Synthetic Data for Power Grid R&D

Tom Overbye and Adam Birchfield
Texas A&M University
overbye@tamu.edu, abirchfield@tamu.edu

July 20, 2017

Background

- Access to data about the actual power grid is often restricted because of requirements for data confidentiality (e.g., critical energy infrastructure)
- For power system community to engage in research that adheres to the scientific principle of reproducibility of results, we need common access to models the mimic the grid complexity
- Work presented here is mostly based an ongoing ARPA-E Grid Data Project involving researchers at TAMU, Illinois, Cornell, ASU and VCU

Overview

- The overall goal of the project is the creation and dissemination of synthetic (fictional) models and scenarios associated with the high voltage grid
  - The models will be of varying size and complexity, ranging from 200 buses up to 100,000 buses; the models will also include contingencies and extra parameters for transient stability and GMD analysis
  - All delivered models with have geographic coordinates
  - Scenarios will be hourly for a year, SCOPF solved
  - Validation metrics are also a key consideration

Approach

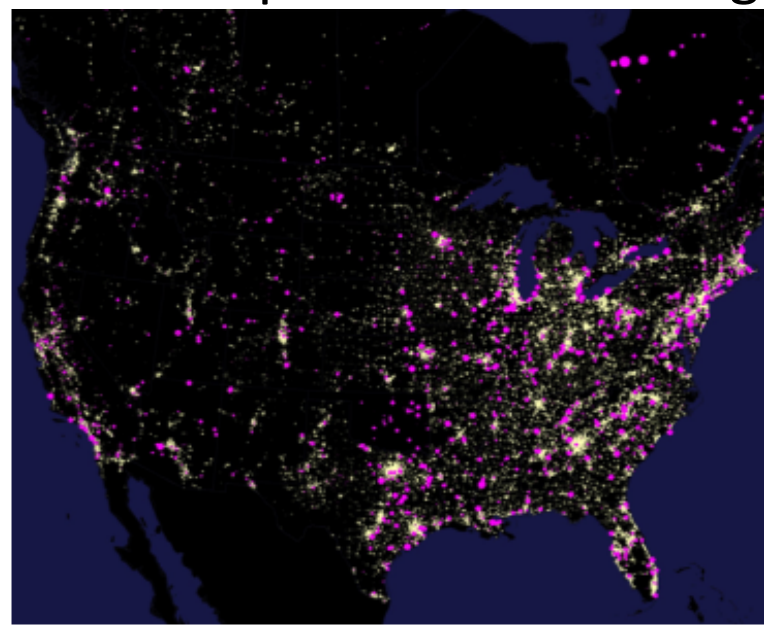
- The approach is to make models that look real and familiar by siting these synthetic models in North America, and serving a population density the mimics that of North America
  - The transmission grid is, however, totally fictitious
- This approach is predicated on gathering statistics associated with actual grids, and then using those statistics as appropriate
  - Actual grids have idiosyncratic characteristics that need to be considered; outlier characteristics can be quite important!

Model Complexity Examples

- A recent 76,000 bus Eastern Interconnect (EI) power flow model has 27,622 transformers including 98 phase shifters
  - Impedance correction tables are used for 351, including about 2/3 of the phase shifters; tables can change the impedance by more than two times
- The voltage magnitude is controlled at about 19,000 buses (by Gens, LTCs, switched shunts)
  - 94% regulate their own terminals with about 1100 doing remote regulation. Of this group 572 are regulated by two or more devices, 277 by three or more, 12 by eight or more, and three by 12 devices!

Geography is Key!

- Actual power grids are geographically consistent
  - This is an inherent characteristic that has profound modeling implications
  - Examples include line impedance, and constraints such as lakes and mountains
- Traditionally power system planning models did not usually include location



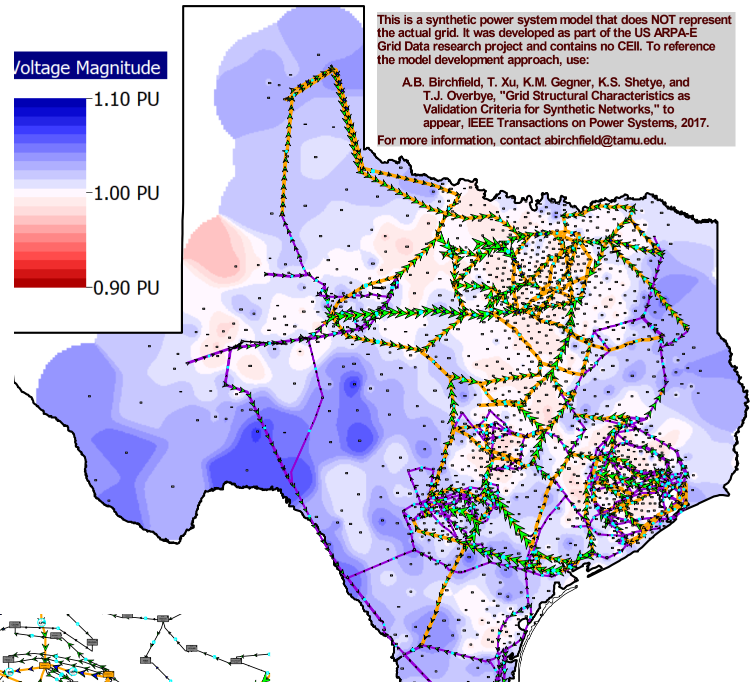
- This is now changing, partially because of GMD studies that require geography

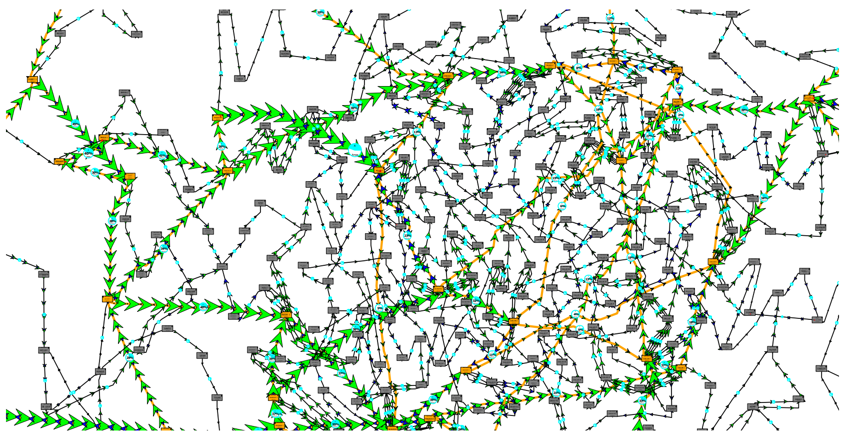
Current Status

- We've created 200, 500 and 2000 bus systems, with a 10,000 bus model almost finished
  - The 200 bus model has OPF, transient stability GMD data, contingencies plus 8760 hourly scenarios
  - The 500 bus has OPF, transient stability and GMD data
  - The 2000 bus has OPF data
  - Additional scenarios and models are being developed
- All models are (or will be) publicly available in a variety of common formats at different locations
  - https://electricgrids.engr.tamu.edu/

Example: 2000 Bus Model

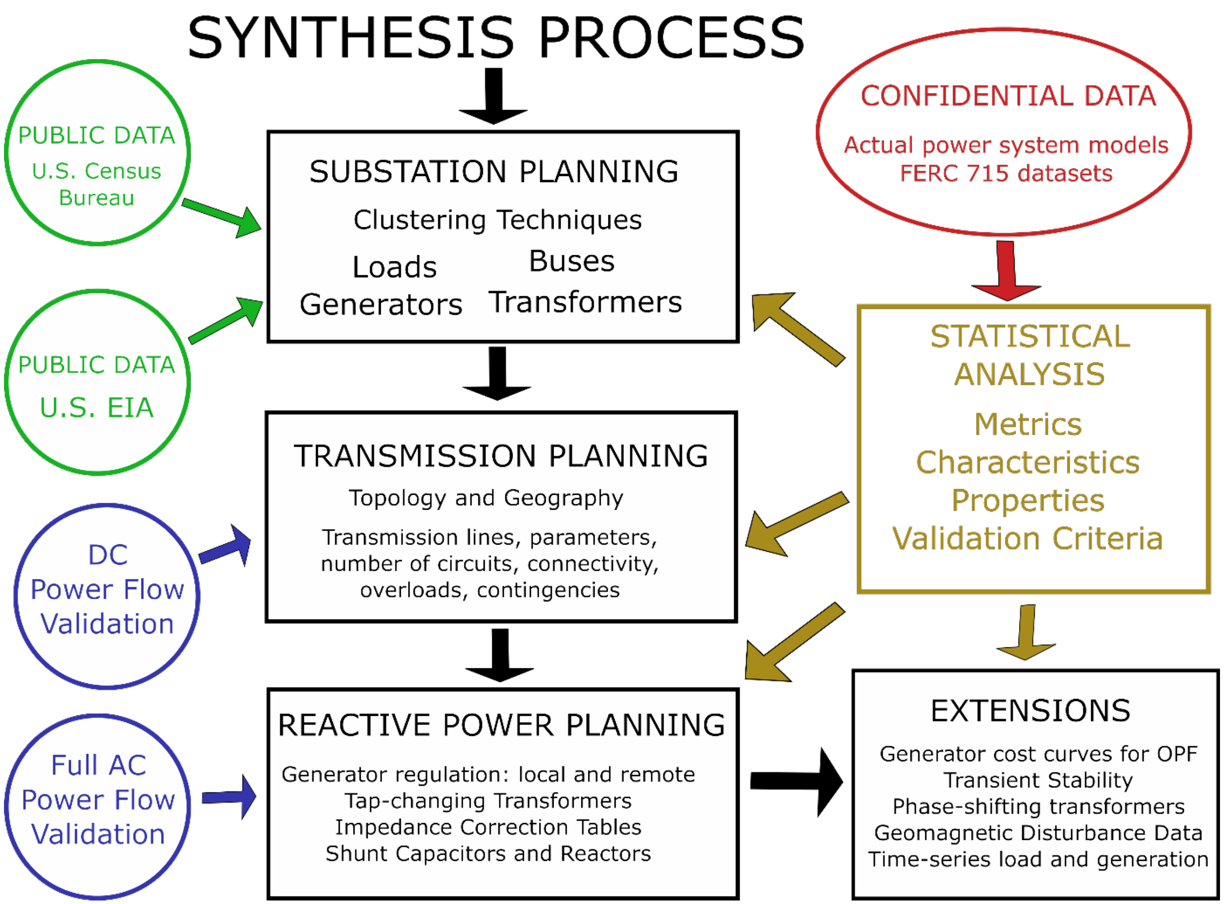
- Geographic footprint is part of Texas (corresponding with ERCOT footprint)
- Four voltage levels: 115 kV, 161 kV, 230 kV, and 500 kV
- Eight areas, appealing onelines





Zoomed View of North Texas

Synthetic Model Design Process



The assumed peak load is based on population, scaled by geographic values

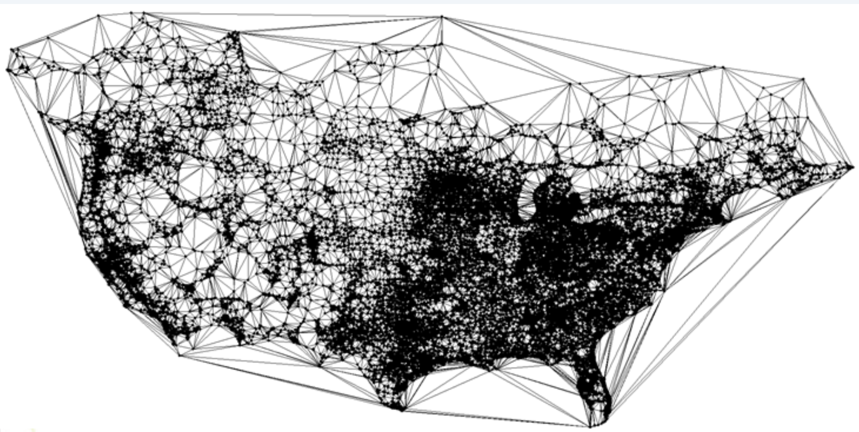
Generation is partially derived from public data (e.g., EIA-860)

Substation oriented approach

Matching Transmission Systems

- We use a number approaches to better mimic the transmission system structure including Delaunay Triangulation, minimum spanning trees, and minimum cycle distribution

Below image shows Delaunay triangulation of 42,000 North America substations; statistics only consider single voltage levels; this is computationally fast (order n ln(n))



The minimum spanning tree (MST) is a subset of the Delaunay triangulation

MST for the EI 500 kV grid; black actual on MST, green other

IEEE PES Power & Energy Society®

IEEE

Validation is Key!

- To-date we've developed about 20 metrics that cover the proportions, size, and structure of actual power grids models, with more coming!
- For example:
  - Buses/substation, Voltage levels, Load at each bus
  - Generator commitment, dispatch
  - Transformer reactance, MVA limit, X/R ratio
  - Percent of lines on minimum spanning tree and various neighbors of the Delaunay triangulation
- Statistics collected from 3 actual interconnects and 12 subset cases from FERC 715 data

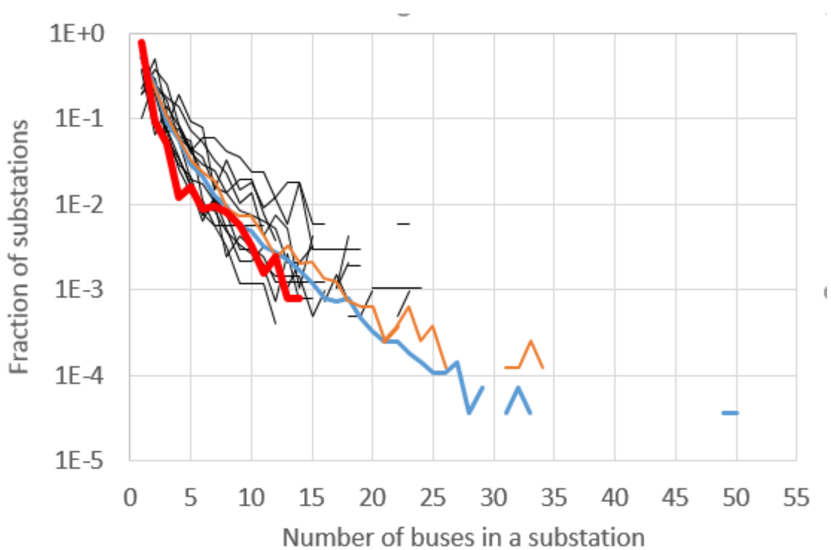
Validation of the 2000 Bus Model

#	Validation Metric	Criteria	ACTIVSg2000
1	Buses per substation	Mean 1.7-3.5	1.6
		Exponential decay	(next slide)
2	Percent of substations containing buses in kV range	< 200 kV, 85-100%	100%
		> 201 kV, 7-25%	16%
3	Substations with load	75-90%	90%
4	Load per bus	Mean 6-18 MW	33.6 MW*
		Exponential decay	(next slide)
5	Generation capacity / load	1.2-1.6	1.46
6	Substations with generators	5-25%	15%
7	Generator Capacities	25-200 MW, 40+%	55%
		200+ MW, 5-20%	30%*
8	Committed Generators	60-80%	79%
9	Generators dispatched > 80%	50+%	67%
10	Generator MaxQ/MaxP	0.40-0.55, > 70%	75%

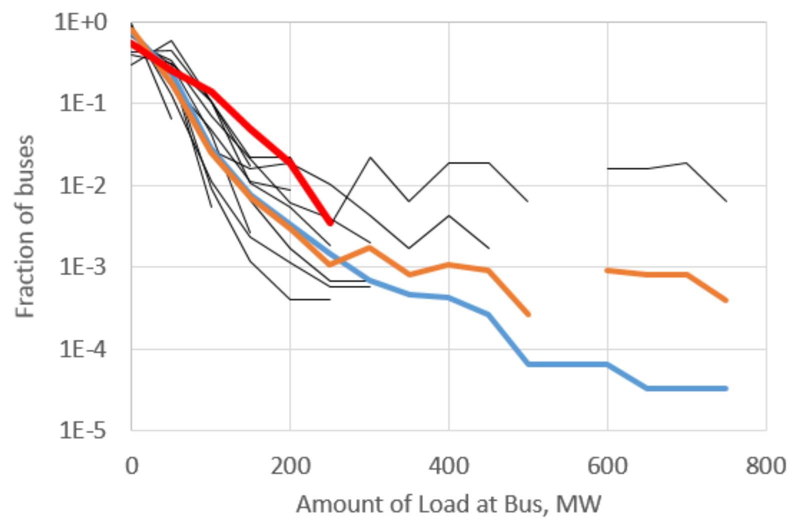
*Note: this model represents a higher load-bus density than actual by design

Validation of the 2000 Bus Model

Number of buses in substation



Amount of load per bus



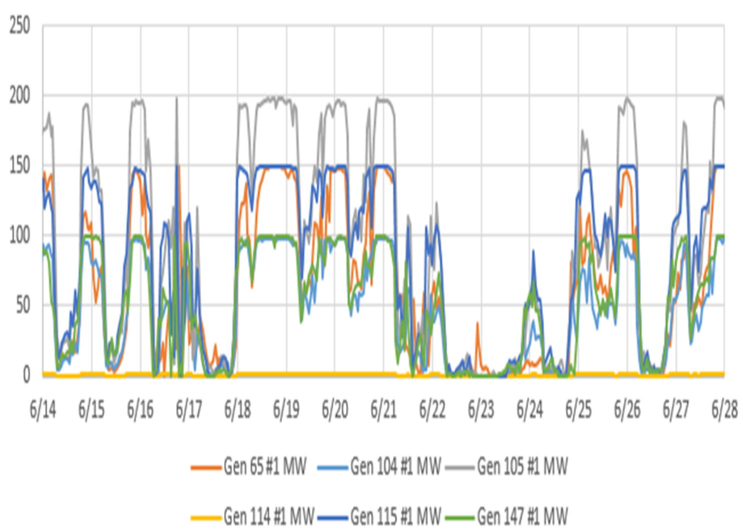
Orange/blue/black actual data, red synthetic

Validation of the 2000 Bus Model

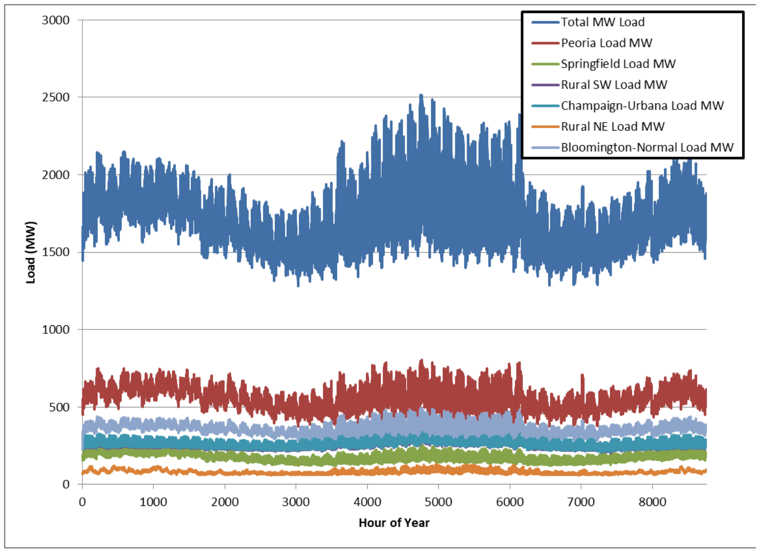
#	Validation Metric	Criteria	ACTIVSg2000			
11	Transformer per-unit reactance, on own base.	80% within [0.05, 0.2]	500 kV	230 kV	161 kV	115 kV
			99%	94%	90%	82%
12	Transformer X/R ratio and MVA limits, by kV level (Table II)	40% below median	51/49	46/41	50/45	48/45
		40% above median	50/51	54/59	50/55	52/55
		80% within 10-90 range	97/98	100/94	100/98	99/82
13	Line per-unit, per-distance reactance, by kV level (Table III)	70% within 10-90 range	96%	98%	100%	100%
14	Line X/R ratio and MVA limits, by kV level (Tables IV and V)	70% within 10-90 range	98/97	100/95	100/100	85/99
15	Lines / Substations, by kV level	1.1-1.4	1.29	1.26	1.23	1.25
16	Lines on min. spanning tree	40-50%	49.0%	49.7%	50.0%	49.5%
17	Distance of line along Delaunay triangulation, by kV level	1, 65-80%	75.2%	75.4%	75.3%	74.8%
		2, 15-25%	20.3%	19.9%	19.8%	20.2%
		3+, 3-10%	4.6%	4.7%	4.9%	5.0%
18	Total line length / MST	1.2-2.2	2.15	1.71	1.62	1.66

Development of Model Scenarios

- Goal of project is to develop models and hourly variation scenarios, SCOPF validated for a year
- This is being done initially for the 200 bus model



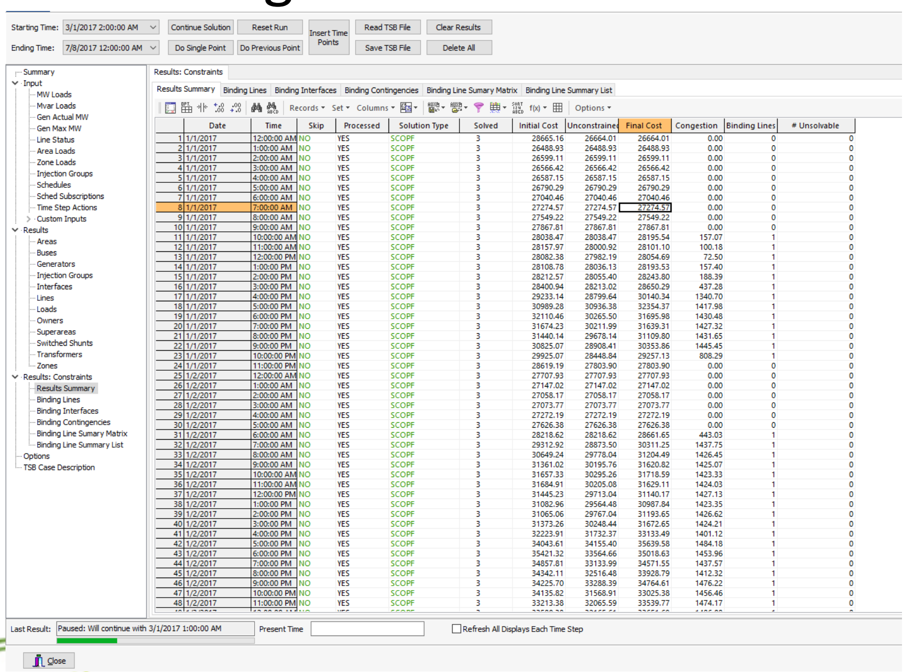
Assumed Hourly Maximum Wind Generation



Assumed Hourly Load

200 Bus SCOPF Validation

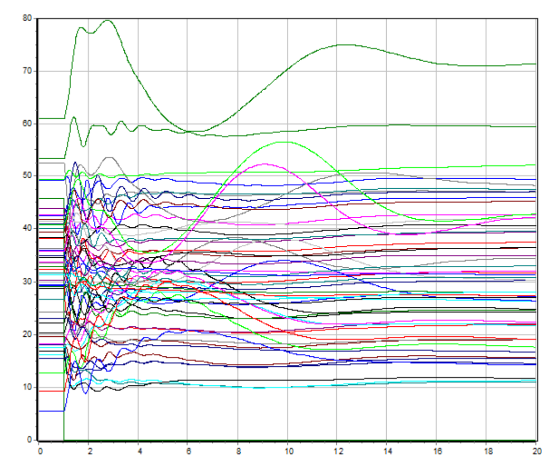
- We are currently combining the 200 bus model, the load schedules, the wind generation variation and the contingencies to do SCOPF validation on model
  - Using PowerWorld Simulator Time Step Simulation

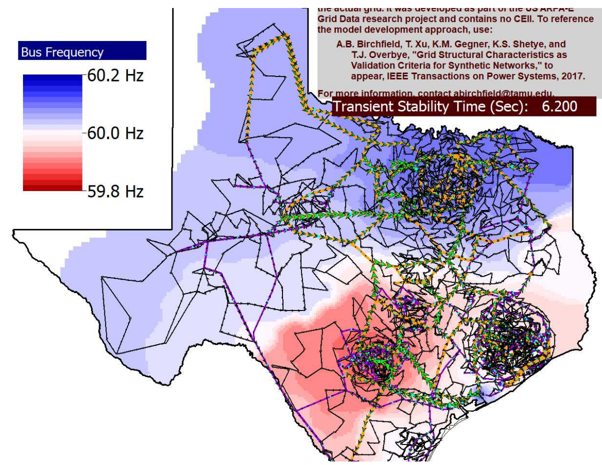


The PowerWorld Simulator tool allows us to easily sequentially solve ac SCOPFs for each of the hourly time points

Transient Stability Models

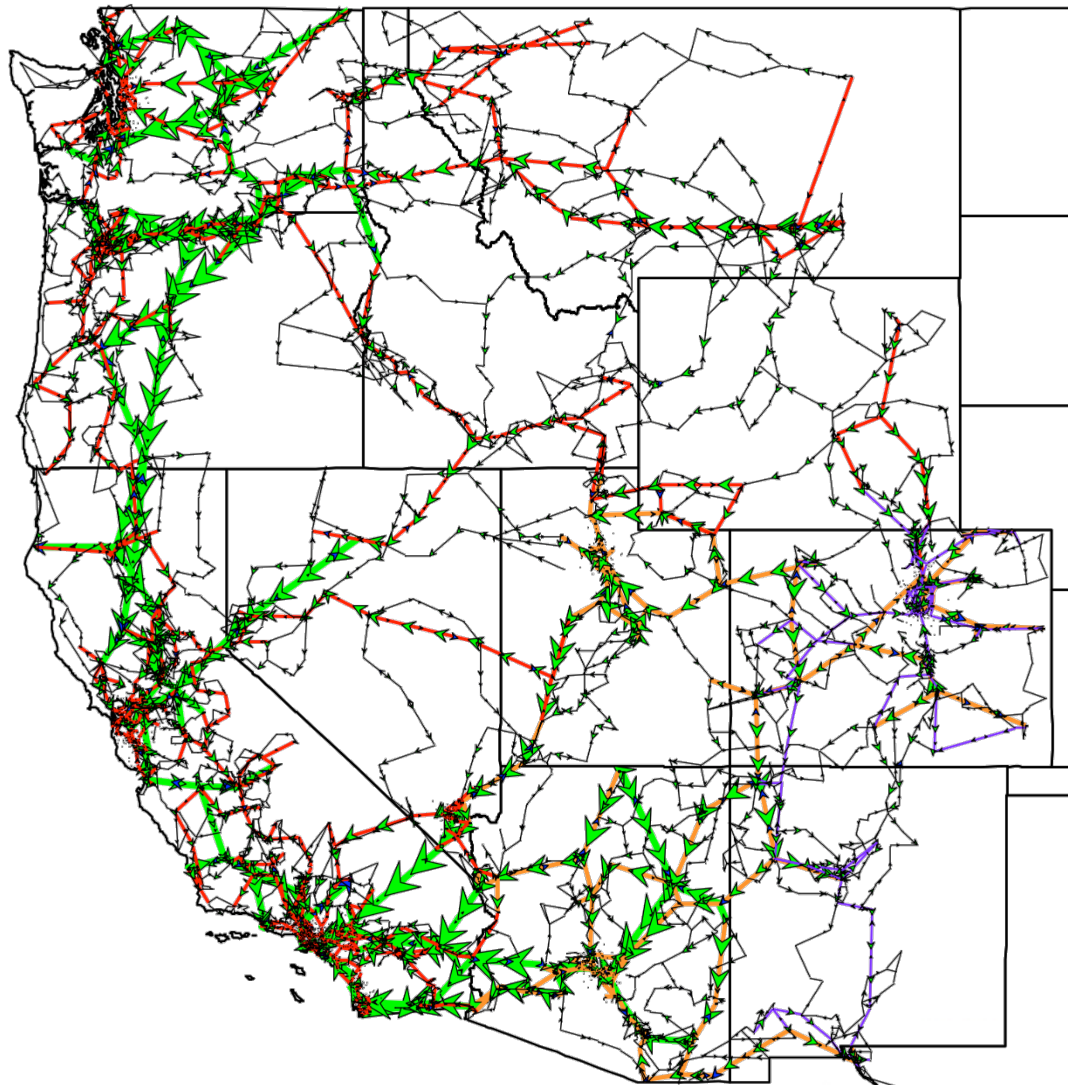
- Ultimately all models will also include transient stability models; this is done for the 200 and 500 bus models, with the 2000 bus almost done
- We're developing models with increasing complexity in model diversity, while using models that run in PowerWorld, PSSE and PSLF
- Validation metrics are also being developed





Next Up: Ten Thousand Bus Case

- Model has 10K buses, 4700 substations, 16 areas, seven nominal transmission voltages (765, 500, 345, 230, 161, 138 and 115kV); total peak load is 150GW
- Green arrows show initial MW flows
- Model should be publicly available soon!



Applications

- To be widely used synthetic models need to be high quality and available in common formats
- Researchers will be able to exchange models and demonstrate techniques on realistic models
- Larger models can be used to enhance teaching
  - TAMU use of 2000 bus model in undergrad courses
  - Vendors adopting synthetic models for training
- Utilities and ISOs may want models adopted to their particular footprints and needs
  - Can be provided to a wider range of potential vendors

Conclusion

- Having access to realistic synthetic models can be an important driver of innovation
- The field of synthetic electric grid models is rapidly developing
  - Creating realistic, large-scale synthetic model is difficult but not impossible
  - Larger scale models are now starting to be released
- There is still much research still to be done especially in the areas of algorithms to create large models and validation metrics!

Thank You!

Questions?

Models are available at
https://electricgrids.engr.tamu.edu