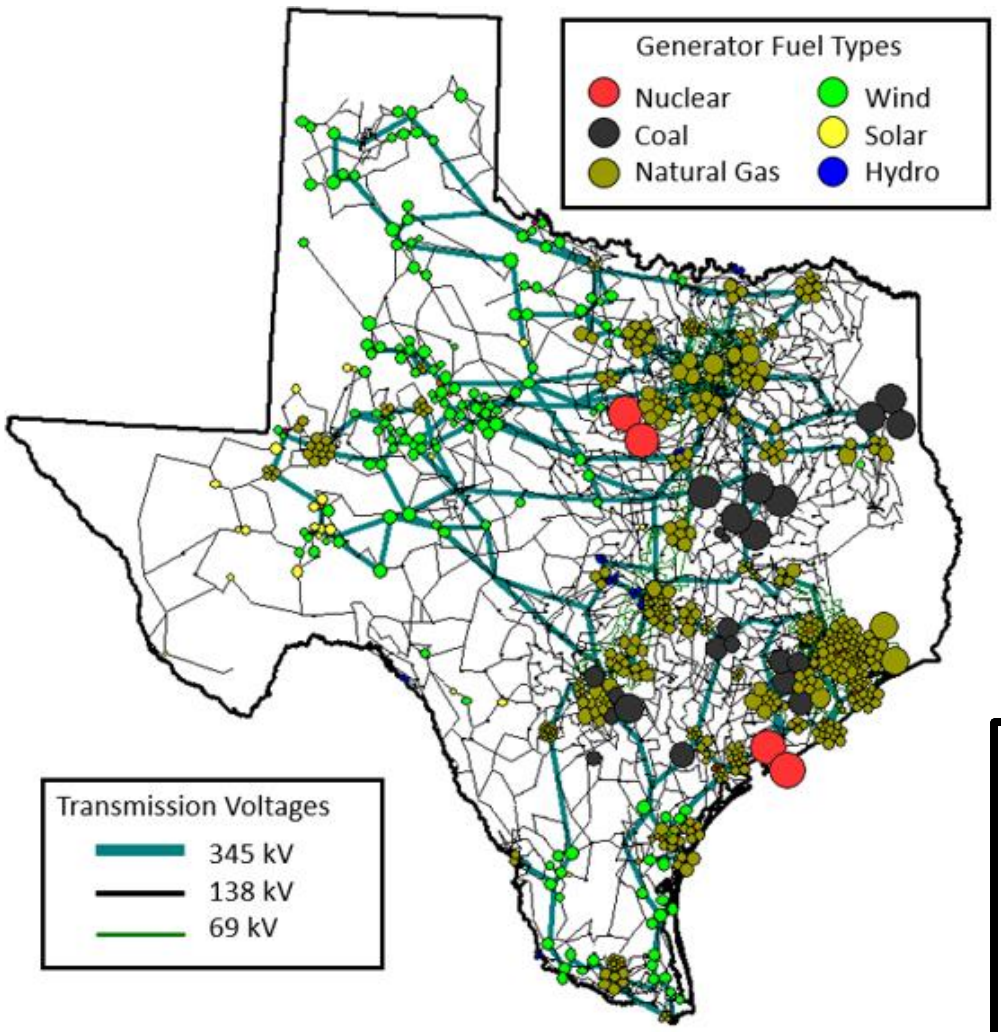
- The goal here is to contribute to existing studies by:
  - Introducing a method that automates the pipeline building process
  - Creating a test case with available data that can be used for future studies
  - Creating a geographically accurate, synthetic, combined system using publicly available data
  - A priority-queue greedy approach works quite well
- The result is a combined simulation of a 47-node gas test and a 173-bus electric system of the Travis County Area
- We have also finished a draft of a full Texas case that will couple in perfectly with our 7000-bus electric grid case
  - As of Jan 2023, a draft of the case is available including storage reservoirs and out-of-state pipeline connections, which solves in a steady-state combined systems sense

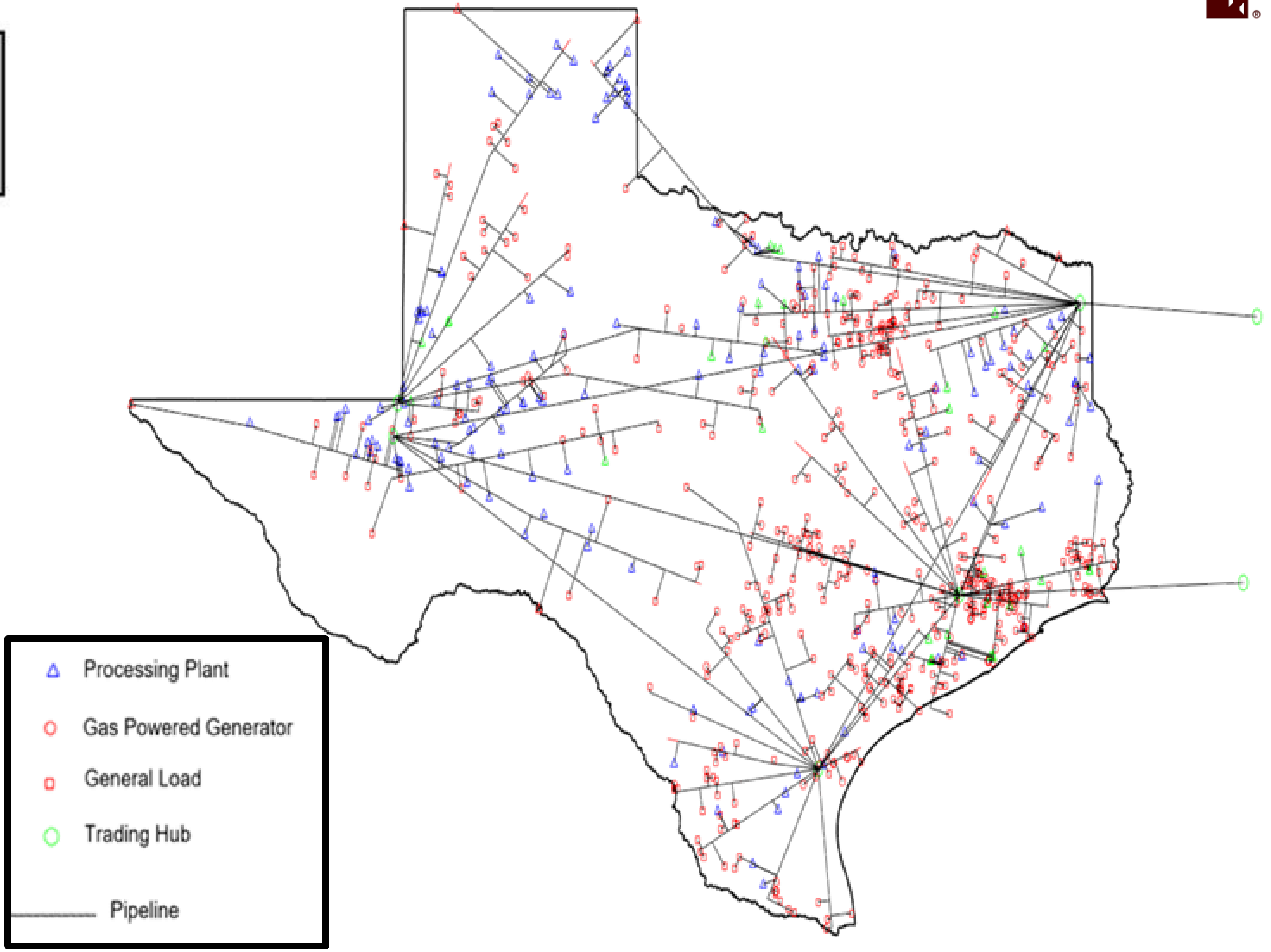
Synthetic Electric-Gas Case in Travis County, TX





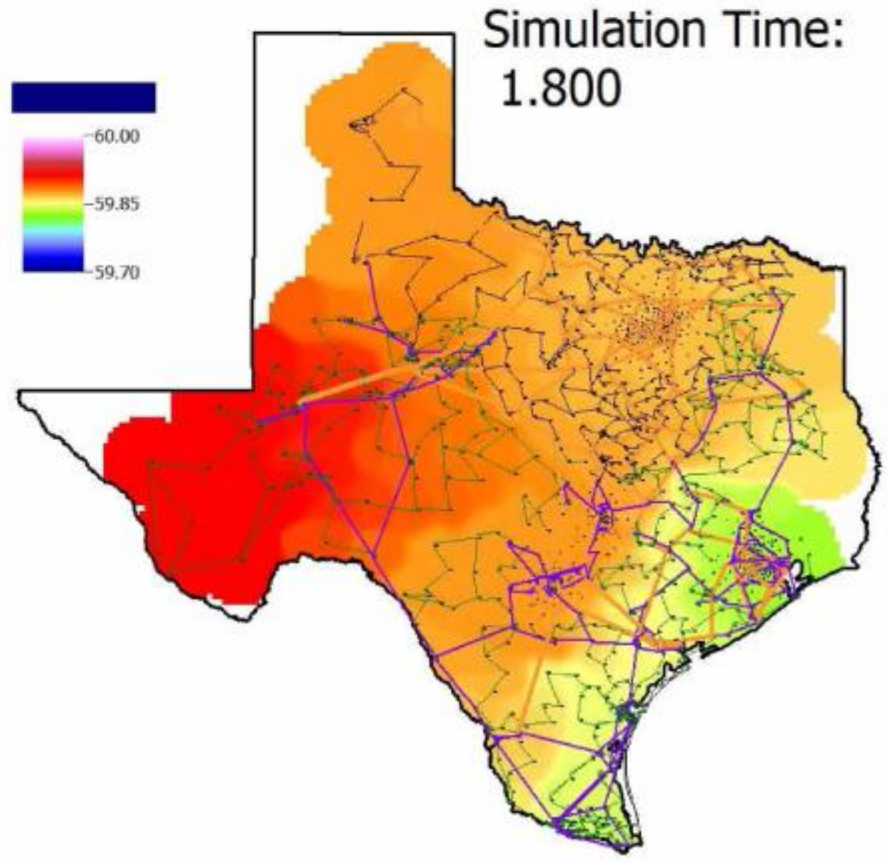
Full Texas Synthetic Combined Electric-Gas Network





Adding Dynamic Models to Synthetic Grids

- Motivation: Support and spur innovation in large transmission system stability and control research
- Objective: Show that the case is realistic
- Challenges: Tons of parameters with limited observability, everything is intertwined, and we want to match wide-area and local performance across multiple dimensions
- Prior work: Statistical analysis of model types, hand-tuning to eliminate oscillations and match global metrics



T. Xu, A. B. Birchfield, K. S. Shetye, and T. J. Overbye, "Creation of synthetic electric grid models for transient stability studies," in The 10th Bulk Power Systems Dynamics and Control Symposium (IREP 2017), 2017, pp. 1–6.

Frequency Response and Inertia

- A major wide-area effect that must be met for realistic synthetic dynamic models is frequency response, driven by the distribution of system inertia
- After a major loss of generation, instantaneous frequency response is driven by synchronous inertia
- At a system level, approximate ROCOF is related to inertia by rearranging the swing equation and summing up H's

$$\dot{\omega} = \frac{1}{2H}(T_M - T_E) = \frac{-P_{loss}}{2H}$$

$$ROCOF = \frac{-f_{base,Hz} P_{loss,MW}}{2\sum H_g S_{g,base_g}}$$

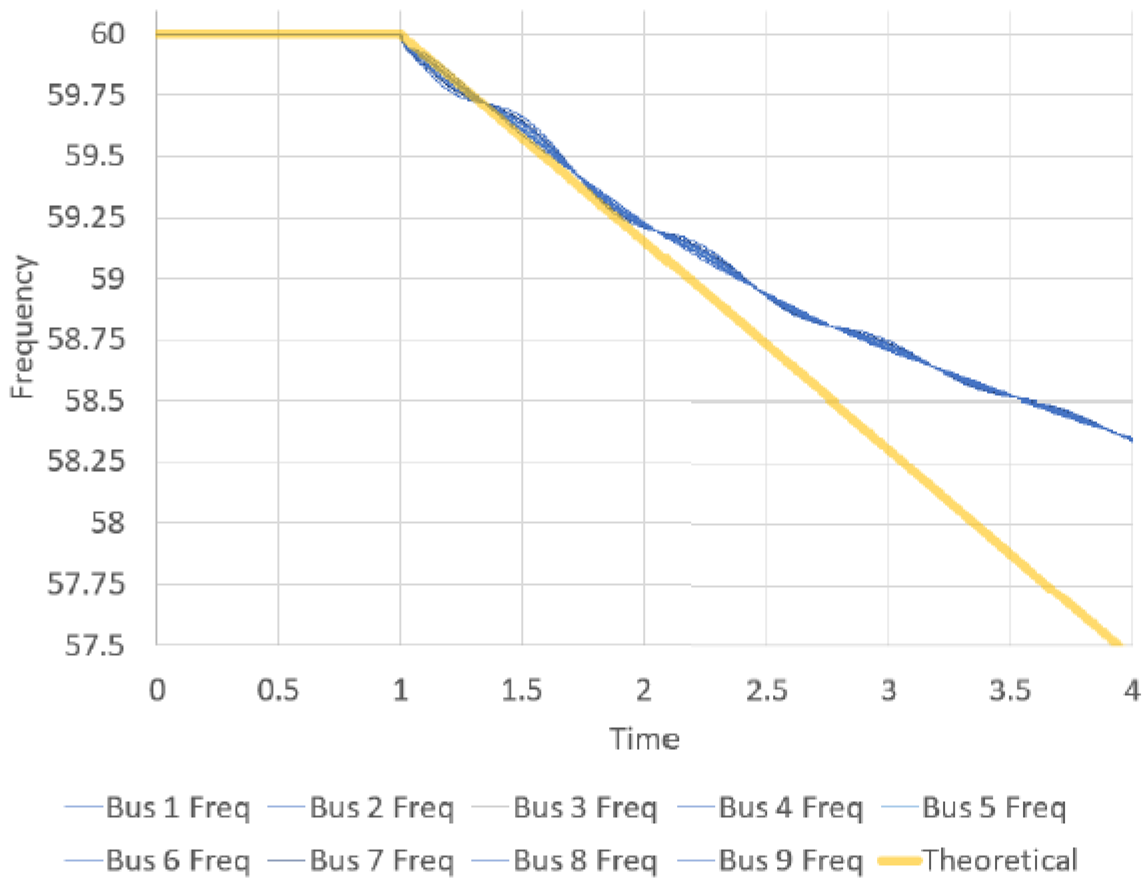
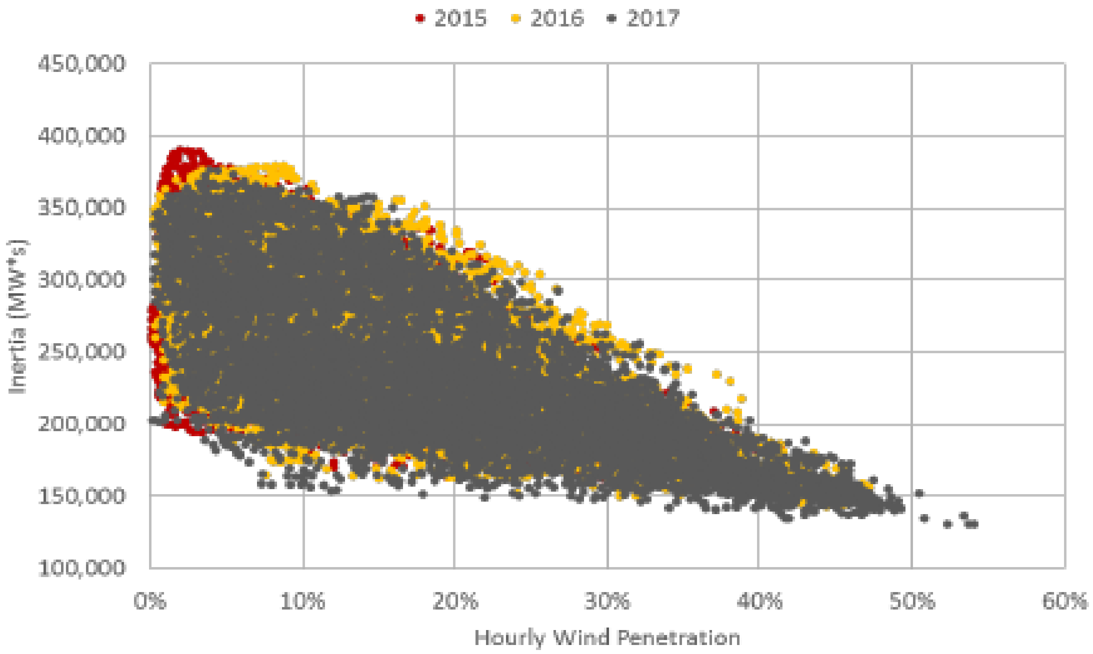


Fig. 1. Frequency response of 9-bus case to loss of generator three. Yellow trace shows theoretical response calculated by system-wide ROCOF.

Inertia Adequacy Analysis in ERCOT

- ERCOT measures system-wide inertia with $\Sigma(\mathrm{H} \cdot S_{base})$ for all online generators
- Value varies based on load, wind, dispatch; minimum in 2017 was 130 MVA-s
  - In March 2021, this got down to 116.6 MVA-s
- Dynamic simulations were performed, concluding that, with current operational practice, inertia levels can be supported down to 100 MVA-s.
  - Based on system-wide frequency decline
- Other grids (like Great Britain and Australia) are looking at regional inertia



Source: ERCOT, "Inertia: Basic concepts and impacts on the ERCOT grid," Electric Reliability Council of Texas, Austin, TX (United States), Tech. Rep., 2018.

Inertia Adequacy Measure on a Locational Basis

- Goal here is to keep the simplified zero-order inertia-ROCOF relationship, but add in network effects, represented by **Y**.
- Start with current injection model, swing equation, definition of electric torque and current injection

$$\mathbf{Y} \cdot V = I \qquad I_i = \frac{E'}{X'_d} \angle(\delta_i - \frac{\pi}{2}) \qquad T_e = Re[VI^*] \qquad \dot{\omega} = \frac{1}{2H}(T_M - T_E)$$

- Known steady-state initial conditions, then define contingency via **Y** and I and calculate new values of V, T_e, and $\dot{\omega}$.
- Now, find second derivatives $\ddot{I}$, $\ddot{V}$ (note that instantly after disturbance $\omega = \dot{\delta} = \dot{I} = \dot{V} = 0$)

$$\mathbf{Y} \cdot \ddot{V} = \ddot{I}$$

$$\frac{\partial I}{\partial \delta} = \frac{E'}{X'_d} \angle \delta_i \qquad \dot{I} = \frac{\partial I}{\partial \delta} \cdot \frac{\partial \delta}{\partial t} = \frac{\partial I}{\partial \delta} \omega$$

$$\frac{\partial^2 I}{\partial \delta^2} = \frac{E'}{X'_d} \angle(\delta_i + \frac{\pi}{2}) \qquad \ddot{I} = \frac{\partial I}{\partial \delta} \dot{\omega} + \omega \cdot (\frac{\partial^2 I}{\partial \delta^2} \omega)$$

A. B. Birchfield, "Inertia adequacy in transient stability models for synthetic electric grids," in The 11th Bulk Power Systems Dynamics and Control Symposium (IREP), 2022, pp. 1–8.

Inertia Adequacy Measure on a Locational Basis, cont.

- Finally, get second derivative of voltage angle, which is related to an approximation of bus-level ROCOF
- $V = V_r + jV_i = V_m \angle V_a$

$$\dot{V}_a = \frac{1}{1+(V_i/V_r)^2}\left(\frac{V_r\dot{V}_i - V_i\dot{V}_r}{V_r^2}\right) = \frac{V_r\dot{V}_i - V_i\dot{V}_r}{V_r^2 + V_i^2}$$

$$\ddot{V}_a = \frac{V_r\ddot{V}_i - V_i\ddot{V}_r - 2\dot{V}_r\dot{V}_i}{V_r^2 + V_i^2} - \frac{2V_rV_i(\dot{V}_i^2 - \dot{V}_r^2)}{(V_r^2 + V_i^2)^2}$$

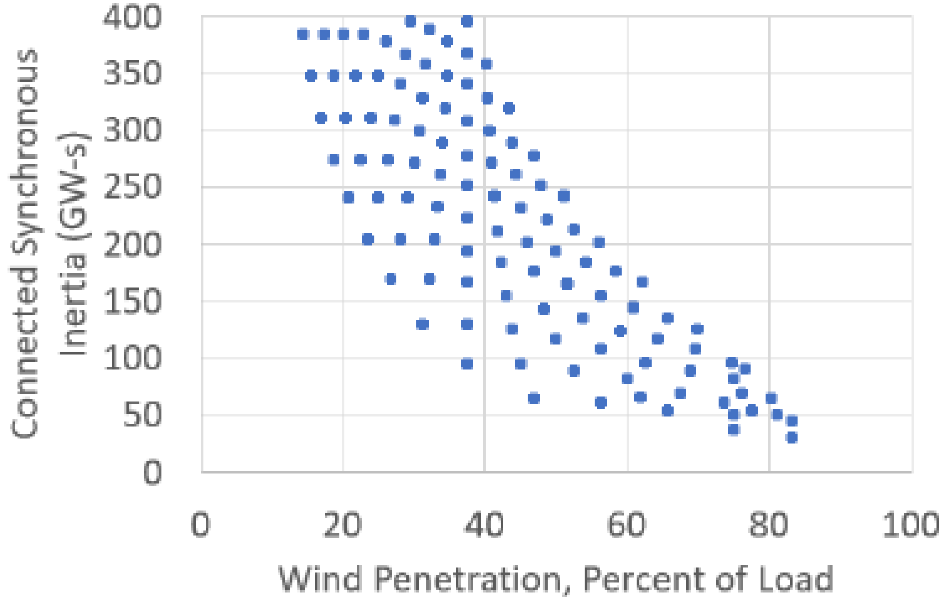
- Remembering that $\dot{V} = 0$,

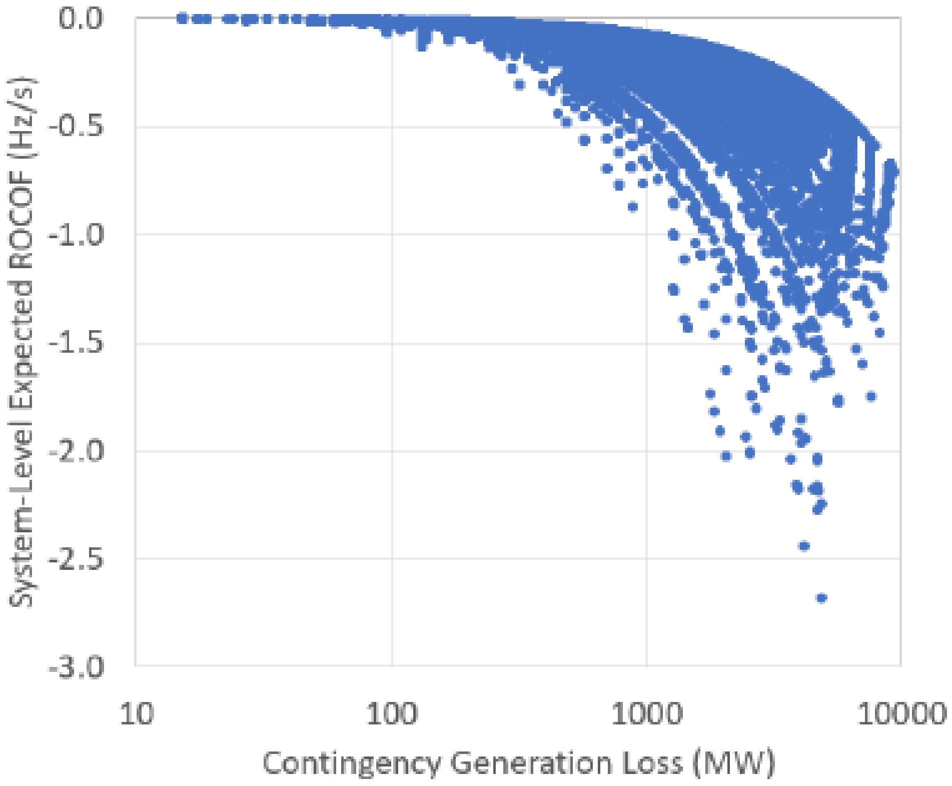
$$\ddot{V}_a = \frac{V_r\ddot{V}_i - V_i\ddot{V}_r}{V_r^2 + V_i^2}$$

This is a potential screening mechanism, with computational order faster than power flow (just two matrix solves)

20,000 Test Cases

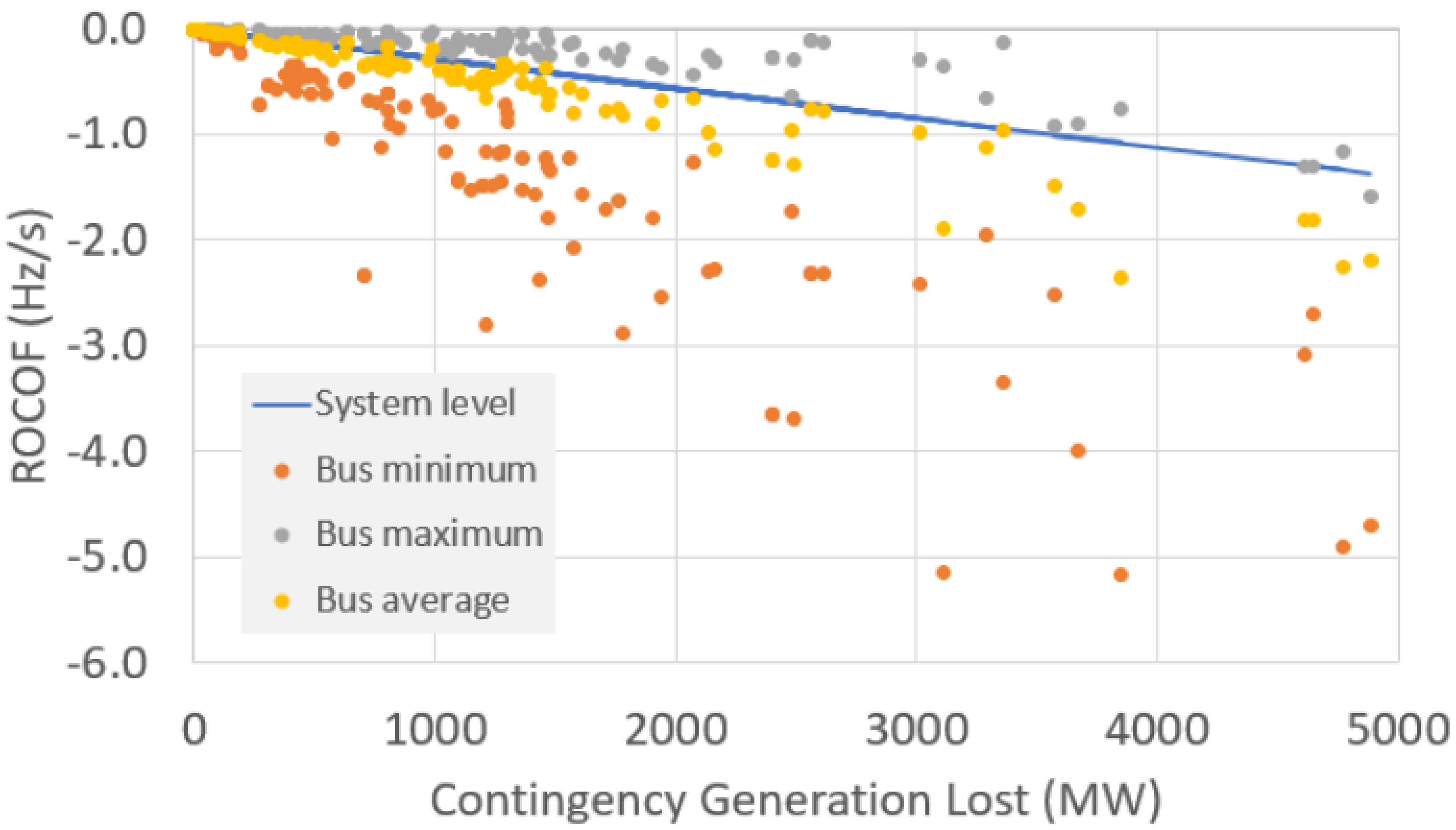
- 7000 bus synthetic case, 125 base loading cases, 163 contingencies
- Base cases cover realistic plus extreme loading (15-75 GW) and wind (10-30 GW)
- Contingencies cover loss of generation, up to benchmark event (2750 MW) and beyond, up to about 5000 MW
- Lots of variety in geographic distribution of inertia and outage





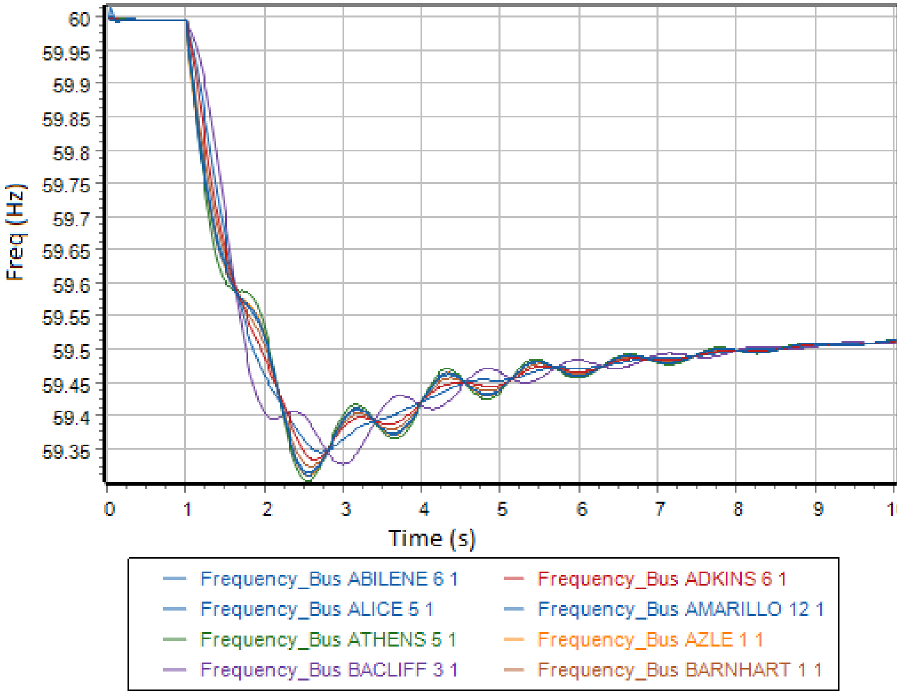
Results

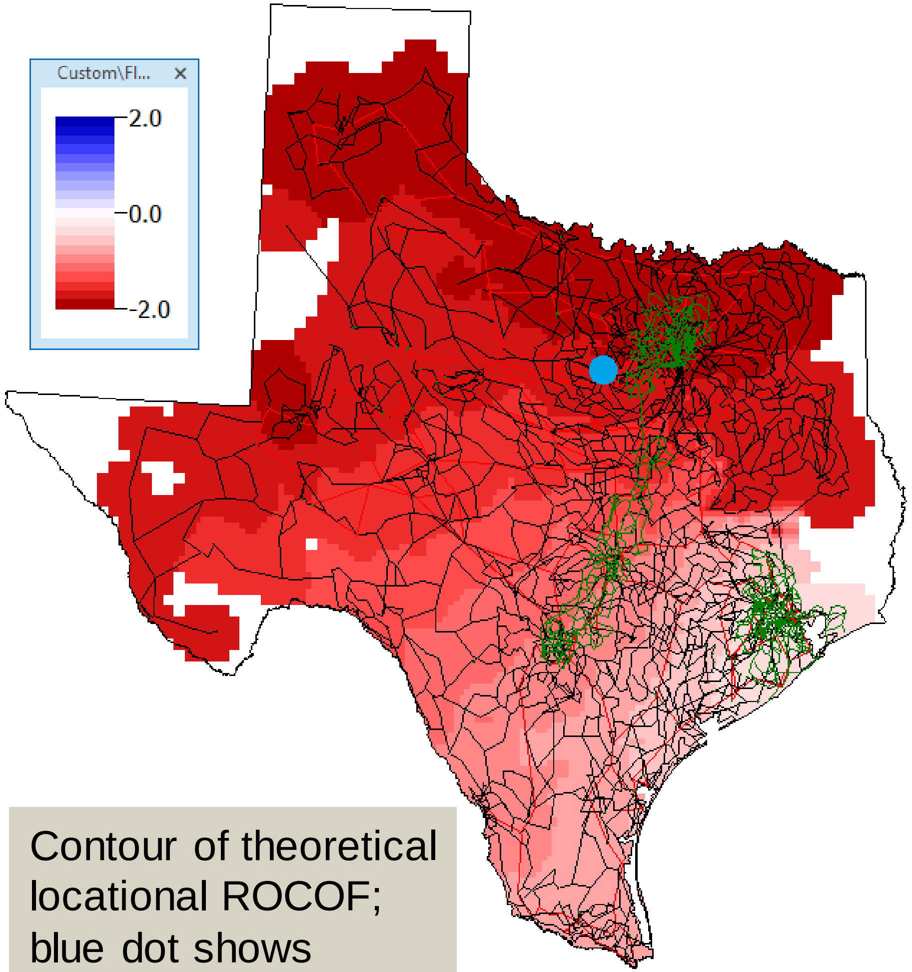
- Runs in about 3 hours, suboptimal implementation on a laptop (much better than 20,000 steady-state contingencies)
- Expected initial ROCOF for every bus in every contingency available
- Plot on the right shows max, min, average for a selection from one base loading case



Results, cont.

- Picking one 1300 MW scenario with interesting results; comparing with full dynamic simulation
- UFLS should not have triggered based on system-level assumptions, but it did here (in a few locations North and East)
- This metric gives a fast way to screen scenarios that may potentially have surprising hidden failure modes





Contour of theoretical locational ROCOF; blue dot shows contingency location

Thanks!

- Our datasets are available for your research!
- To download our datasets, visit our website: electricgrids.engr.tamu.edu
- My email: abirchfield@tamu.edu