Electric Grid Modeling and Data Challenges Associated with Machine Learning Applications

Tom Overbye
Erle Nye ‘59 Chair for Engineering Excellence
Texas A&M University
overbye@tamu.edu

NSF ML+DS4PES Workshop
March 5, 2020

TEXAS A&M UNIVERSITY

Introduction

- There are many potential machine learning applications in power systems!
- However, there are also many potential pitfalls particularly in the systems area
- In forging collaborations between research communities it is useful to understand the real problems
- We are ultimately trying to model one of the world's most complex machines
  - In North America our large-scale electric interconnects have billions of constantly changing components, and can be very nonlinear (e.g., when a transmission line opens)

A Challenge: Addressing Some Fundamental Issues

- We need more of the academic research community focused on solving large-scale, realistic problems
  - Much is research is done utilizing small-scale and/or simplistic electric grids that fit well with graduate level research
  - Real grids can be quite complex and nonlinear!
  - Not understanding algorithm scaling and the range of operating conditions that need to be considered
- When large, realistic grid models are used they cannot be shared because of legitimate Critical Energy/Electricity Infrastructure Information (CEII) concerns

A Common Request From Industry

- We're gathering all this data. Can you help us figure out what to do with it?
- However, often, after a lot of talks with lawyers to develop NDAs, at most a subset of the desired data and/or models is available
  - The data and models usually cannot be shared, and often a lot of the associated metadata is removed
  - There is a high cost to industry to provide this data
  - **It isn't just the amount of data (e.g., who has the most terabytes!) but rather the breadth of the data and whether the data is coupled to models!!**
- Starting with just a subset of information can result in less useful results

Changing the Question

- Assume you have all the information possible about the electric grid (data, models, coupled infrastructures) and can share it.
- And it can be provided with low access cost.
- **What would you do with it?**
  - This allows for large amounts of data fusion!
- This is crucial for machine learning applications!
- The focus of this talk is mostly on how to develop this seemingly impossible information set
  - And, of course, one can never know everything!

A Related Issue: Getting Operational Experience

- Many students graduating from electric power programs have little or no experience operating an electric grid, real or simulated
  - This often includes students whose research has focused on improving electric grid operations
- We're developing a realistic simulation environment in which individuals and/or teams of students can gain this experience
- An added benefit is getting lots of operational data for use in machine learning applications

The Simulation Environment

A Recent Success: Synthetic Electric Grids and Datasets

- Synthetic electric grids are fictional representations that are free from confidential information and hence can be freely shared
- Two recent National Academies Reports (2016 and 2017) called for greater research in this area
- Over the last four years tremendous progress has been made through various ARPA-E and NSF projects in the creation of large-scale, high quality, realistic synthetic grids at both the transmission and distribution levels
- Goal is that innovation done with these grids can be directly applied to the actual grid

Synthetic Electric Grids

- Synthetic electric grids are models of electric grids that were not created to represent any actual electric grid
- The below image shows the five bus synthetic grid I used as an undergraduate

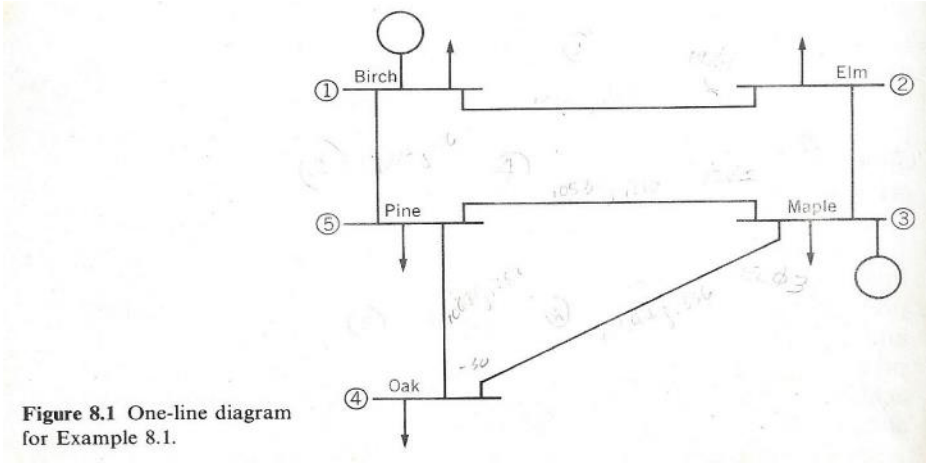


Figure 8.1 One-line diagram for Example 8.1.

Image Source: W.D. Stevenson, *Elements of Power Systems*, Fourth Edition, McGraw-Hill Book Company New York, 1982 (the first edition was in 1955)

Geographically-Based Synthetic Electric Grids

- Synthetic electric grids can be created with or without reference to actual geography
- The image shows an early geographically-based synthetic electric grid
- This grid was designed to show concepts to regulators

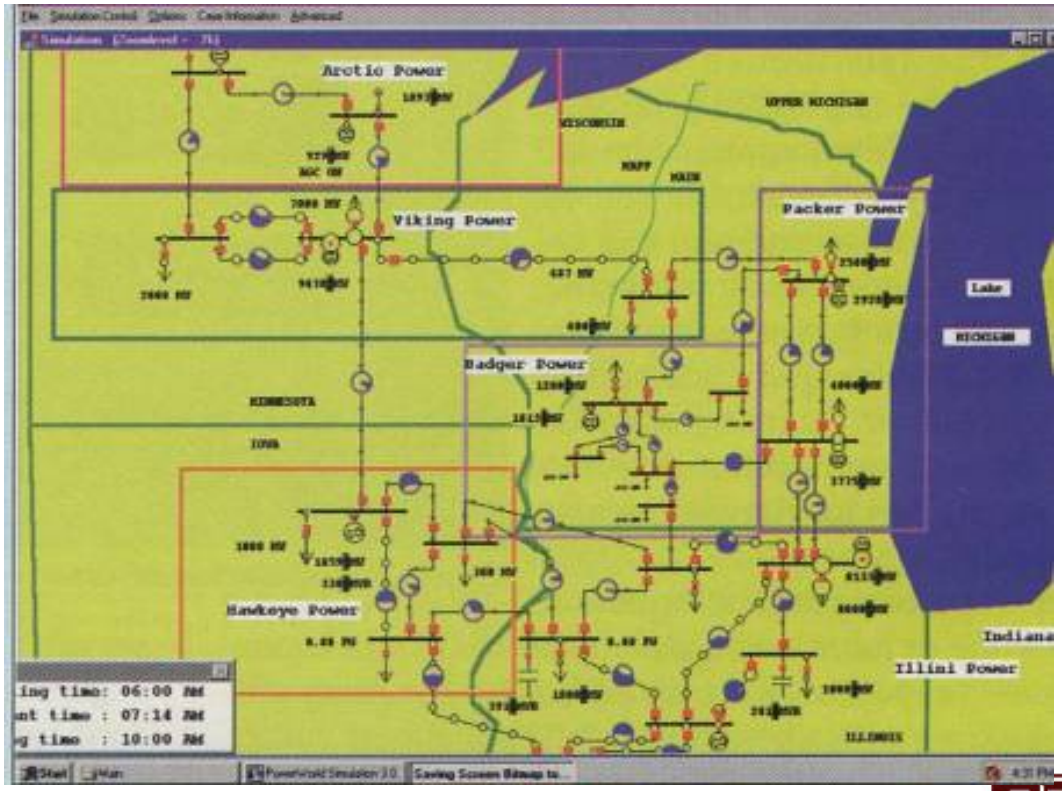


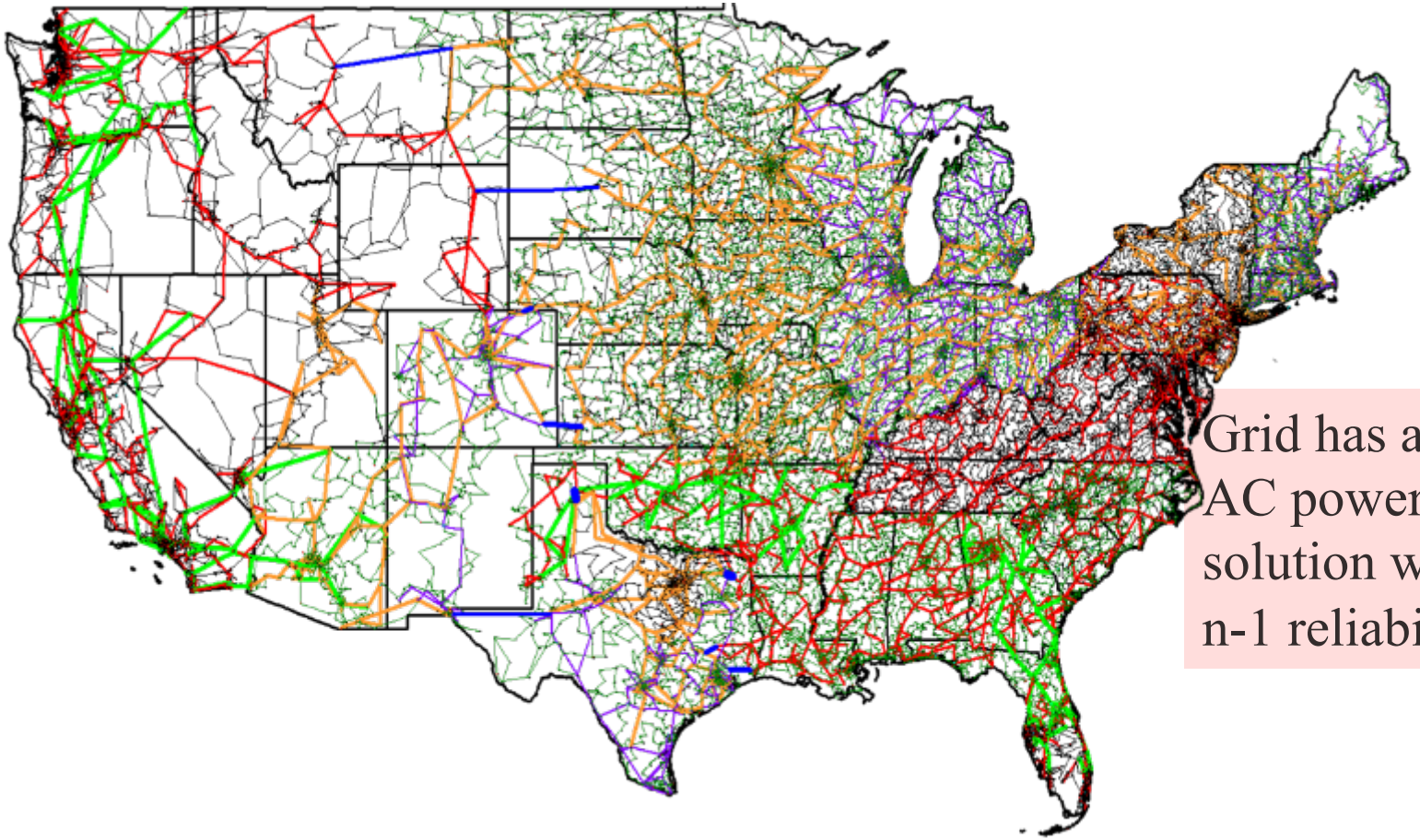
Image Source: PowerWorld Corporation, 1995

High-Quality, Geographically-Based Synthetic Electric Grids

- High-quality synthetic electric grids are designed to have a wide range of characteristics that are similar to those found in actual electric grids
  - "Realistic but not real" to quote Wisconsin colleagues
  - Fictional, but hopefully good fiction
  - **Developed techniques can be applied to real grids**
- However, importantly these grids are not designed to try to duplicate any actual grid
- Over the last three years tremendous progress has been made through ARPA-E at both the transmission and distribution levels

Current Status: Large-Scale Grids are Now Available

This is an 82,000 bus synthetic model that we publicly released in summer 2018 at **electricgrids.engr.tamu.edu**

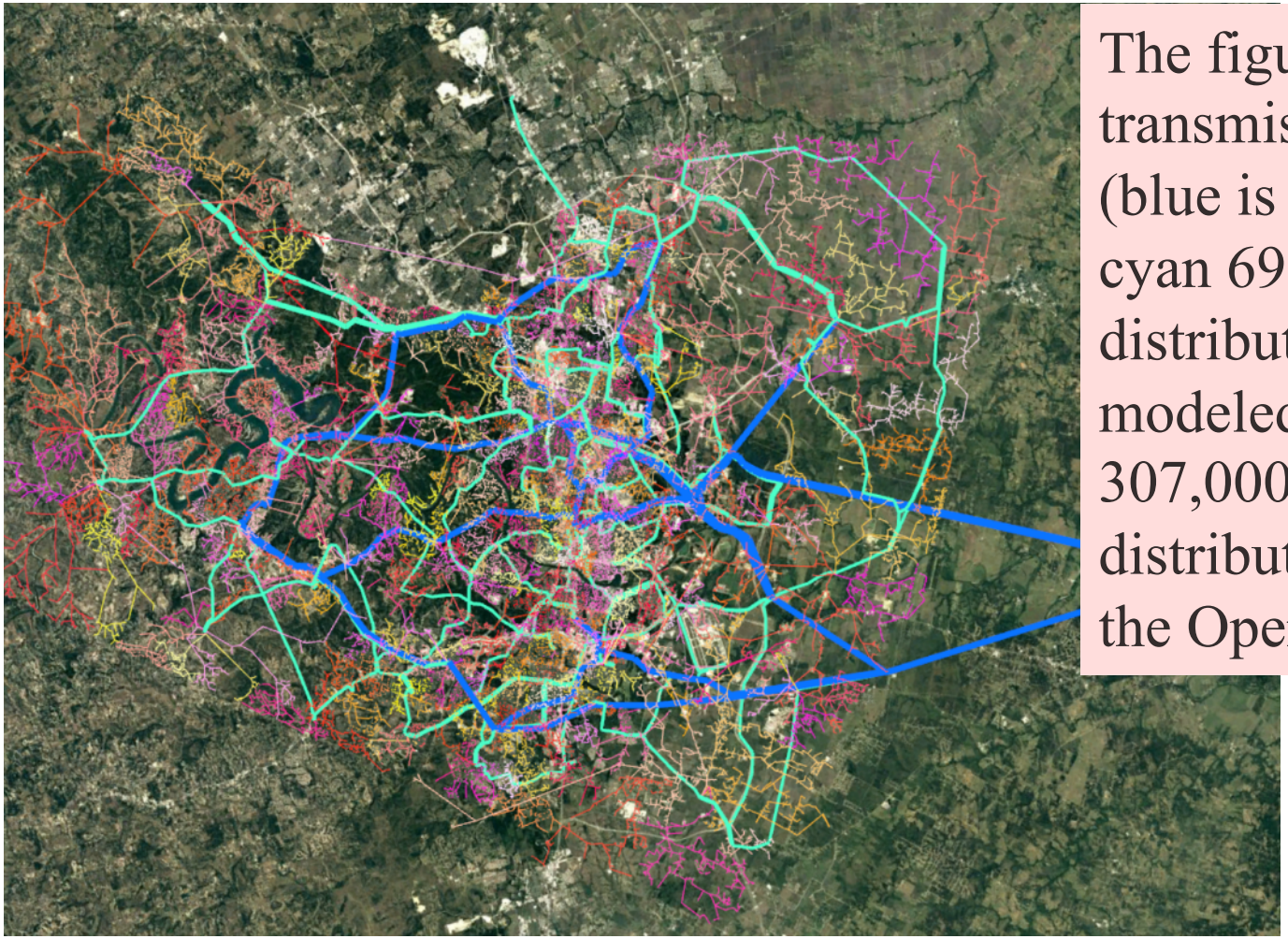


Grid has an AC power flow solution with n-1 reliability

Highly Detailed Combined Transmission and Distribution Grids

- Previous transmission grids were geographic to the zip code level
- On current ARPA-E project we (with NREL, MIT and Comillas-IIT University) are developing “down to the meter” synthetic grids
- Actual parcel data is used to determine location of the electric meters. The parcels are connected by a distribution system, and the distribution system by a transmission grid
- Currently we have about much of the load in the southern half of Texas done

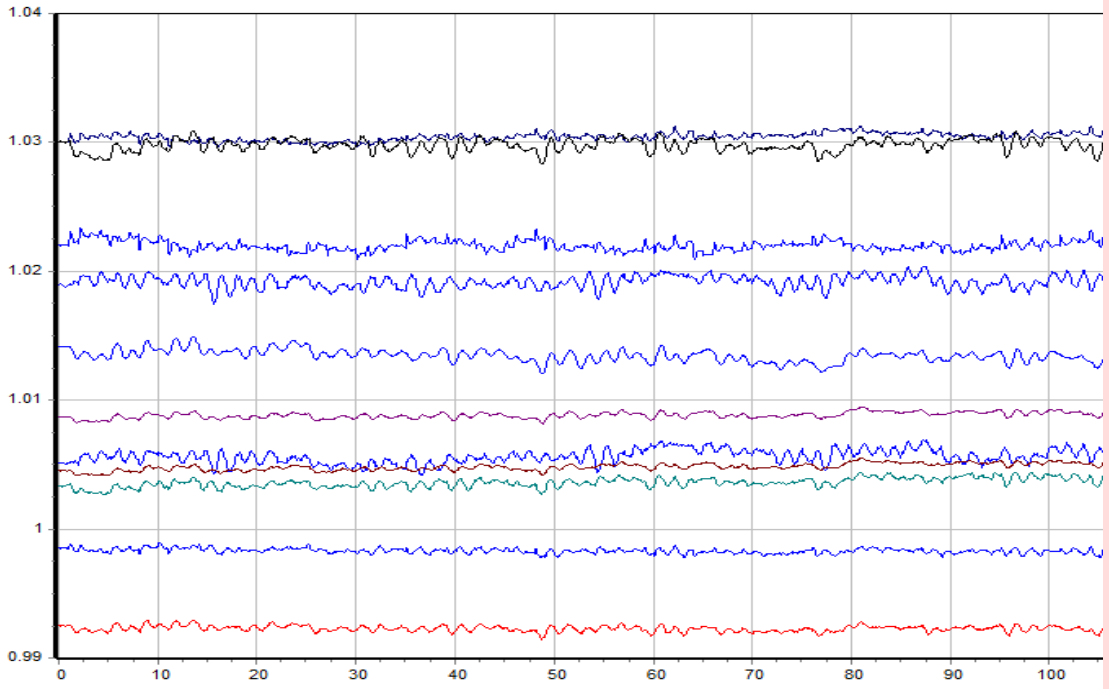
Travis County, Texas (location of Austin, TX)



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.

Synthetic PMU Data

- To create synthetic PMU data we use dynamic simulations with realistic variation in the load and generation, and appropriate PMU errors.



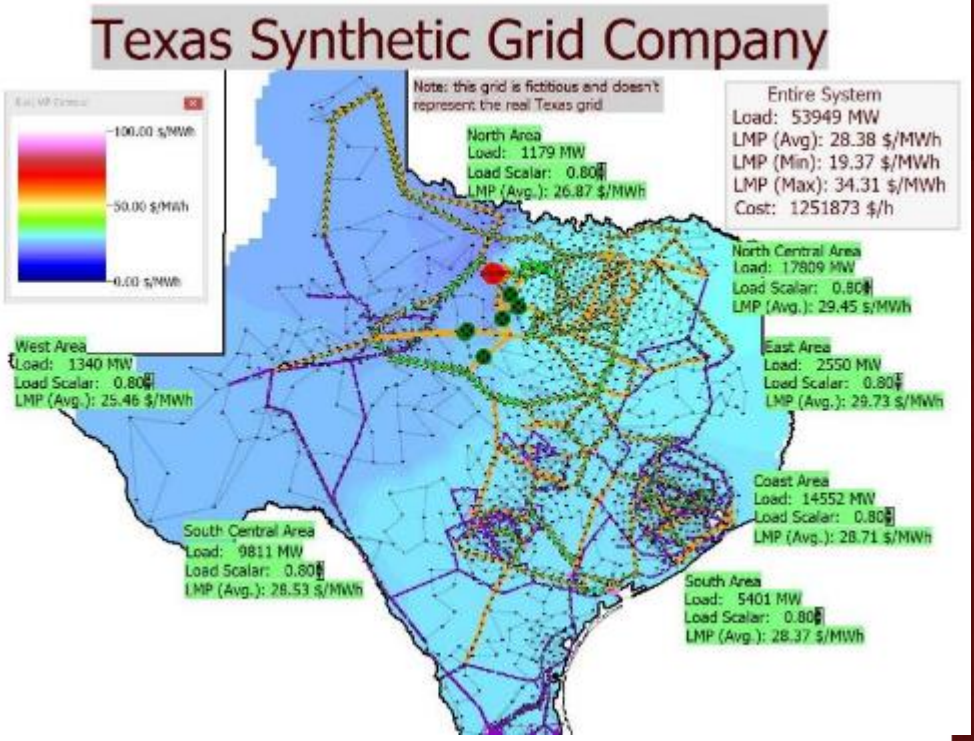
On our 2000 bus Texas synthetic grid was can provide more then 10,000 signals at PMU data rates (30 times per second) in real-time. This is data creation at about 4GB per hour.

Different Levels of Modeling

- Just because we have detailed grids, doesn't mean we always simulate the coupled transmission and distribution models. Other options are
  - Transmission only
  - Distribution only
  - Full transmission with distribution topology; this can be quite useful for doing multi-infrastructure simulation in which we just need to know what parts of the distribution system are out-of-service

Synthetic Grid Applications: Innovative Electric Power Education

- Lab assignments involving a 2000 bus case have been integrated into Texas A&M's power classes
- Class includes large-system exercises for power flow, economic dispatch, contingency analysis, SCOPF, and transient stability



Applications: Visualization Research

- With the now wide spread availability of large-scale, geographically based public synthetic grids, there has been increased need for better wide-area visualizations
- Visualizations involve many trade-offs, with the best approach ultimately application dependent
- Recent research using the synthetic grids is focused on how to maximize display space, while retaining some geographic context

Pseudo-Geographic Mosaic Displays Can Maximize Screen Usage

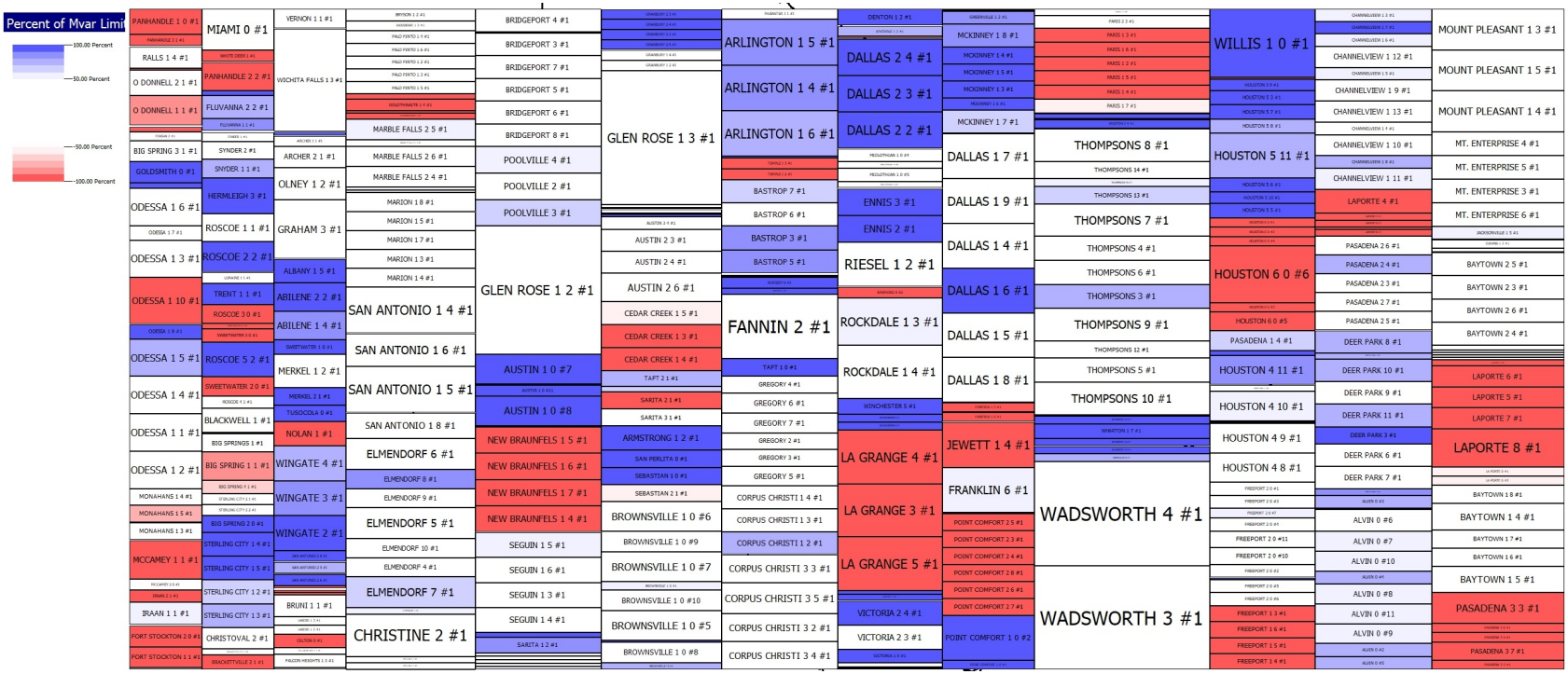
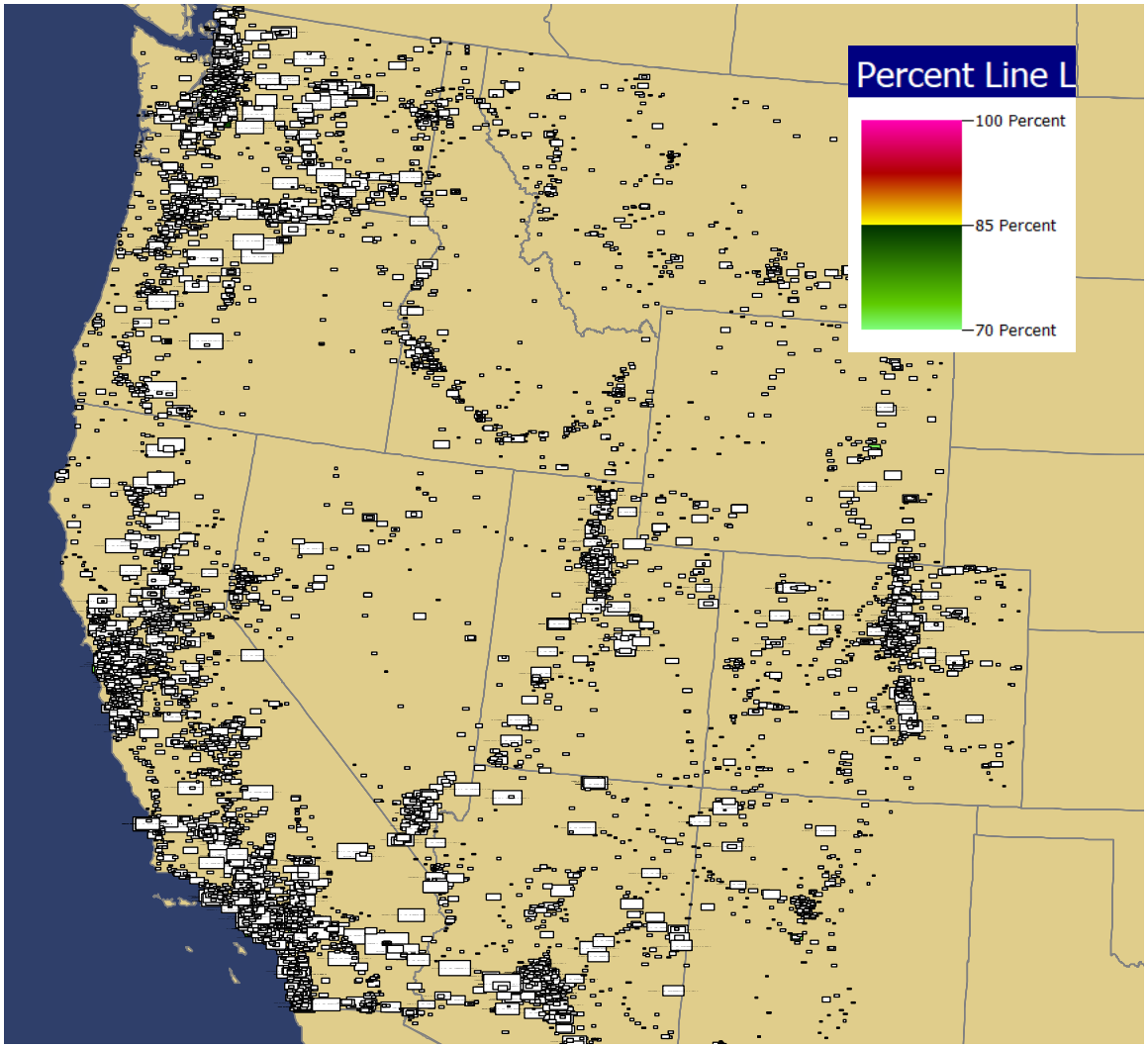


Image visualizes percentage generator reactive power output. Image source: Overbye, Wert, Birchfield, Weber, “Wide-Area Electric Grid Visualization Using Pseudo-Geographic Mosaic Displays,” Proc. 2019 North American Power Symposium (NAPS), Wichita, KS, Oct 2019

Visualizing 12,700 Line Flows:



Large-scale synthetic grids can help guide new visualization solutions. This image attempts to show 12,700 line flows, with object size based on the line's MVA flow and its coloring based on percentage loading.

Visualizing 12,700 Line Flows:

Visualization using a pseudo-geographic mosaic approach with all rectangles linked to provide drill-down details

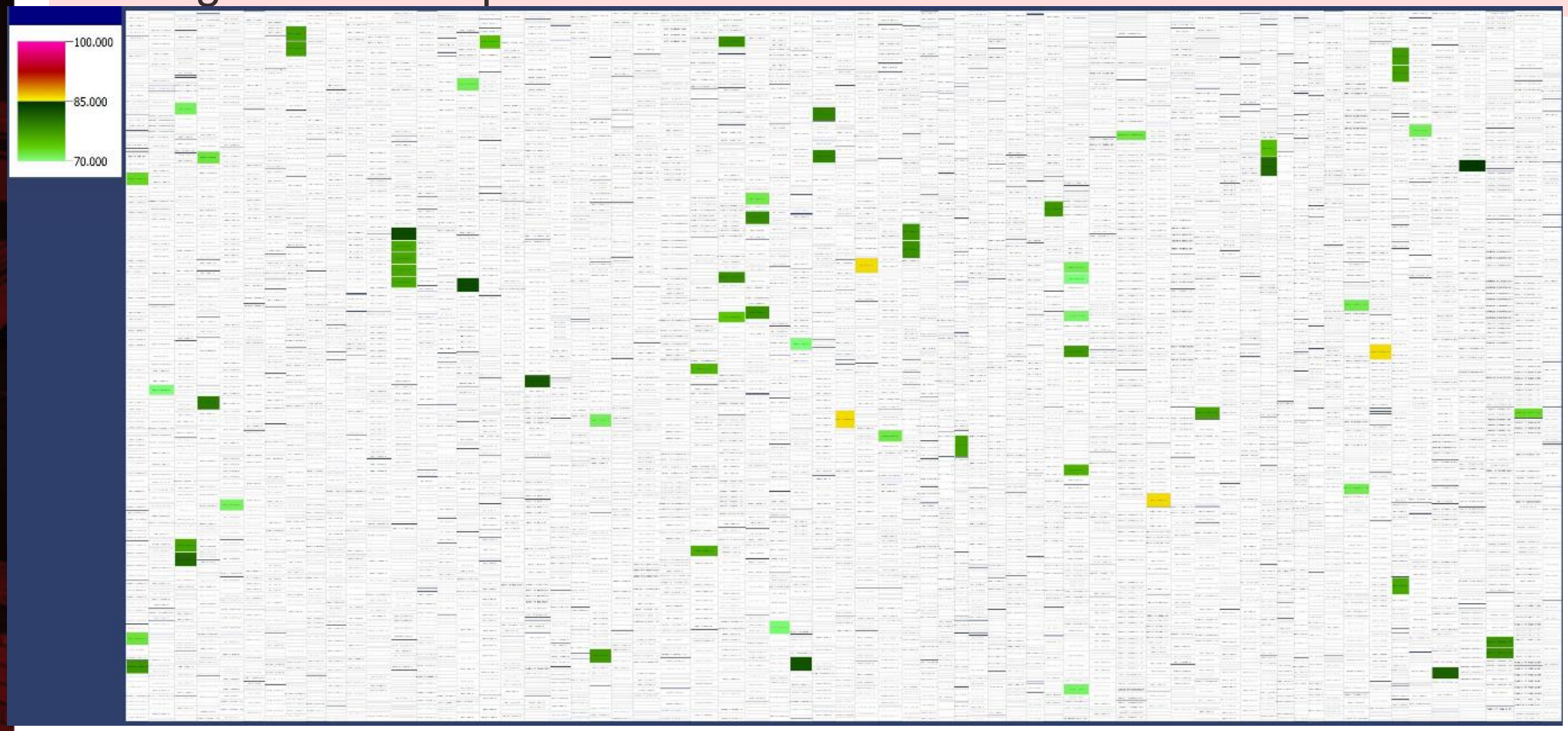
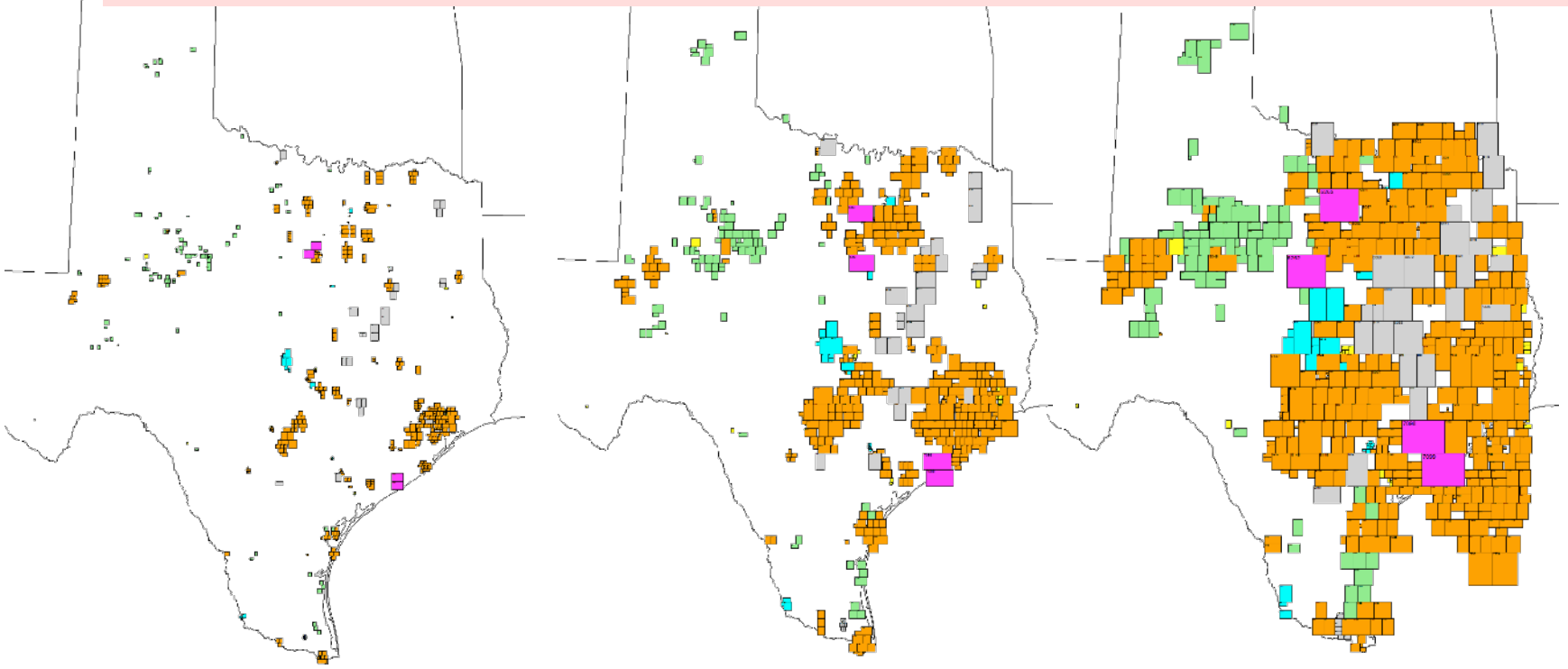


Image source: Overbye, Wert, Birchfield, Weber, “Wide-Area Electric Grid Visualization Using Pseudo-Geographic Mosaic Displays,” Proc. 2019 North American Power Symposium (NAPS), Wichita, KS, Oct 2019

Auto Screen Layout Designed to Percentage Screen Fill

Images show generation by capacity and fuel type using different percentages of the available screen space



2%

10%

30%

Image source: Adam Birchfield

Summary

- Thanks to investment primarily by ARPA-E and NSF, large-scale, highly detailed synthetic grids are a reality
  - Ongoing research on enhancements and application
- These can be used to generate large amounts of context-aware data, helping to enable a large amount of power system research particularly machine learning
- Coupled infrastructure research is now potentially available to a much broader audience

Thank You!

Questions?