New Developments in the Visualization of Wide-Area Electric Grid Information

Thomas J. Overbye
Erle Nye ’59 Chair for Engineering Excellence
Texas A&M University

Texas A&M Smart Grid Center Webinar
June 24, 2020

Acknowledgments

- Work presented here has been supported by a variety of sources including the Texas A&M Smart Grid Center, PSERC, DOE, ARPA-E, NSF, EPRI, BPA, and PowerWorld. Their support is gratefully acknowledged!
- Slides also include contributions from many of my students, postdocs, staff and colleagues at both TAMU and UIUC
- Thanks for human factor aspects from Prof. Esa Rantanen, Rochester Institute of Technology

Our Energy Future Could be Quite Bright!

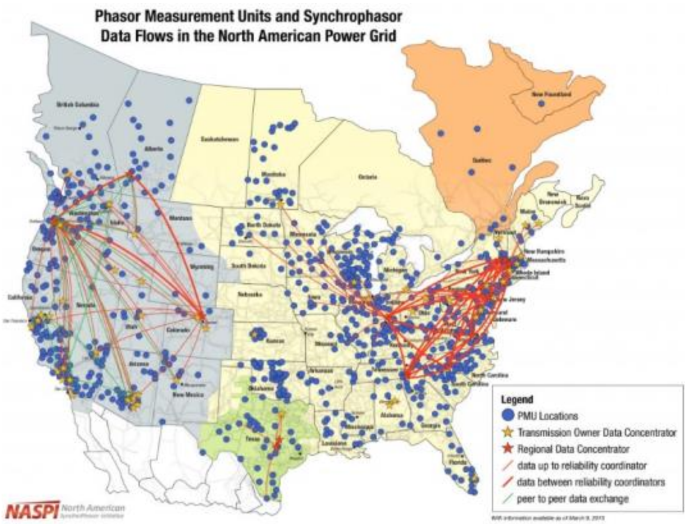
- My professional goal is to help in the development of a sustainable and resilient electric infrastructure for the entire world.
- Electric grids are in a time of rapid transition, with lots of positive developments.
- I think our electric energy future could be quite bright! But there are lots of challenges with this transition, including maintaining human situational awareness, particularly during times of stress.

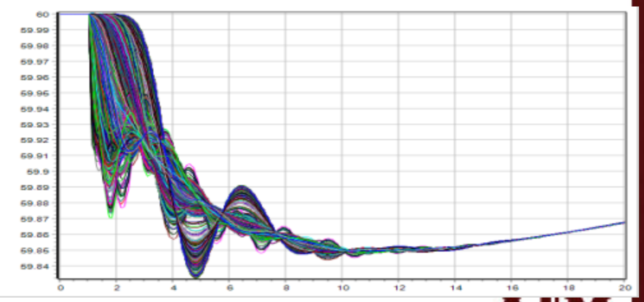
Overview

- Presentation focuses on how wide-area electric grid visualization can help
- Grids are getting increasingly complex, particularly with many more automatic controls, and there is a concern about whether anybody fully really knows what’s going on
- How data is transformed into actionable information is a crucial, yet often unemphasized, part of the software design process
- Focus here is more on visualization for engineers, as opposed to operators

Examples of Power System "Big Data"

- Power system operations and planning are a rich source of data
  - SCADA has traditionally provided a grid data at scan rates of several seconds
  - Thousands of PMUs are now deployed providing data at 30 times per second
  - In planning many thousand of studies are now routinely run, with a single transient stability run creating gigabytes





A Specific Example: One Dynamics Study

- One of our projects is looking at dynamic aspects of an ac interconnection of the Eastern Interconnect and the WECC
- We're doing lots of dynamic simulations (some 30 seconds, soon to be much longer)
- The Model has 110,000 buses, 244 different types of dynamic models, 48,000 model instances and 194,000 states
- A human factors challenge in doing a study is for the engineer to know what happened

Visualization Software Design

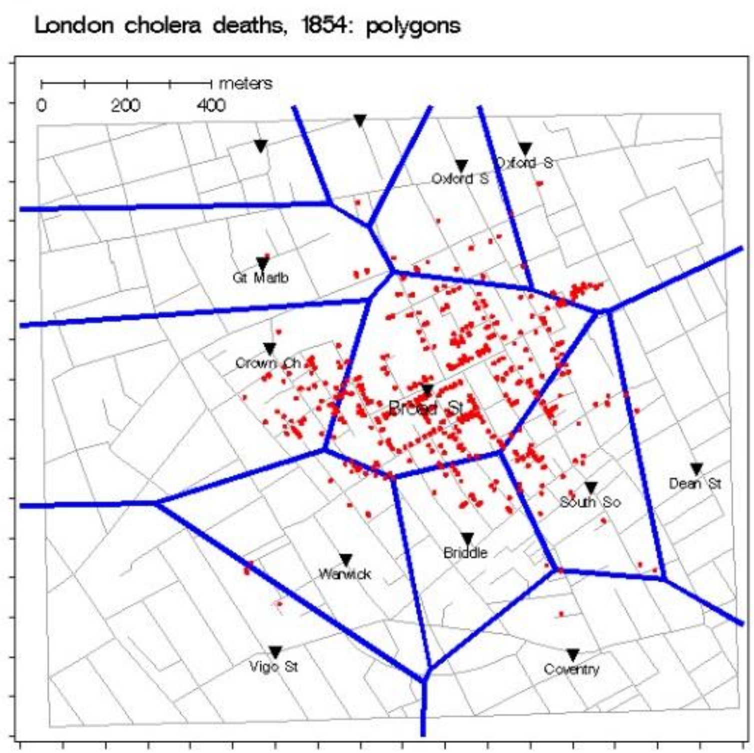
- Key question: what are the desired tasks that need to be accomplished?
  - Needs for real-time operations might be quite different than what is needed in planning
- Understanding the entire processes in which the visualizations are embedded is key
- Software should help humans make the more complex decisions, i.e., those requiring information and knowledge
  - Enhance human capabilities
  - Alleviate their limitations (adding up flows into a bus)

Some Useful General References

- Colin Ware, *Information Visualization: Perception for Design*, Fourth Edition, 2021
- Edward Tufte, *Envisioning Information*, 1990
- Edward Tufte, *Visual Explanations: Images and Quantities*, 1997
- Edward Tufte, *The Visual Display of Quantitative Information*, 2001
- Edward Tufte, *Beautiful Evidence*, 2006
- Claus Wilke, *Fundamentals of Data Visualization*, 2019

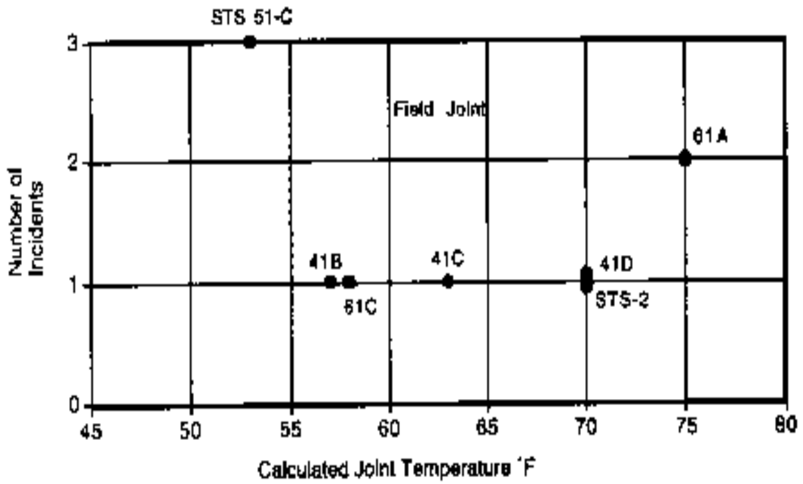
Example: Visualization and Cholera in Central London, 1854

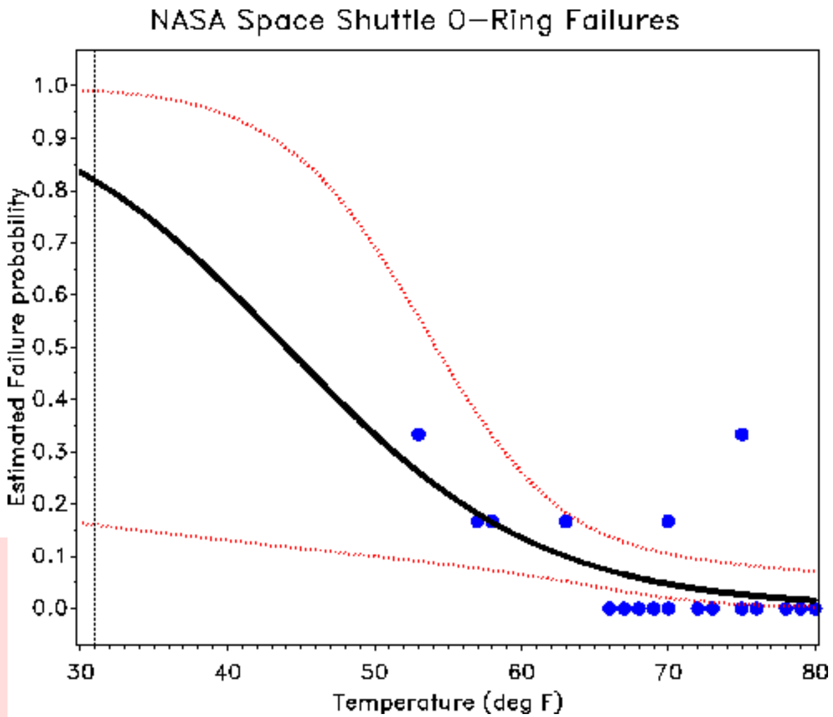
Dr. John Snow helped to end an epidemic by noting that the deaths were clustered about the Broad Street Water pump



Images source: http://www.datavis.ca/gallery/historical.php

Example: Visualization and Challenger, 1986





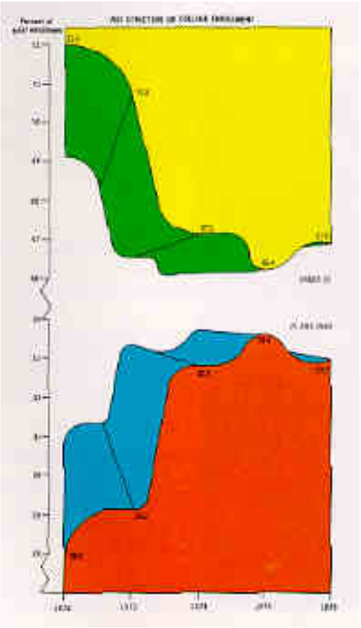
Graph used to determine whether to launch Challenger in 1986; shows O-ring failure vs. temperature but neglects launches with no failures

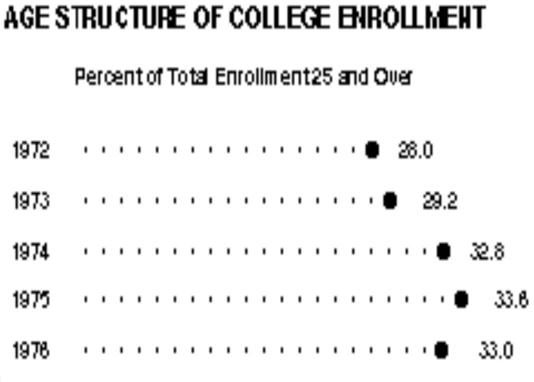
A better graphic including all data. Would you launch at 32ºF?

http://www.datavis.ca/gallery/missed.php

Visualization Cautions!

- Just because information can be shown graphically, doesn't mean it should be shown
- Three useful design criteria from 1994 EPRI visualization report:
  1. Natural encoding of information
  2. Task specific graphics
  3. No gratuitous graphics





Tufte: "may well be worst graphic ever"

Source: E. Tufte, The Visual Display of Quantitative Information, Graphics Press, Cheshire, CT, 1983.

Visualization and User Familiarity

- Visualizations do not exist in a vacuum; the prior experience of the users is a key consideration
  - QWERTY keyboard arrangement is a classic example, in which a design that was originally setup in the 1870's to prevent mechanical problems is still used today
- Using existing visual metaphors in new designs help them seem more familiar (like a folder)
  - A skeuomorphic design retains no longer needed structures that were inherent in the original, usually to make them more familiar (using gauges, sliders, buttons and analog clocks in visualizations are examples)

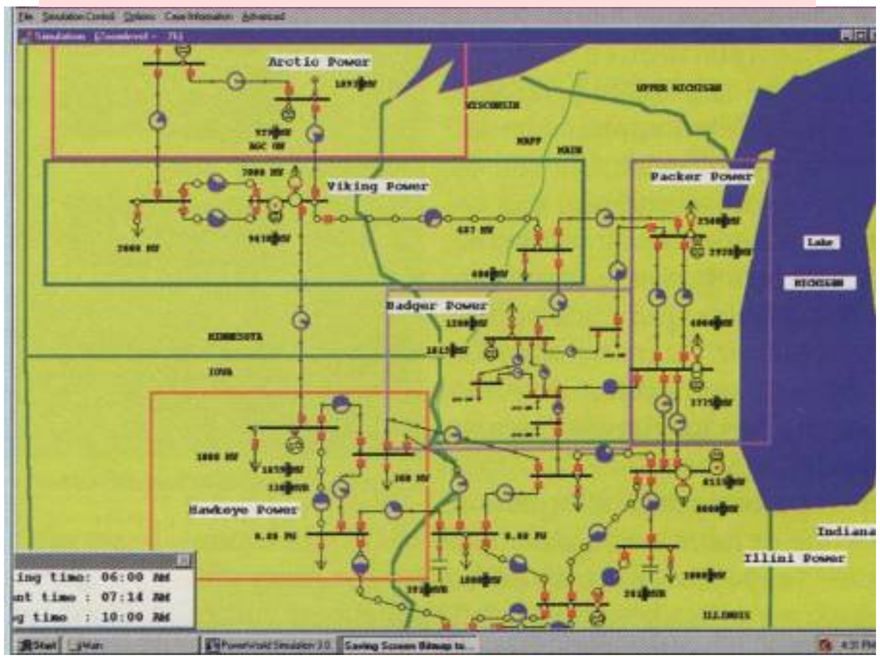
Synthetic Models and Visualization

- Access to actual power grid models is often restricted (CEII), and this can be a particular concern with data analysis and visualization since its purpose is to provide insight into the model, including weaknesses
  - Models cannot be freely shared with other researchers, and even presenting results can be difficult
- A solution is to create entirely synthetic (fictitious) models the mimic characteristics of actual models
  - Kudos to the US DOE ARPA-E for funding work over the last five years in this area

Early Synthetic Grids

- Synthetic electric grids are models of electric grids that do not represent any actual electric grid

A pseudo-geographic grid



A non-geographic grid

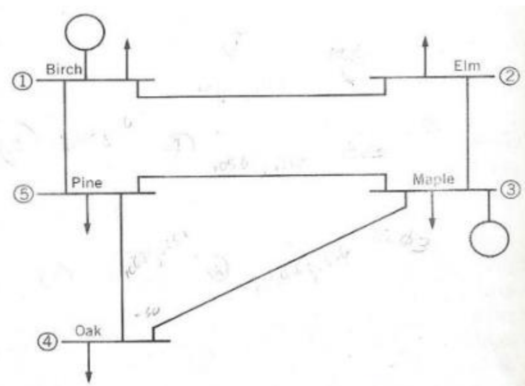


Figure 8.1 One-line diagram for Example 8.1.

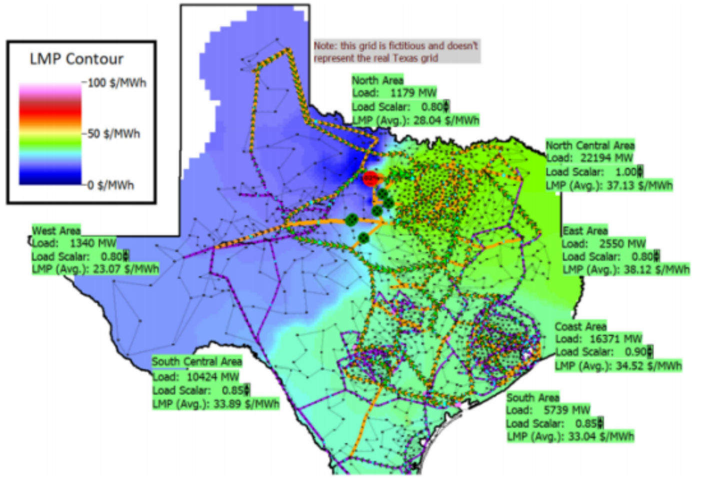
Left Image Source: W.D. Stevenson, *Elements of Power Systems*, Fourth Edition, McGraw-Hill Book Company New York, 1982 (the first edition was in 1955)

Synthetic Models Used Here

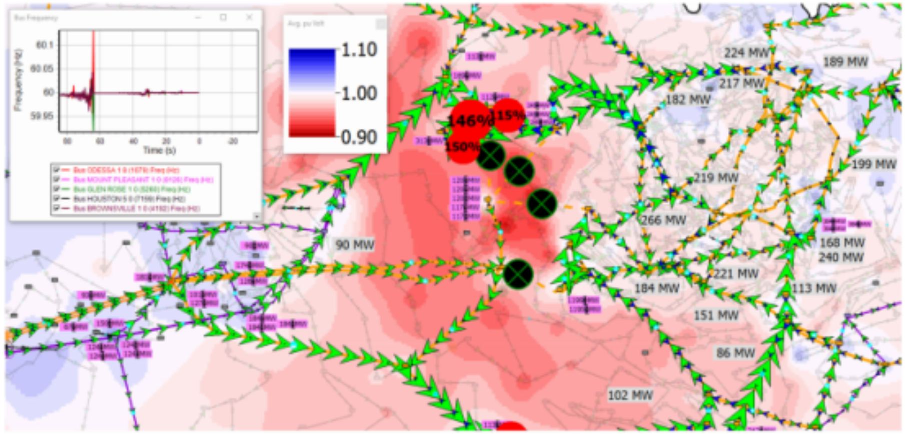
- Examples presented here will be based on either a 42-bus grid, a 2000-bus synthetic grid covering the ERCOT footprint, or an 80,000-bus synthetic grid modeling separate US east and west grids combined by a few ac interties
- All grids have embedded geographic coordinates
- The inclusion of geographic coordinates in actual electric grid models has increased rapidly over the last few years, driven in part by their requirement for geomagnetic disturbance (GMD) impact studies

2000 Bus Texas Synthetic Grid

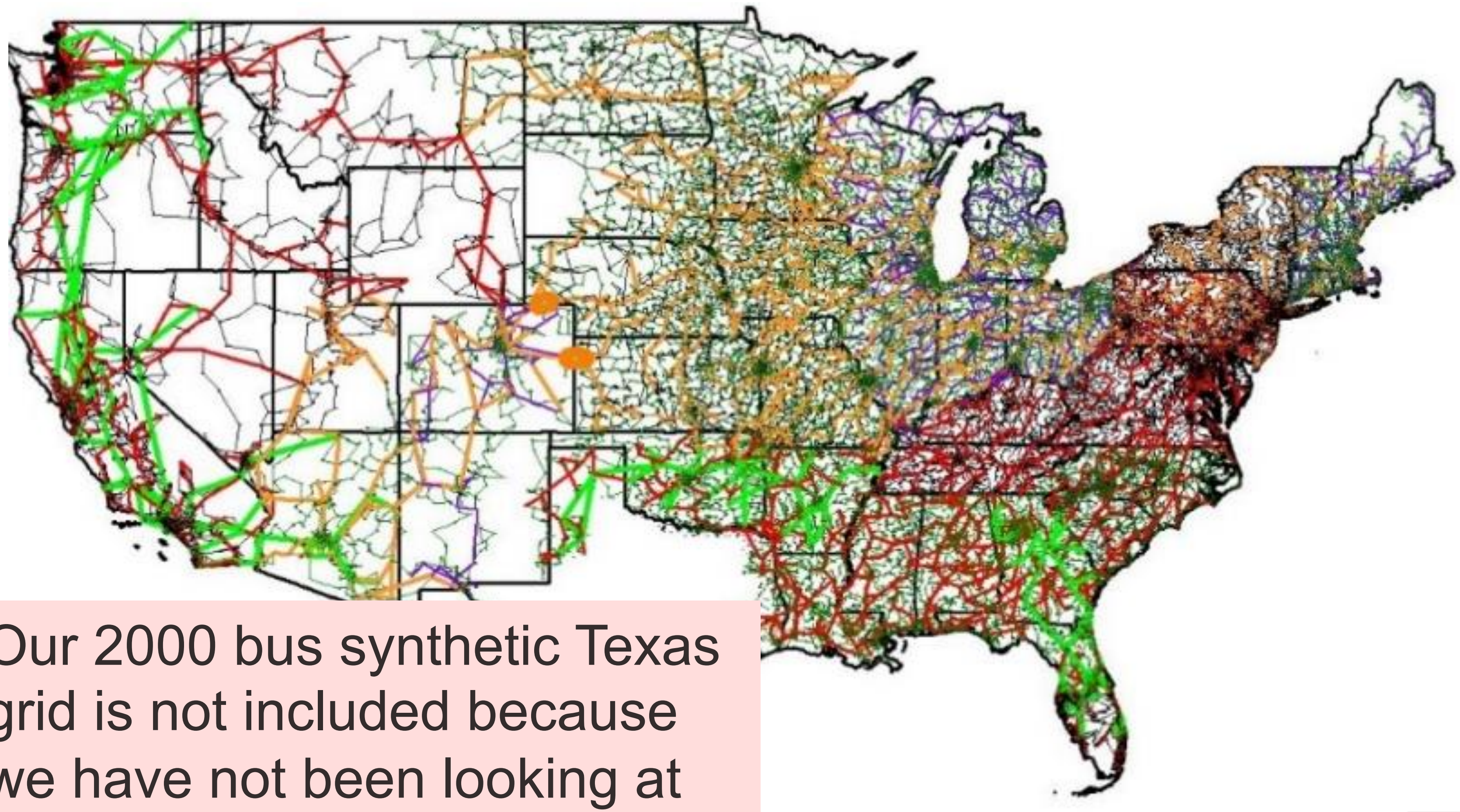
- This fictional grid, which has 2000 buses, is designed to serve a load similar to the ERCOT load with a similar geographic distribution
  - The grid was designed using a 500/230/161/115 kV transmission to be different from the actual grid
  - Public generator information is used



6. Diagram display for optimal power flow lab on the fictitious synthetic 2000-bus system. Green fields provide controls for the load scalar in seven of th
en areas, and report the average LMP for these areas. The background contour [45] shows that the locational marginal prices.



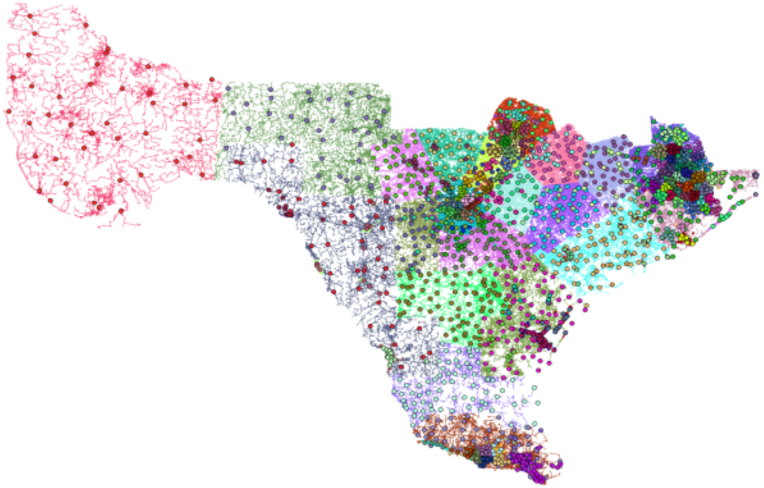
80,000 Bus Synthetic Grid for East-West Interconnection Studies



Our 2000 bus synthetic Texas grid is not included because we have not been looking at an ac intertie for Texas

Combined T&D Synthetic Grids

- While not covered here, distribution systems also need to be considered. Working with NREL and some other universities for ARPA-E we're creating combined T&D synthetic grids covering Texas down to the electric meter
- The finished grid will have about 50 million electric nodes
- The current grid has 8.4 million customers and 21.7 million electric nodes



Decision Making, Data, Information, Knowledge

- Ultimate goal is to help humans make better decisions
- Competing definitions for the process of taking raw “data” and producing something useful
  - Understanding, decisions, wisdom
- Data: symbols, raw, it simply exists
- Information: Data that is given meaning, often in a relational context; some how processed
- Knowledge: Application of information to answer “how.” Connecting patterns.
- Understanding, and/or wisdom at top

The Visualization Process

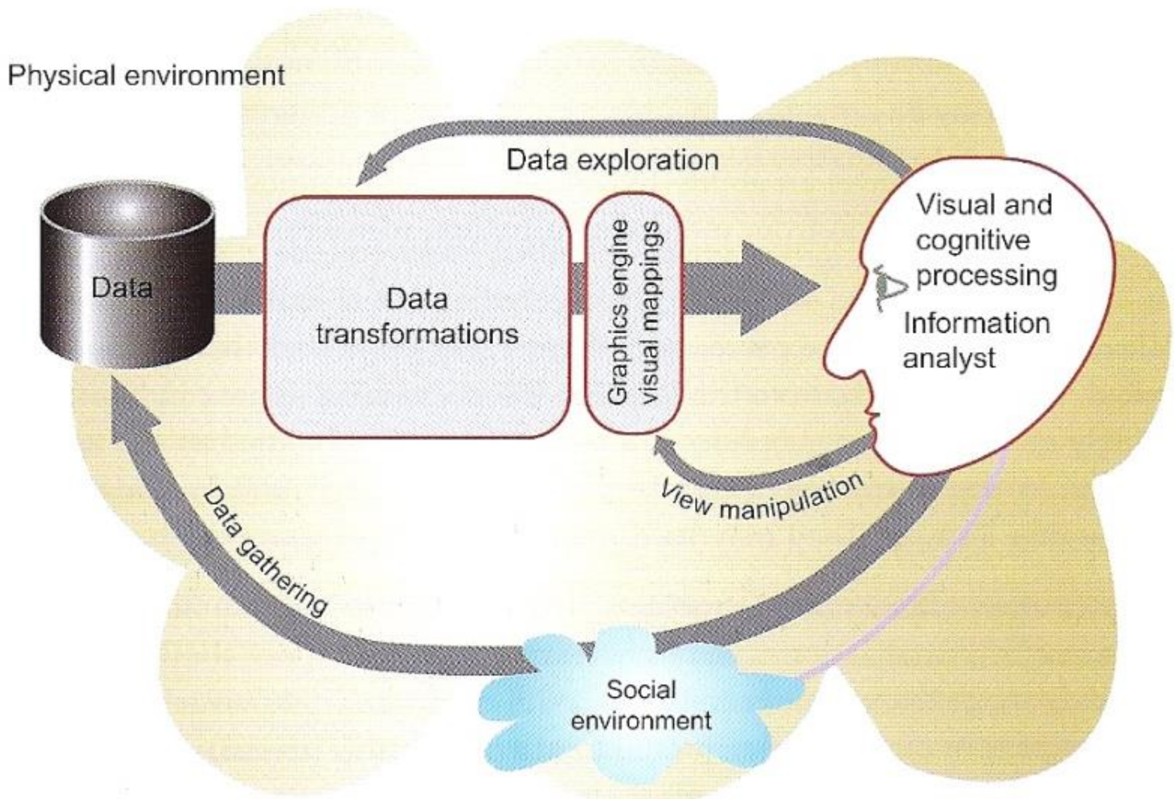


Figure 1.2 The visualization process.

Image source: Colin Ware, *Information Visualization*, Fourth Edition, 2021

Understanding the Entire Process is Key

- Understanding the entire processes in which the visualizations are embedded is key.
  - What is the “information access” cost?
  - How will the information be used and shared?
  - Is it raw data, or derived values?
  - Should the visualizations sit on top of a model, or is a standalone process sufficient?
  - Ultimately, what are the desired tasks that need to be accomplished?
- We’ll start with a brief coverage of some traditional approaches (tabular, graphs and onelines, then go into some newer ones)

Example: Tabular Displays

- In many contexts, tabular displays (particularly with interactive features such as sorting, filtering, drill-down, and the ability to enter data) can be a great way to show data

	Number	Area Name	Name	Nom kV	PU Volt	Volt (kV)	Angle (Deg)	Load MW	Load Mvar	Gen MW	Gen Mvar	Switched Shunts Mvar	Act G Shunt MW	Act B Shunt Mvar	Area Num	Zone Name	Zone Num
1	1001	Far West	ODESSA 2 0	115.00	0.98089	112.802	-30.18	20.78	5.89				0.00	0.00	1	Far West T	9
2	1002	Far West	PRESIDIO 2 0	115.00	1.01218	116.400	-24.75	15.41	4.37				0.00	0.00	1	Far West T	9
3	1003	Far West	O DONNELL 1	115.00	1.00832	115.956	-25.02						0.00	0.00	1	Far West T	9
4	1004	Far West	O DONNELL 1	230.00	1.01000	232.301	-26.84			158.25	-29.07		0.00	0.00	1	Far West T	9
5	1005	Far West	BIG SPRING 5 (	115.00	1.00790	115.908	-22.77						0.00	0.00	1	Far West T	9
6	1006	Far West	BIG SPRING 5 '	13.80	1.00147	13.820	-20.60			25.73	-4.94		0.00	0.00	1	Far West T	9
7	1007	Far West	VAN HORN 0	115.00	1.01973	117.268	-25.10	7.01	1.99			0.00	0.00	0.00	1	Far West T	9
8	1008	Far West	IRAAN 2 0	115.00	1.00133	115.153	-13.78						0.00	0.00	1	Far West T	9
9	1009	Far West	IRAAN 2 1	13.80	1.00000	13.800	-10.41			61.87	-2.55		0.00	0.00	1	Far West T	9
10	1010	Far West	PRESIDIO 1 0	115.00	1.01933	117.223	-23.46					0.00	0.00	0.00	1	Far West T	9
11	1011	Far West	PRESIDIO 1 1	22.00	1.01958	22.431	-22.12			7.50	0.00		0.00	0.00	1	Far West T	9
12	1012	Far West	SANDERSON (	115.00	0.98899	113.734	-29.67	2.99	0.85			9.29	0.00	0.00	1	Far West T	9
13	1013	Far West	MONAHANS 2	115.00	1.00167	115.192	-21.95	29.23	8.28				0.00	0.00	1	Far West T	9
14	1014	Far West	GRANDFALLS (	115.00	1.00324	115.373	-18.04	2.22	0.63				0.00	0.00	1	Far West T	9
15	1015	Far West	MARFA 0	115.00	1.02132	117.451	-24.87	7.51	2.13				0.00	0.00	1	Far West T	9
16	1016	Far West	GARDEN CITY (	115.00	1.01758	117.022	-21.94	2.89	0.82			31.06	0.00	0.00	1	Far West T	9
17	1017	Far West	ODESSA 4 0	115.00	0.98205	112.936	-28.53	18.34	5.20				0.00	0.00	1	Far West T	9
18	1018	Far West	NOTREES 0	115.00	0.99128	113.997	-27.25	0.07	0.02				0.00	0.00	1	Far West T	9
19	1019	Far West	MIDLAND 4 0	115.00	1.00078	115.090	-29.70	61.78	17.50			143.20	0.00	0.00	1	Far West T	9
20	1020	Far West	BIG SPRING 1 (	115.00	1.02190	117.519	-21.73					80.13	0.00	0.00	1	Far West T	9
21	1021	Far West	BIG SPRING 1 '	13.80	1.00000	13.800	-15.11			149.63	-25.59		0.00	0.00	1	Far West T	9
22	1022	Far West	O DONNELL 2	115.00	1.01132	116.302	-24.18						0.00	0.00	1	Far West T	9
23	1023	Far West	O DONNELL 2	13.80	1.01000	13.938	-15.27			135.00	3.21		0.00	0.00	1	Far West T	9
24	1024	Far West	ODESSA 6 0	115.00	0.99425	114.338	-26.17	63.04	17.86				0.00	0.00	1	Far West T	9
25	1025	Far West	BIG SPRINGS 0	115.00	1.01805	117.076	-20.73						0.00	0.00	1	Far West T	9
26	1026	Far West	BIG SPRINGS 1	13.80	1.00000	13.800	-11.21			93.15	-4.41		0.00	0.00	1	Far West T	9
27	1027	Far West	MIDLAND 2 0	115.00	1.01258	116.447	-32.98	101.21	28.68			76.90	0.00	0.00	1	Far West T	9
28	1028	Far West	COAHOMA 0	115.00	1.01371	116.577	-25.80	10.01	2.84				0.00	0.00	1	Far West T	9
29	1029	Far West	MIDLAND 3 0	115.00	1.00868	115.998	-31.93	83.18	23.57			40.70	0.00	0.00	1	Far West T	9

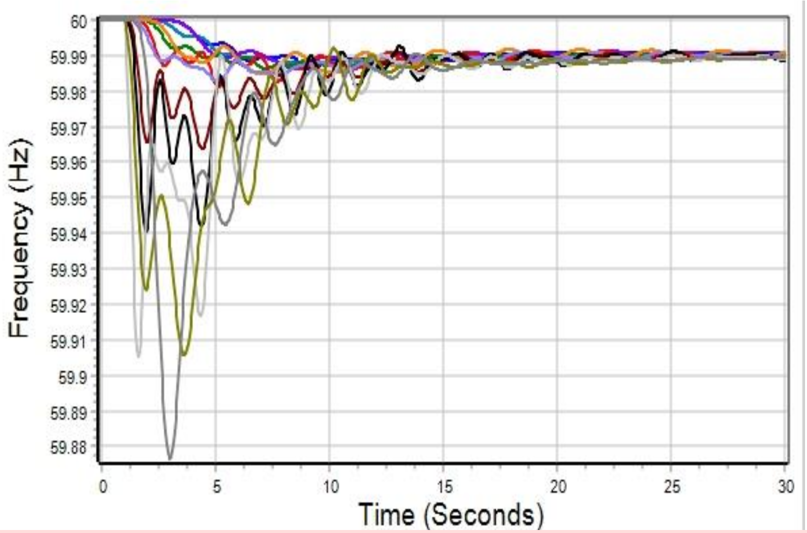
Use of Color

- Some use of color can be quite helpful
  - 10% of male population has some degree of color blindness (1% for females)
- Do not use more than about ten colors for coding if reliable identification is required
- Color sequences can be used effectively for data maps (like contours)
  - Grayscale is useful for showing forms but not values
  - Multi-color scales (like a spectrum) have advantages (more steps) but also disadvantages (effectively comparing values) compared to bi-color sequences

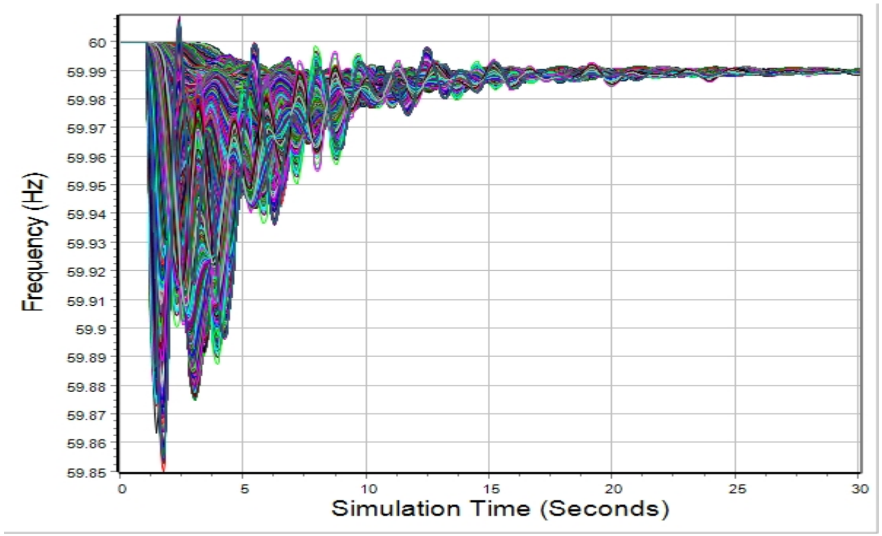
The book by Colin Ware is a great resource

Graphs

- Graphs are also a great way to show information, particularly for time-variation
- The number of curves needs to match the task

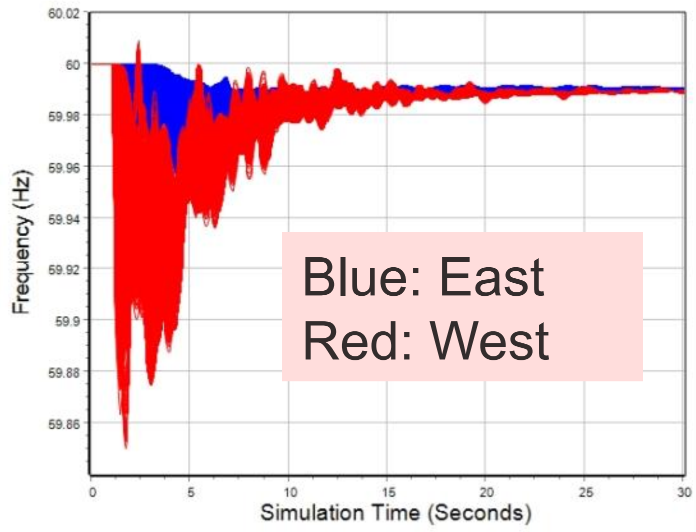


A few curves, detail of each visible, key can identify objects (several thousand values)

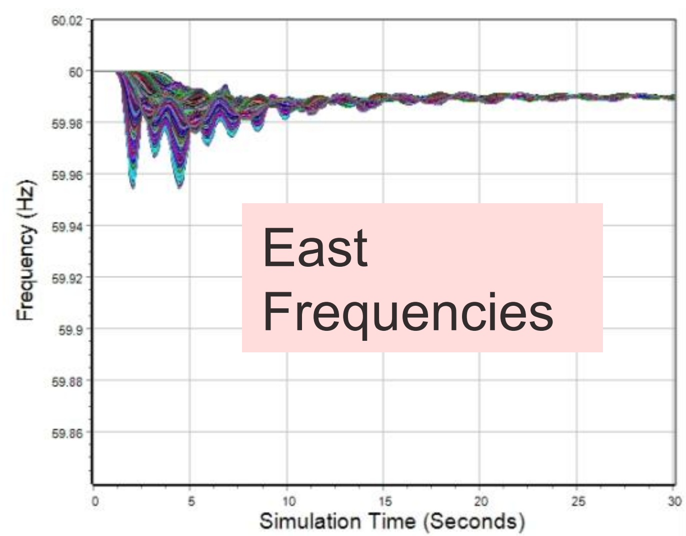


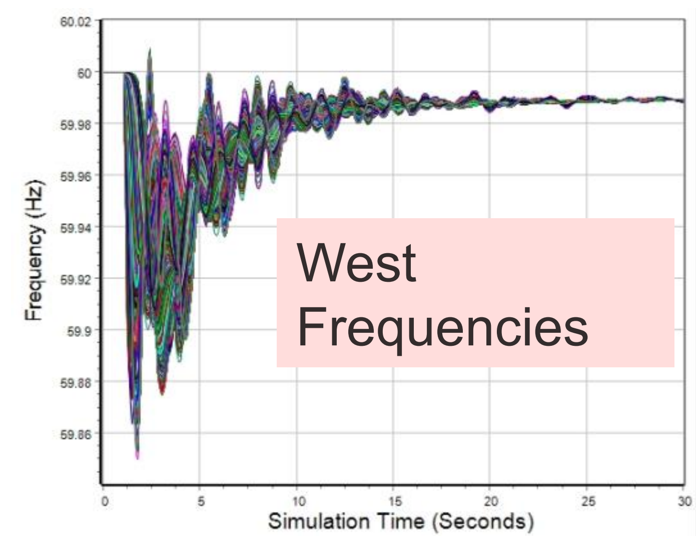
Envelope of response for the 80k bus, 40,000 substation frequencies (24 million values)

Graphs: 40,000 Substation Examples



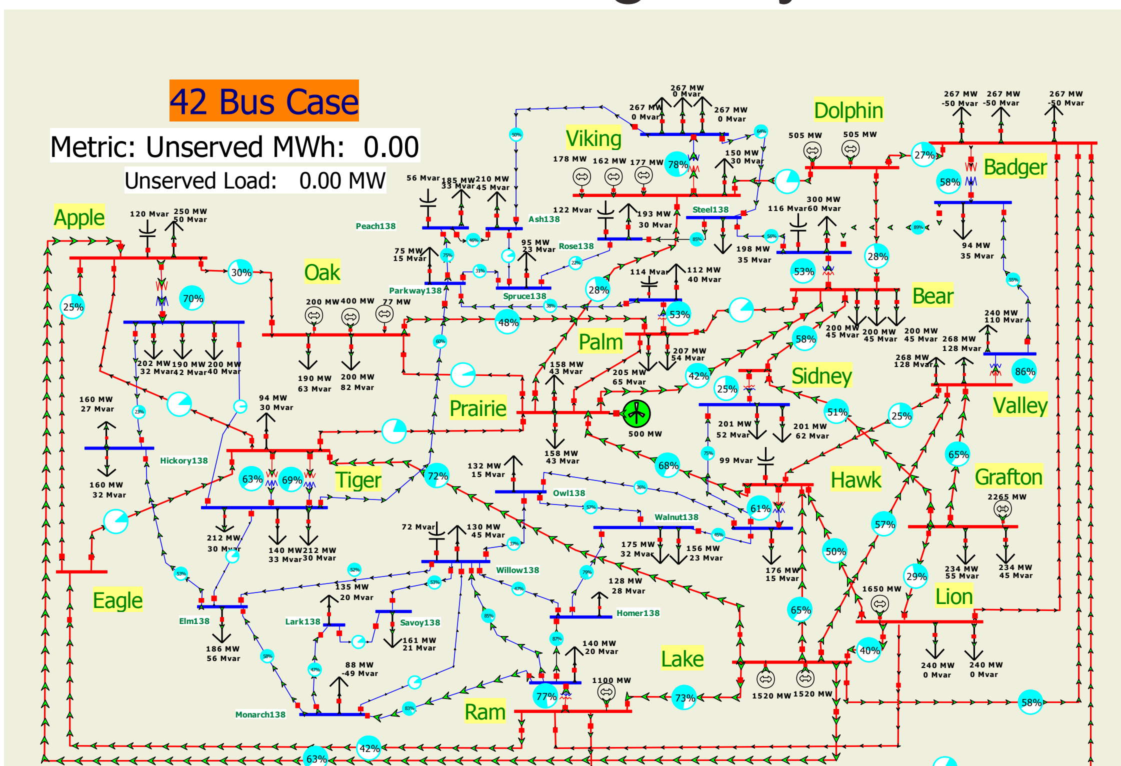
For the 40,000 substation plot, color can be helpful in showing the East response (blue) versus West (red) but the curve order matters. It is probably better to use two plots, with one for the East and one for the West (obviously using the same scale)





Onelines

- Widely used and can be quite effective for showing substations (or local regions) or smaller grids; can be slow on larger systems



Visualization Background: Preattentive Processing

- When displaying large amounts of data, take advantage of preattentive cognitive processing
  - With preattentive processing the time spent to find a “target” is independent of the number of distractors
- Graphical features that are preattentively processed include the general categories of form, color, motion, spatial position

Preattentive Processing Examples

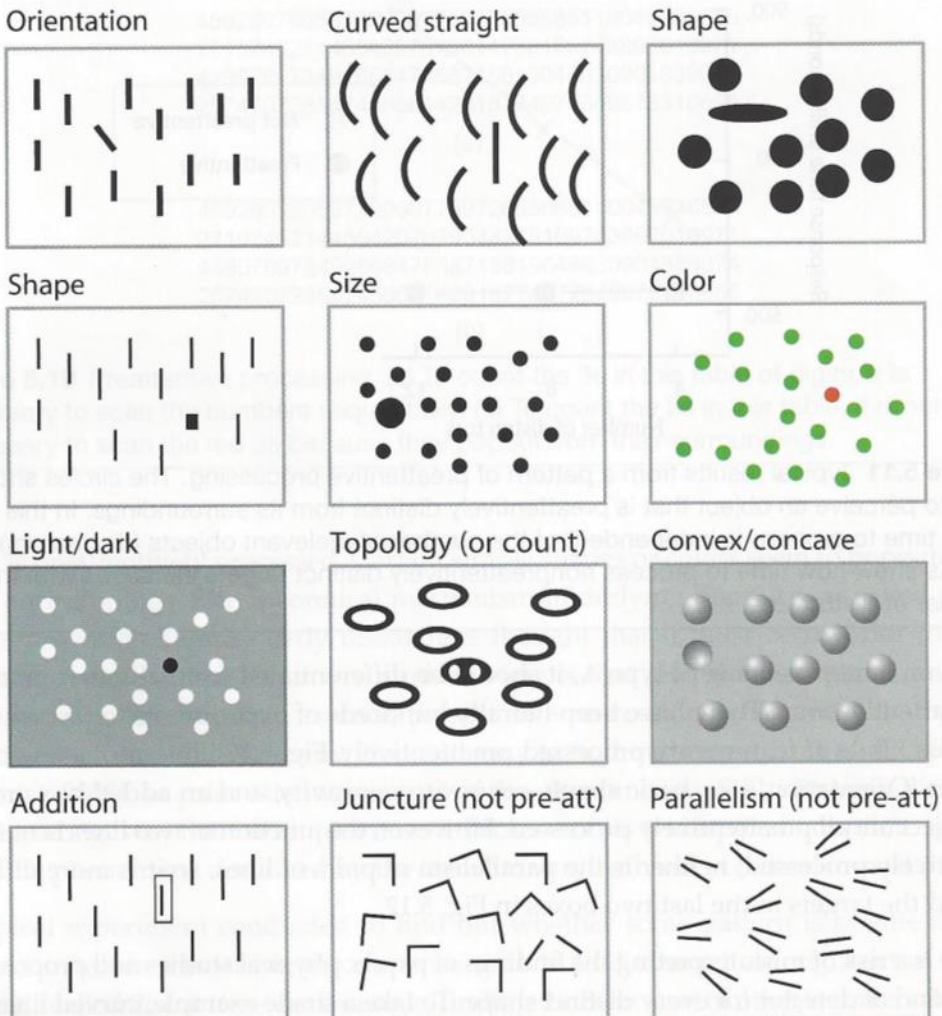
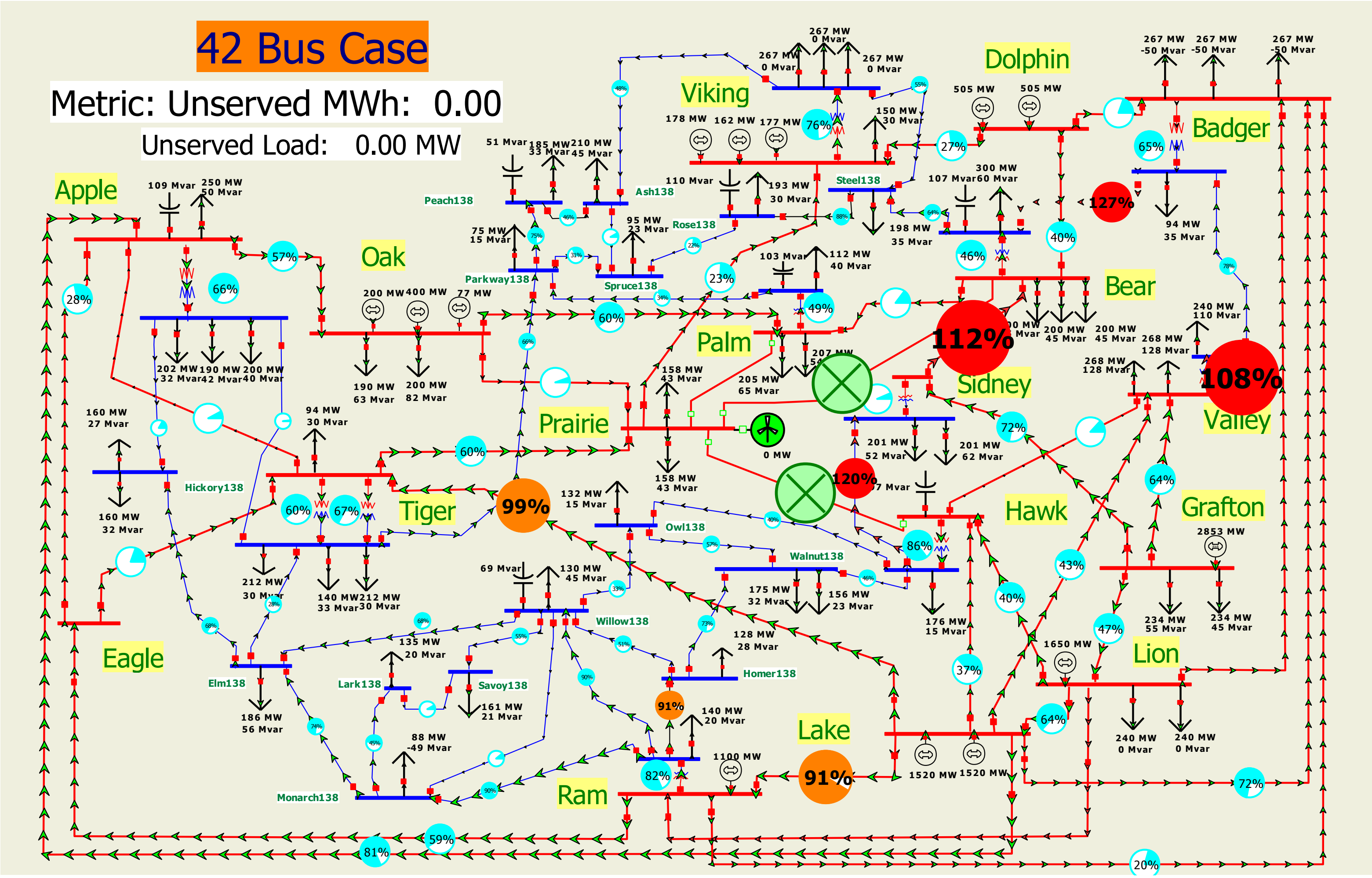


Figure 5.12 Most of the preattentive examples given here can be accounted for by the processing characteristics of neurons in the primary visual cortex.

All are preattentively processed except for juncture and parallelism; however too many can defeat their purpose

Source: *Information Visualization* (Fourth Edition) by Colin Ware, Fig 5.12

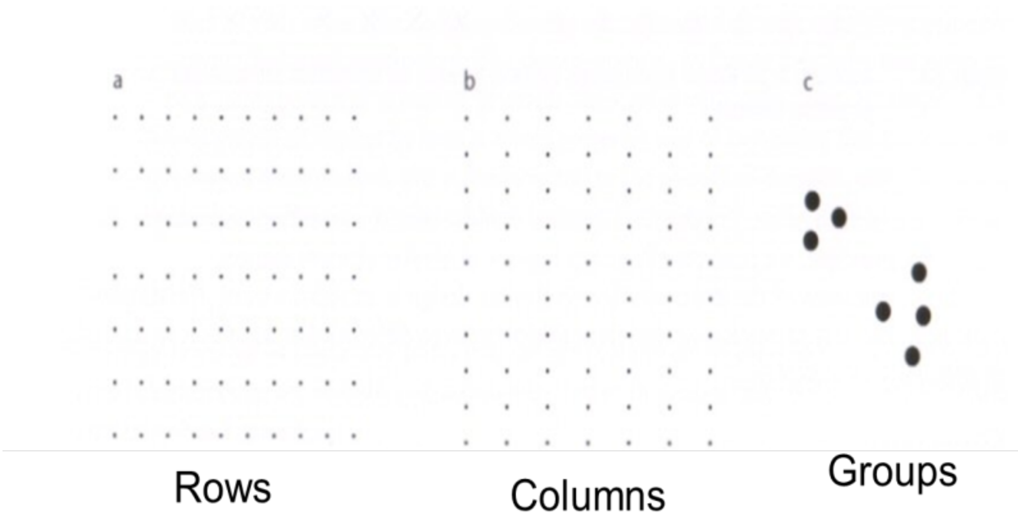
Preattentive Processing with Color & Size

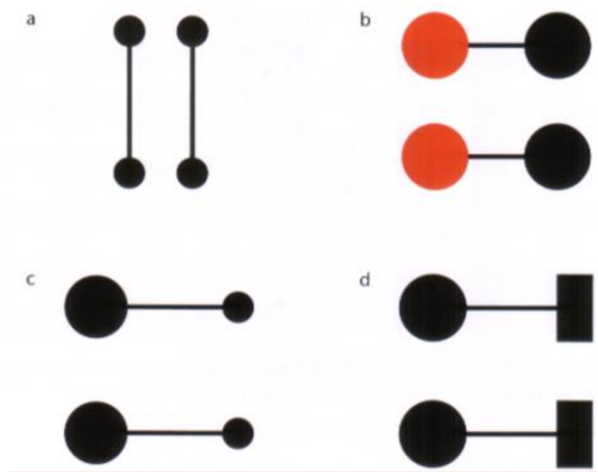


Detecting Patterns

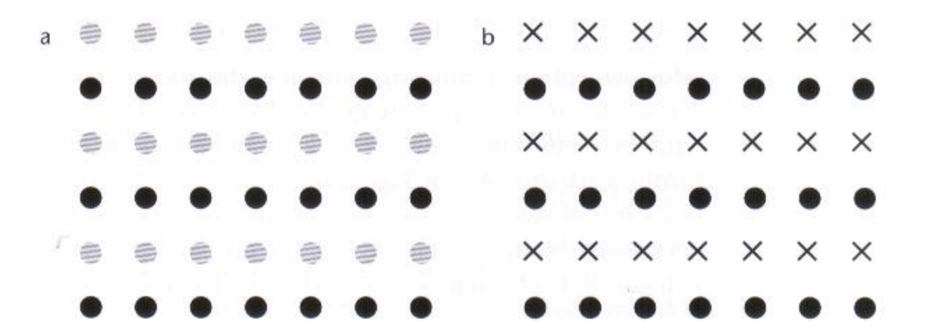
- A large portion of information visualization is associated with detecting patterns
- Gestalt (German for "pattern") Laws
  - Proximity
  - Similarity (we didn't discuss color)
  - Connectedness
  - Common Fate (flows)

Proximity, Similarity, Connectedness,





Connectedness is stronger than proximity, color, shape

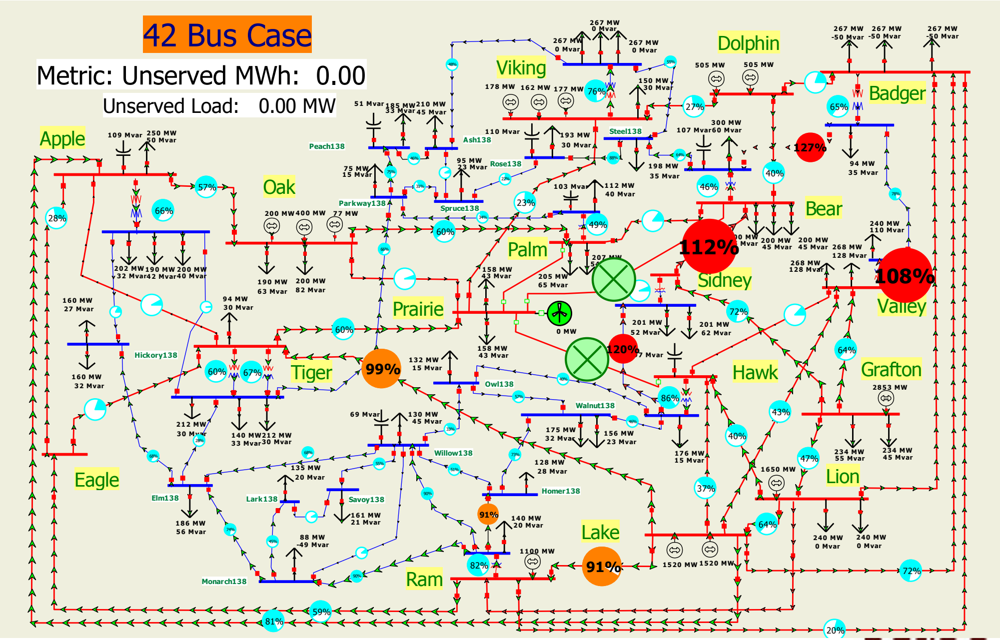


Similarity makes all perceived as rows

Source: *Information Visualization* (Fourth Edition) by Colin Ware, Chapter 6 Images

Common Fate: Patterns in Motion

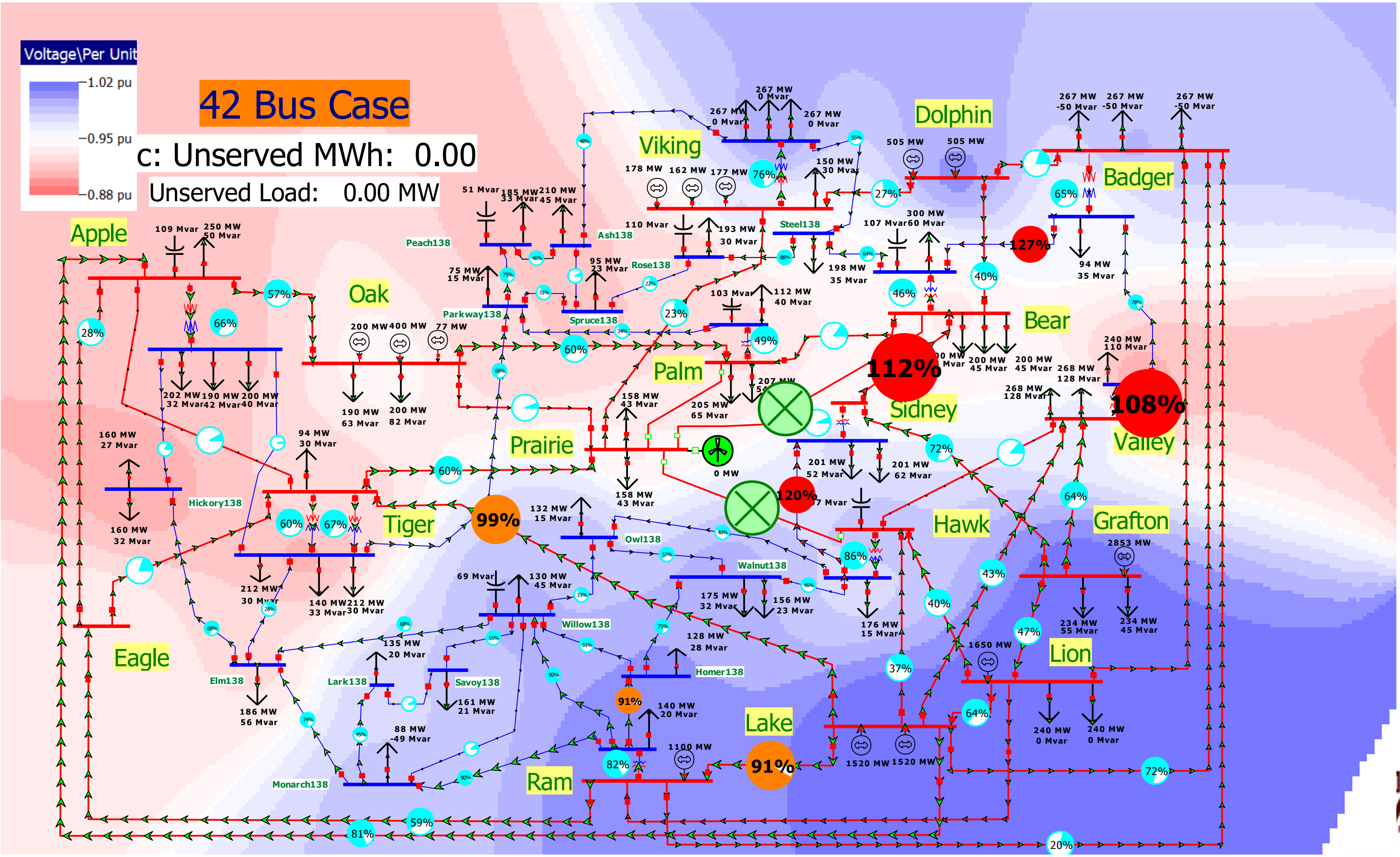
- Motion can be a very effective means for showing relationships between data
- People perceive motion with great sensitivity
- Motion can also be used to convey causality (one event causing another)
- However, too much motion can be a distractor



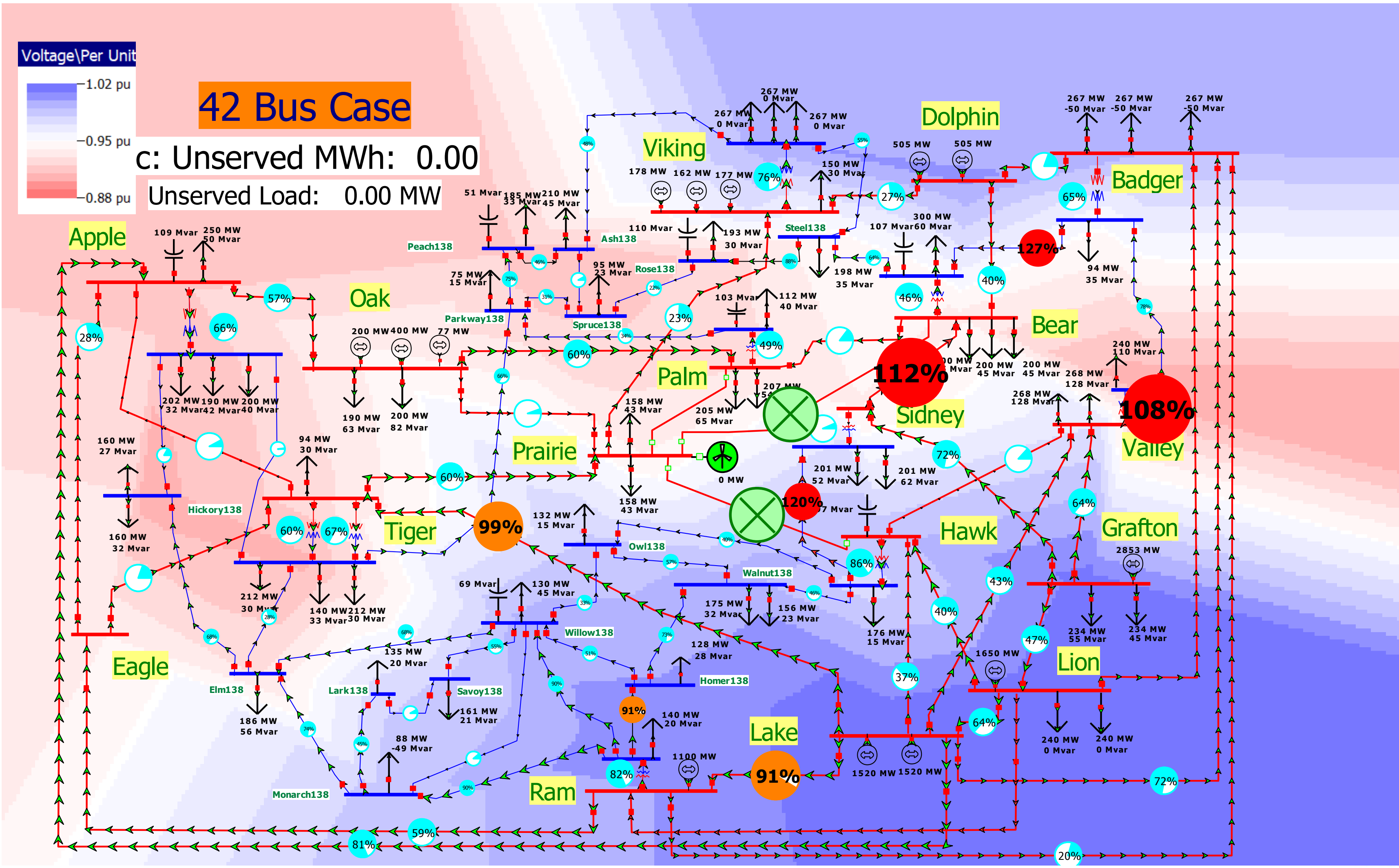
Scattered Data Interpolation (Colored Contouring)

- For wide-area visualization, contours can be effective for showing large amounts of spatial data
  - Takes advantage that as humans we perceive the world in patterns (sometimes even when none exist!)
  - Now widely used
- Scattered data interpolation algorithms are needed to take the discrete power system data and make it spatially continuous
  - Various algorithms can be used include a modified Shephard's and Delaunay triangulation
- A color mapping is needed

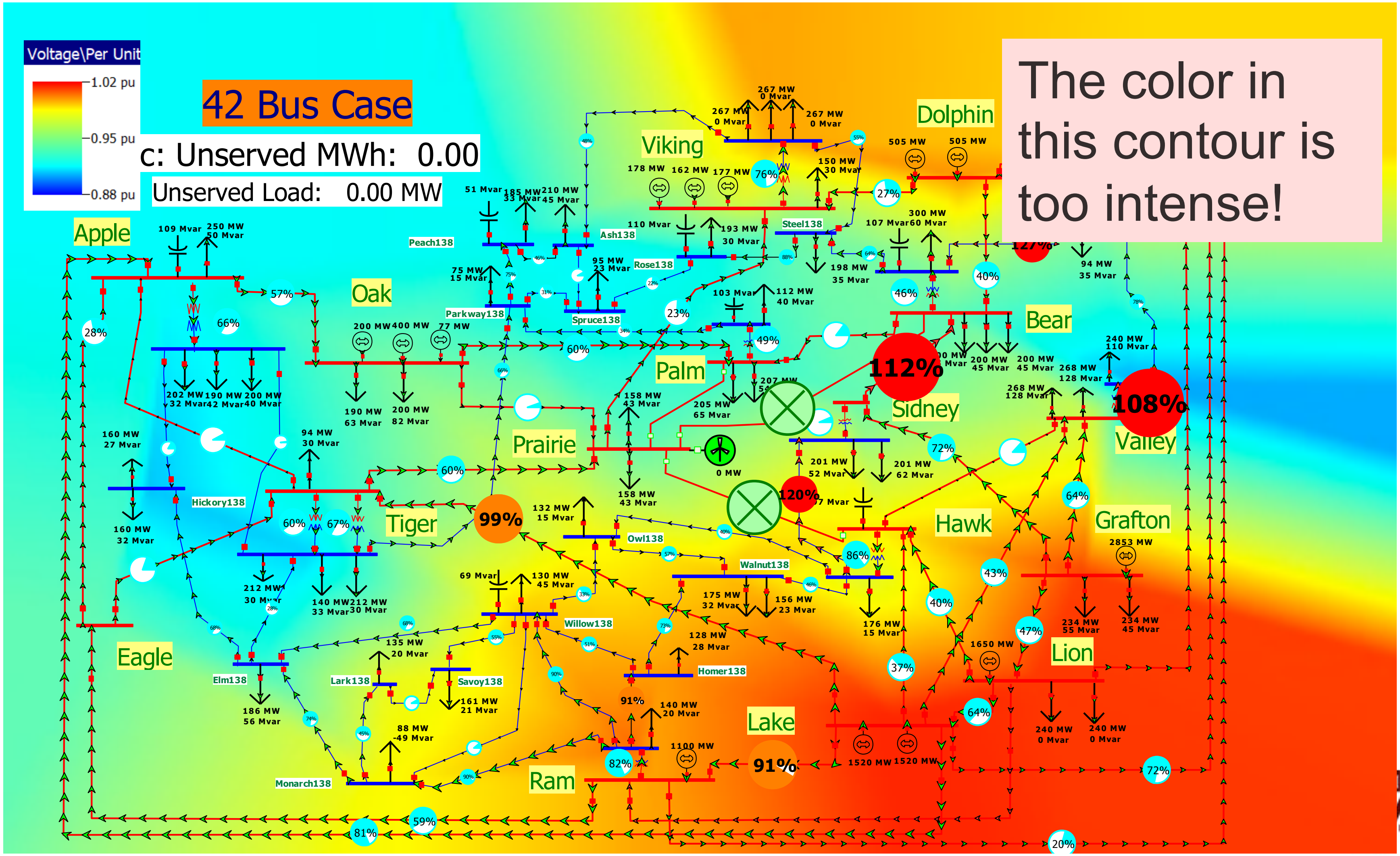
Shepard's Algorithm, Blue/Red Discrete Color mapping



Delaunay Algorithm, Blue/Red Discrete Color mapping



Delaunay Algorithm, Spectrum Continuous Color mapping



The color in this contour is too intense!

Some General Thoughts on Power System Visualizations

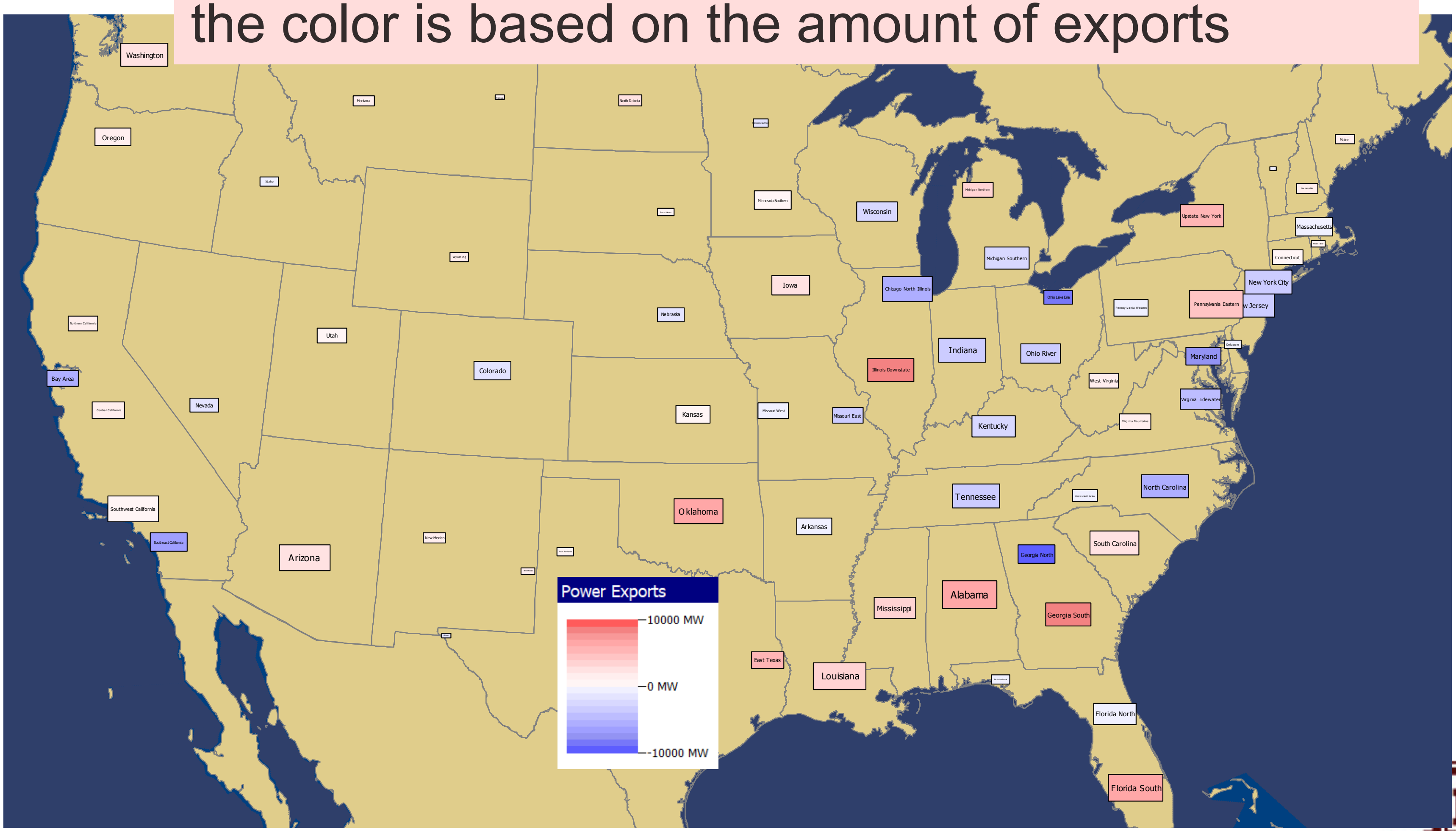
- While the previous techniques can be quite helpful, there is often just too much data to display
- Interactive visualizations, taking advantage of the underlying geographic information, can be quite effective, particularly if the displays can be rapidly customized to show different sets of information
- Also, much of the data should first be pre-processed using potentially quite sophisticated algorithms

Geographic Data Views

- One way to make visualizations more interactive is to use underlying geographic information to quickly auto-create displays
  - Known as geographic data views (GDVs)
- GDVs can be used either on individual objects (like generators, buses, or substations), or on aggregate objects (like areas and zones)
- The GDV display attributes (e.g., size, color) can be used to show object data
- The GDV displays can be saved for later use and links to the underlying objects allow for drilldown

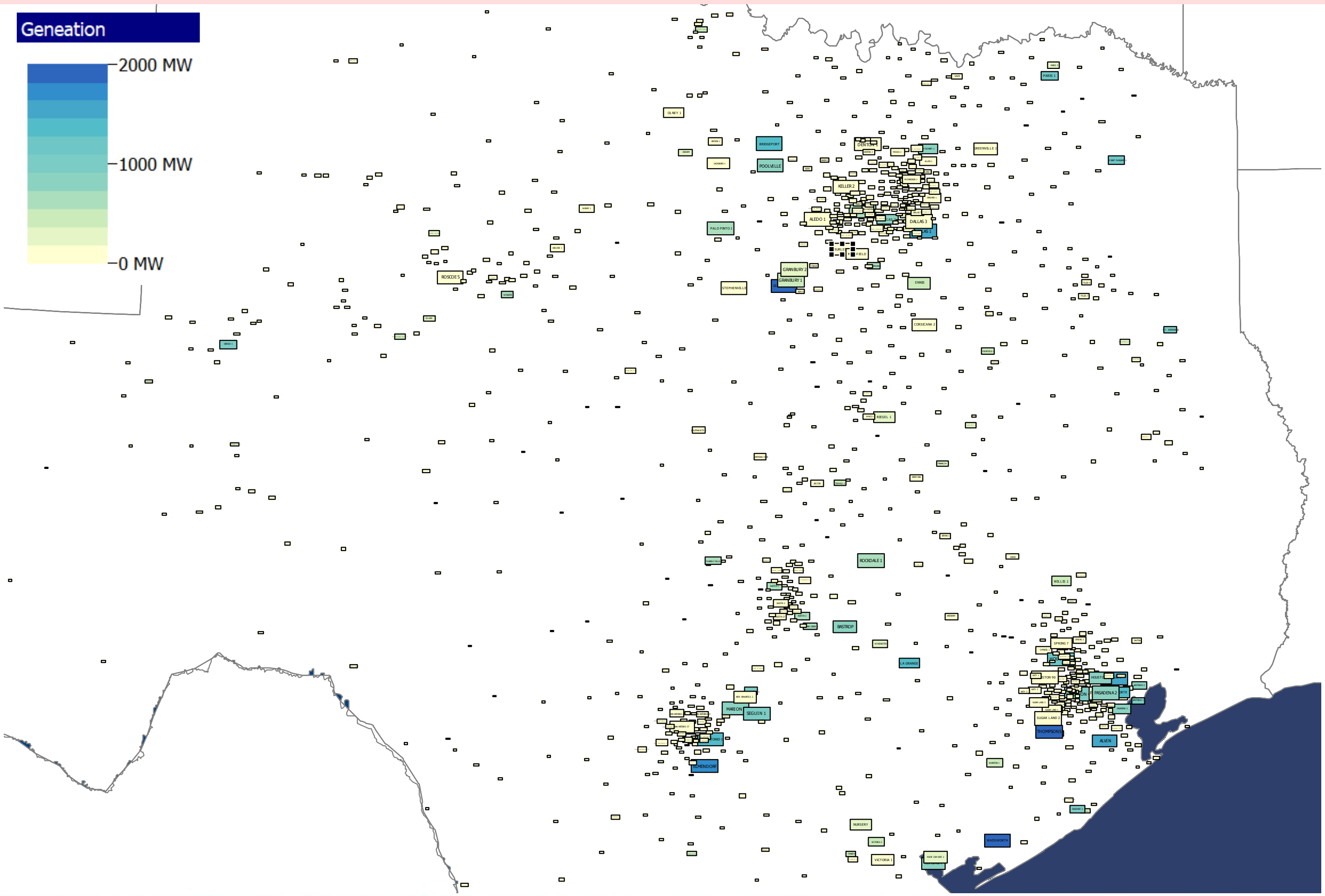
80,000 Bus System Area GDV Example

Size is proportional to the area's generation, while the color is based on the amount of exports



Texas 2000 Substation GDV

Size is proportional to the substation MW throughput, while the color is based on the amount of substation generation



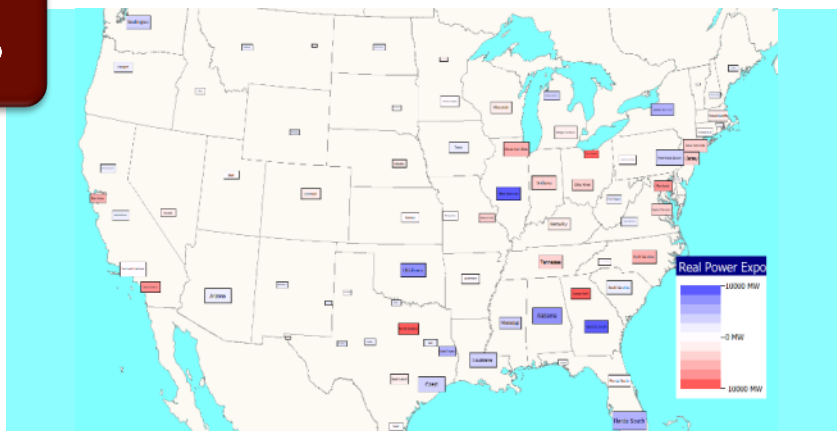
Pseudo-Geographic Mosaic Displays

- GDVs can be quite useful, but there is a tradeoff between geographic accuracy and maximum display space usage
  - Much of the electric grid is concentrated in small (primarily urban) areas
- Pseudo-Geographic Mosaic Displays (PGMDs) utilize a tradeoff of geographic accuracy to maximize display space

80,000 Bus System Area PGMD

GDVs are as before (size = area gen MW, color = interchange); the percentages show the amount of transition

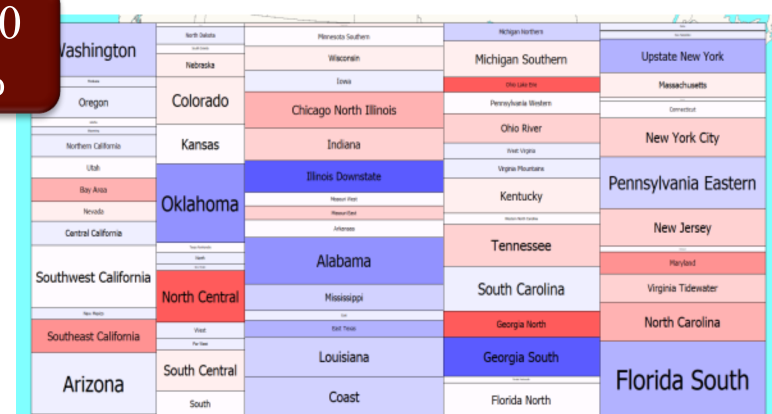
0%



25%

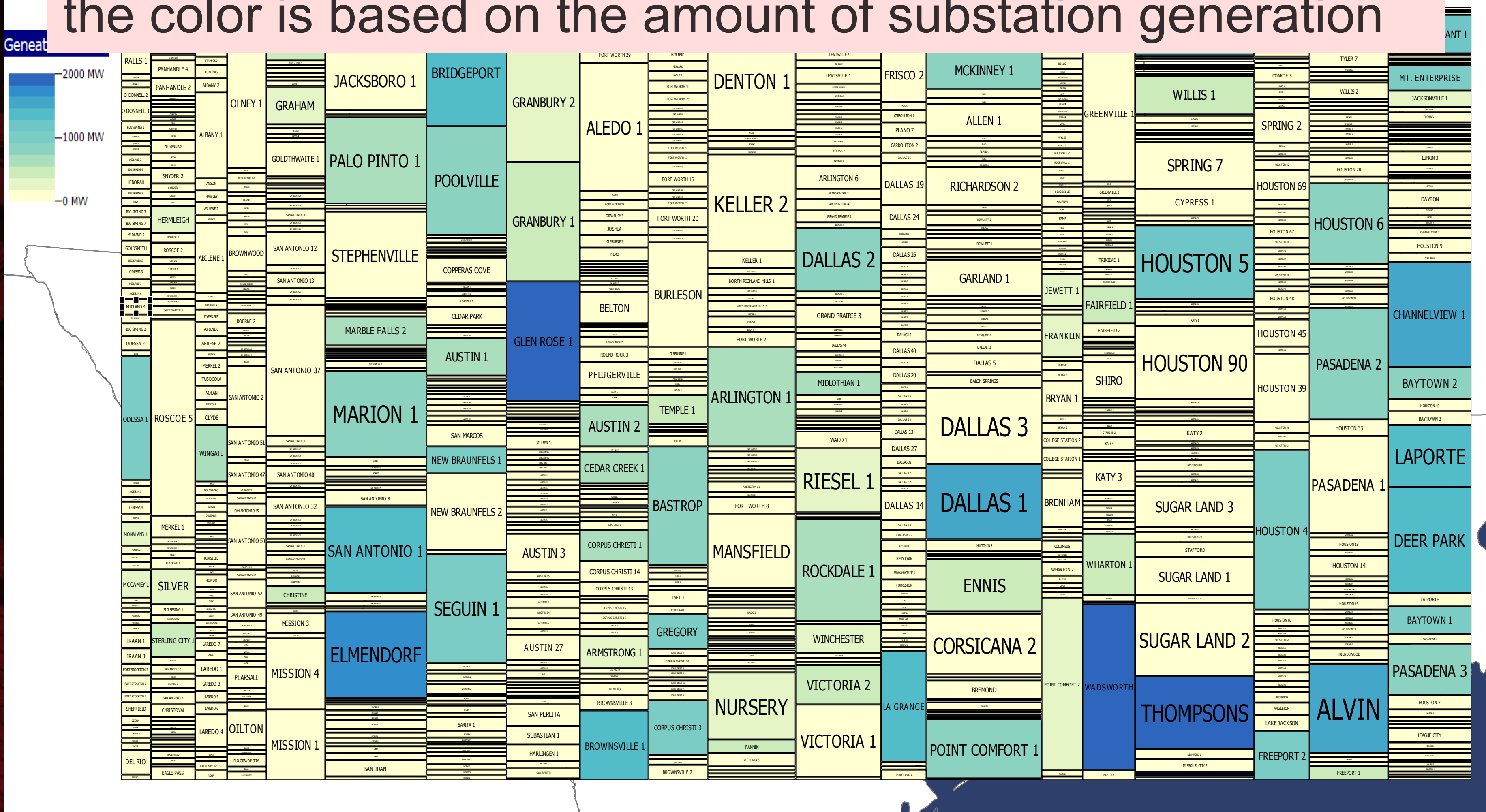
60%

100 %



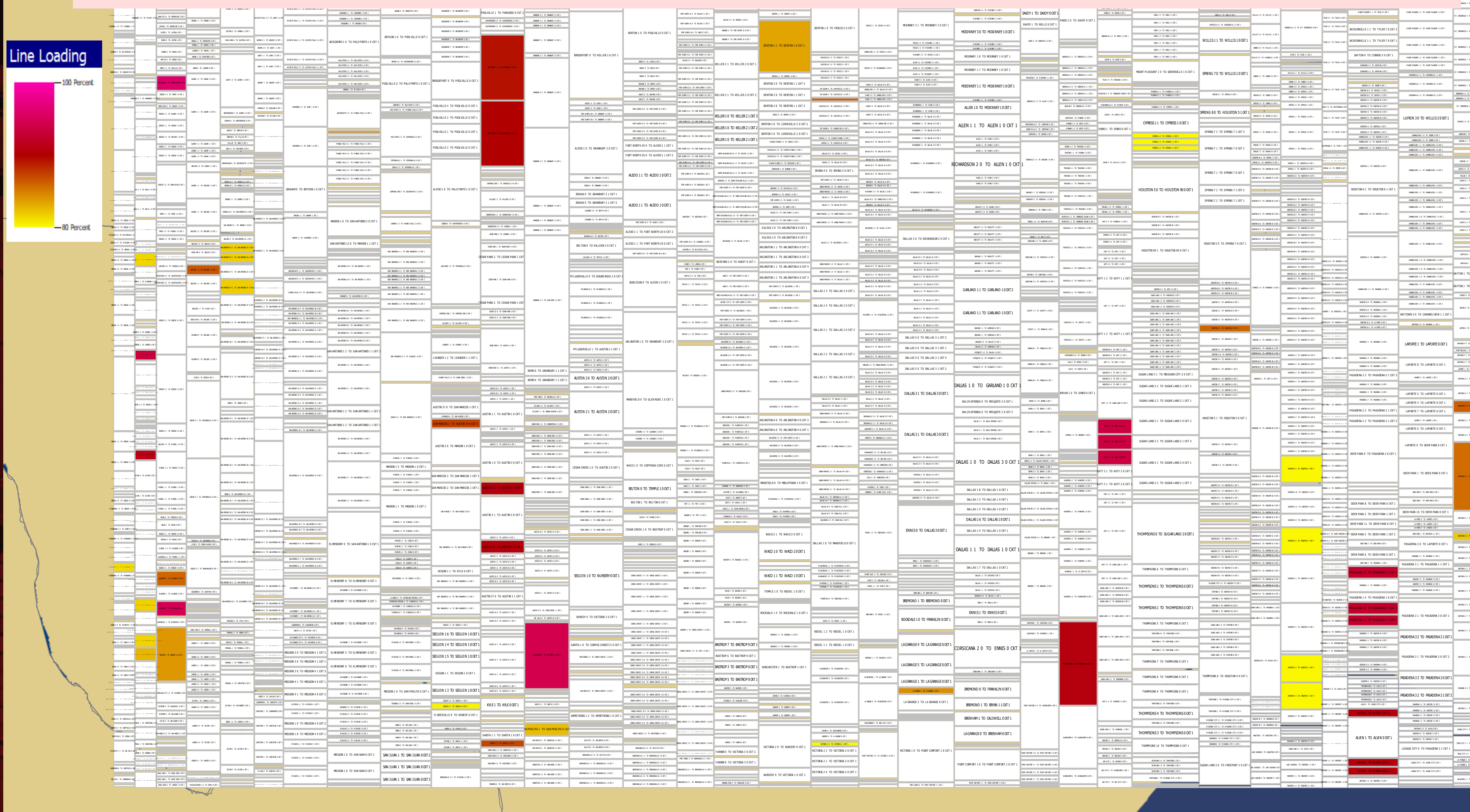
Texas 2000 Substation PGMD

Size is proportional to the substation MW throughput, while the color is based on the amount of substation generation

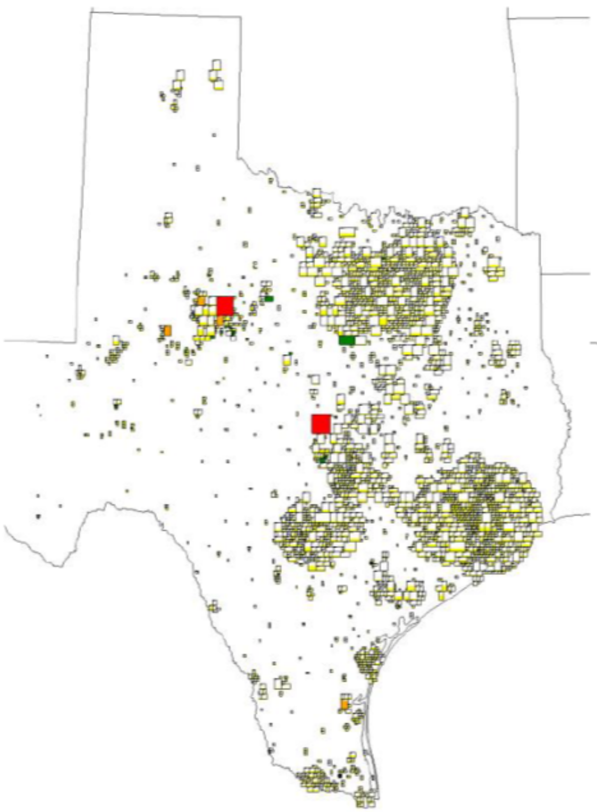


Texas 2000 Line Loading PGMD

Size is proportional to the Line MVA, the color is based on percentage loading; display shows data for 3200 lines



Mosaic Packing Generalize the PGMD to Allow a Specified Percentage Fill



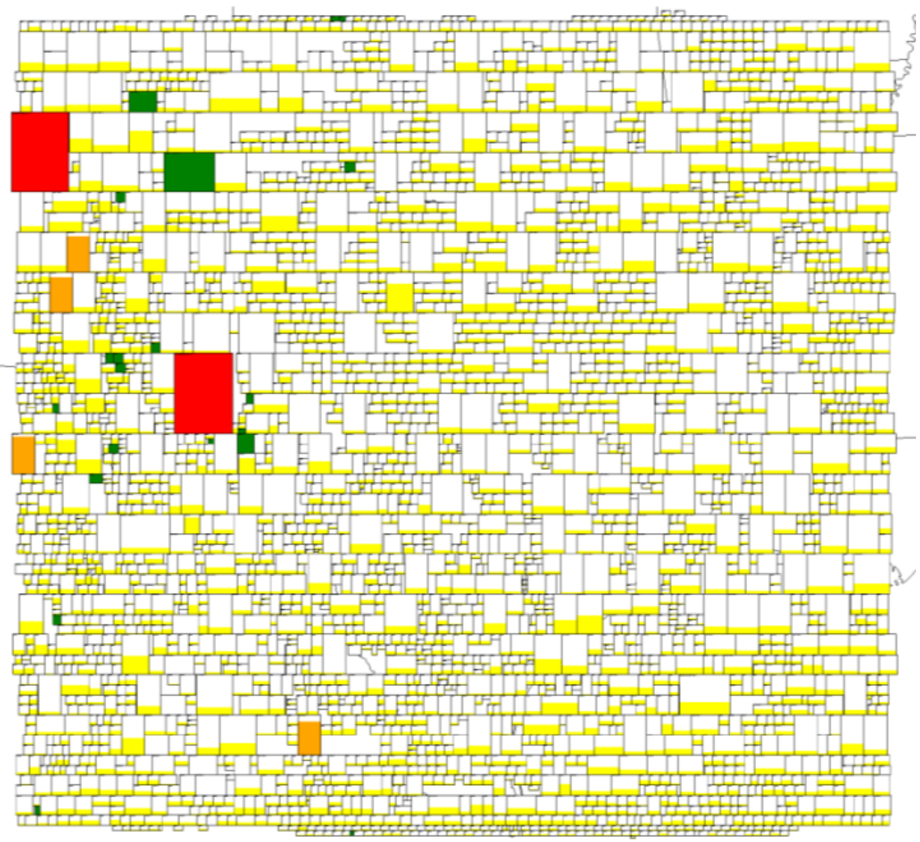


Image is Figure 12 from A.B. Birchfield and T.J. Overbye, “Mosaic Packing to Visualize Large-Scale Electric Grid Data,” accepted in *IEEE Open Access Journal of Power and Energy*, Jun. 2020

Some Techniques for Dealing with Time-Varying Data

- Need to keep in mind the desired task!
- Tabular displays
- Time-based graphs (strip-charts for real-time)
- Animation loops
  - Can be quite effective with contours, but can be used with other types of data as well
- Data analysis algorithms, such as clustering, to detect unknown properties in the data
  - There is often too much data to make sense without some pre-processing analysis!

Animation loops

- Animation loops trade-off the advantages of snapshot visualizations with the time needed to play the animation loop
  - A common use is in weather forecasting
- In power systems applications the length/speed of the animation loops would depend on application
  - In real-time displays could update at either SCADA or PMU rates
  - Could be played substantially faster than real-time to show historical or perhaps anticipated future conditions

Wide-Area Contours Can be Quickly Created Using GDVs

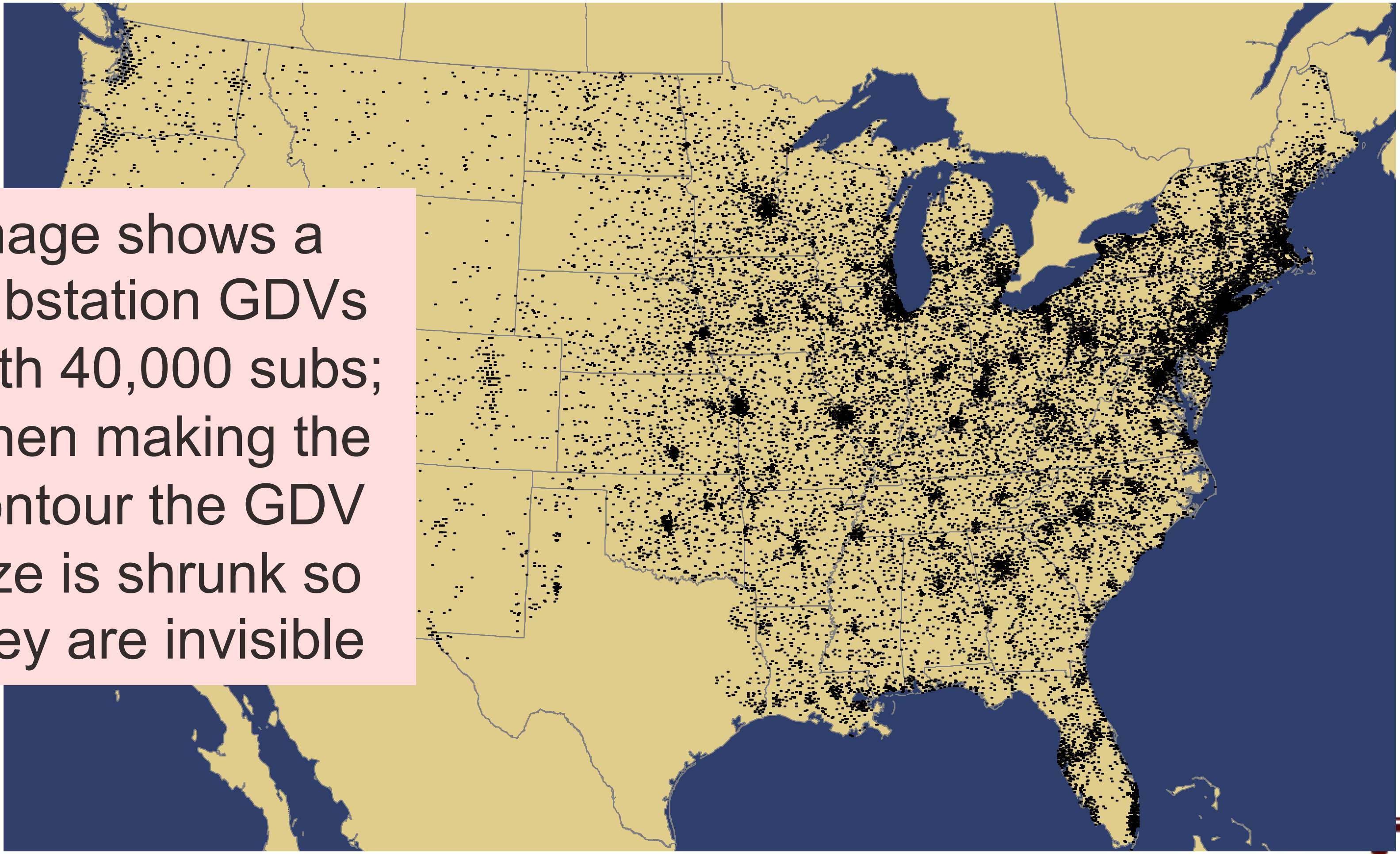
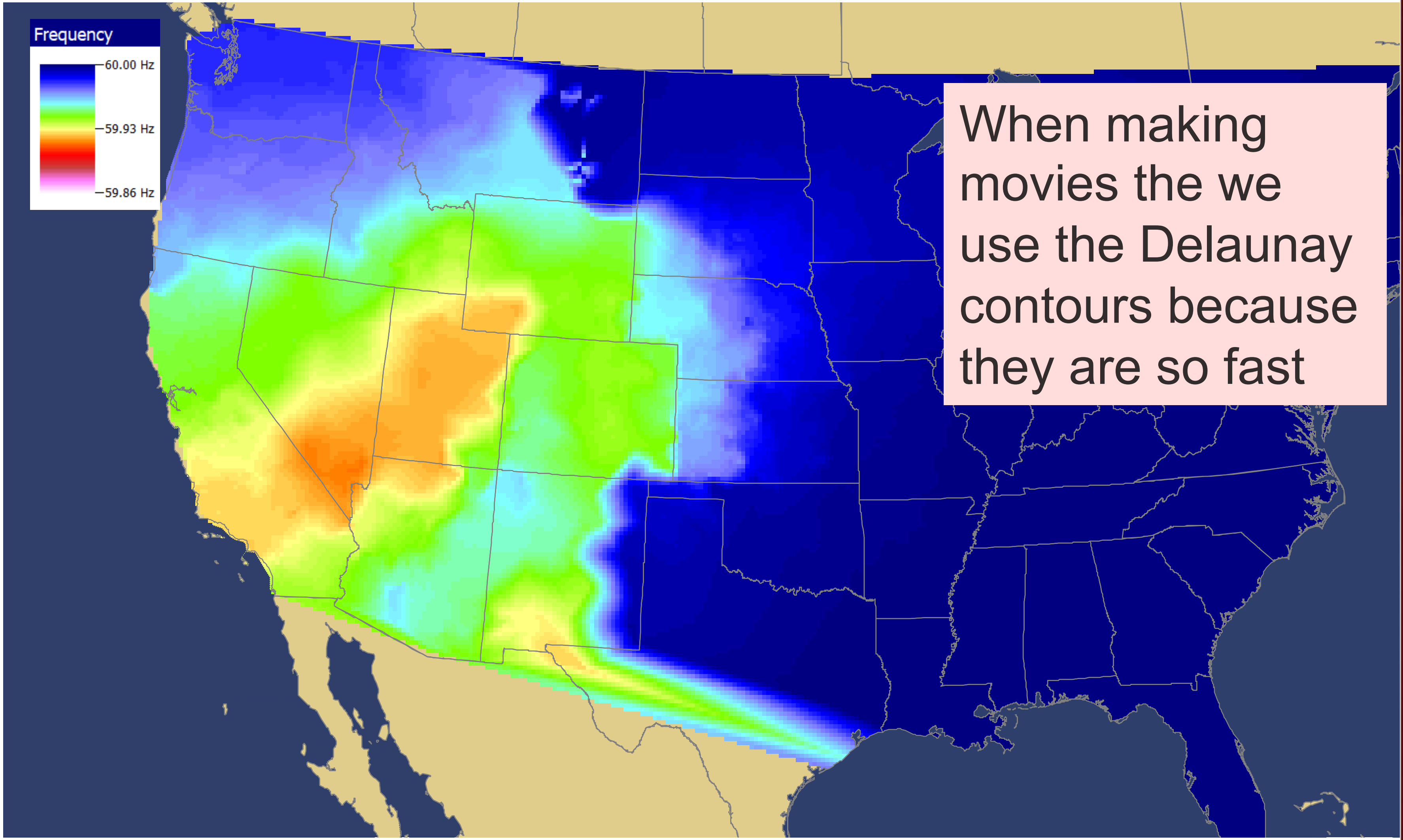


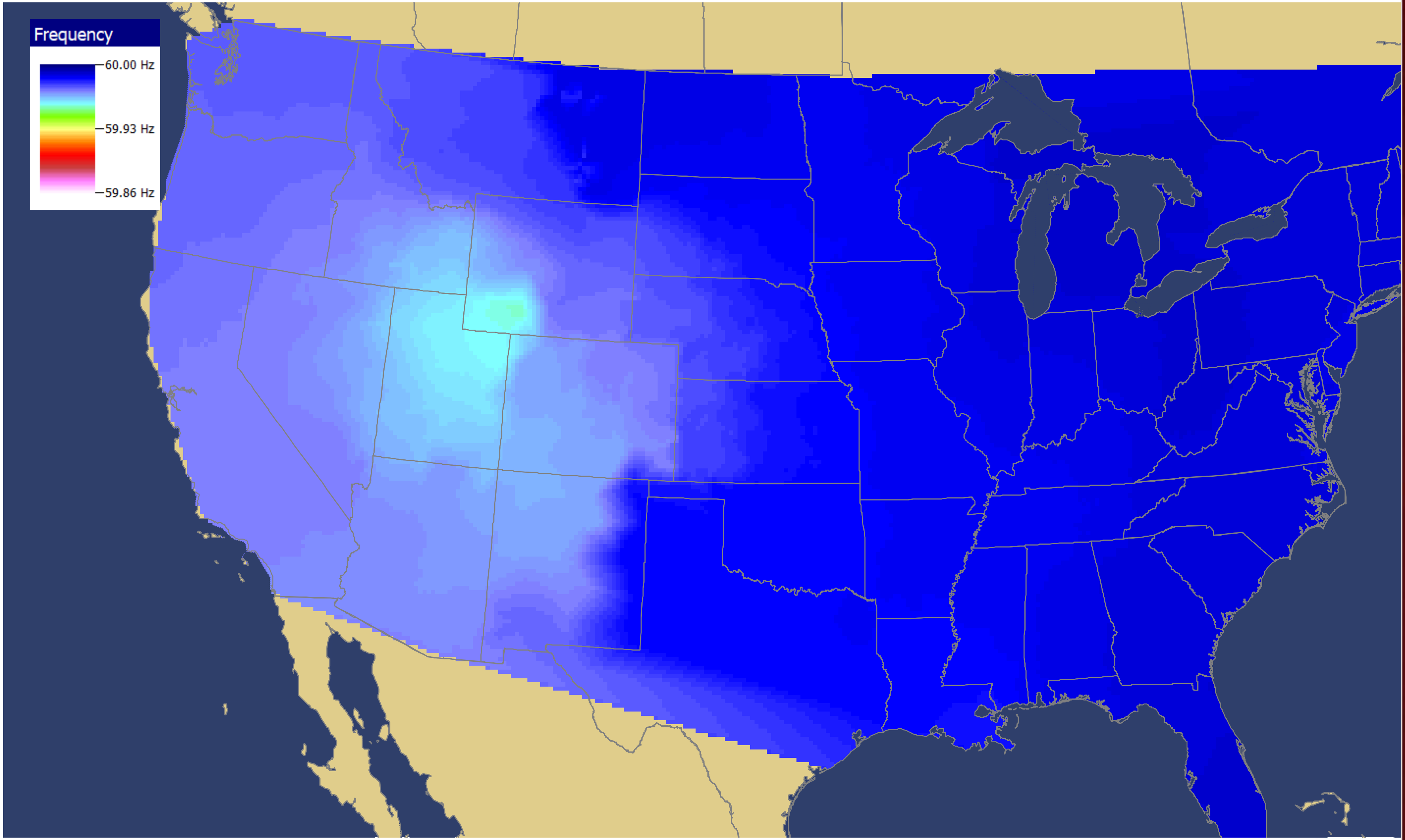
Image shows a substation GDVs with 40,000 subs; when making the contour the GDV size is shrunk so they are invisible

80K-Bus Frequency Variation (Substation Contour; Time = 2.0 Seconds)

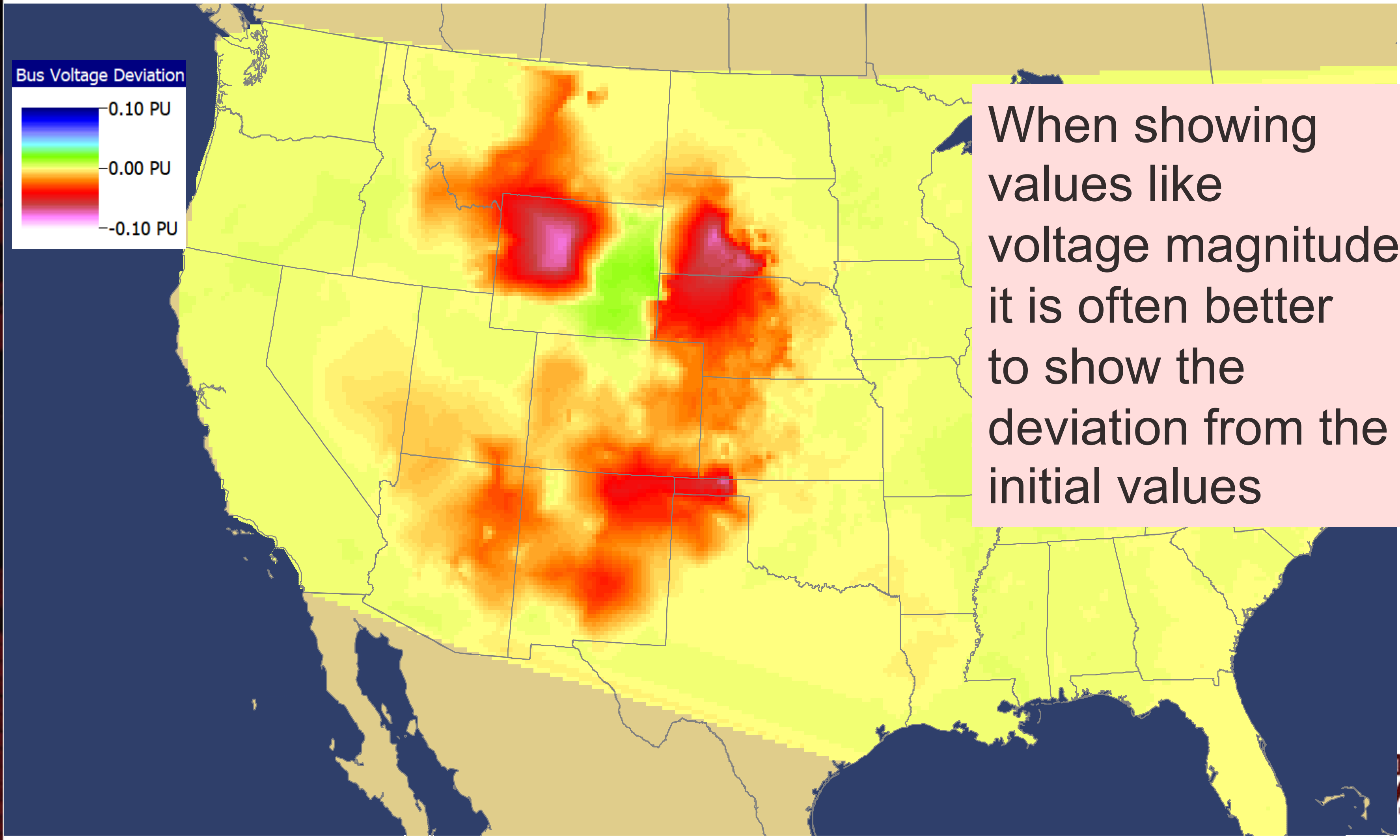


When making movies the we use the Delaunay contours because they are so fast

80K-Bus Frequency Variation (Substation Contour; Time = 8.0 Seconds)



80K-Bus Voltage Magnitude Variation (Substation Contour; Time = 5.0 Seconds)



When showing values like voltage magnitudes it is often better to show the deviation from the initial values

Data Analysis Results Visualization: Example Modal Analysis

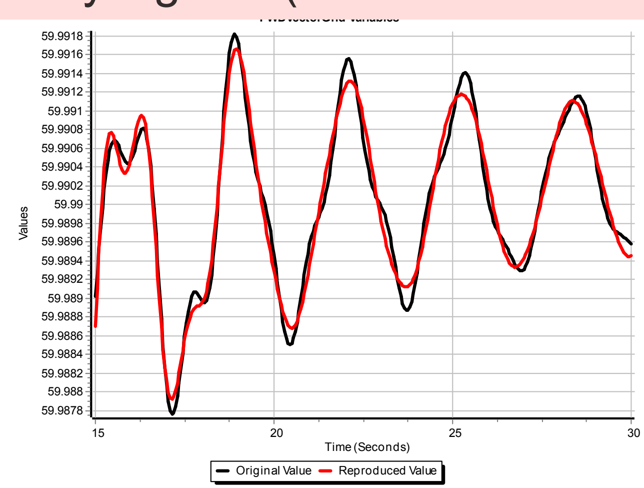
- Idea is to determine the frequency and damping of power system signals after an event
  - Reproduce a signal, such as bus frequency, using exponential functions
- The Iterative Matrix Pencil (IMP) provides a great way to quickly process and verify lots of signals

Number of Complex and Real Modes: 9
Lowest Percent Damping: -13.002
☑ Include Detrend in Reproduced Signals
☐ Subtract Reproduced from Actual
Update Reproduced Signals

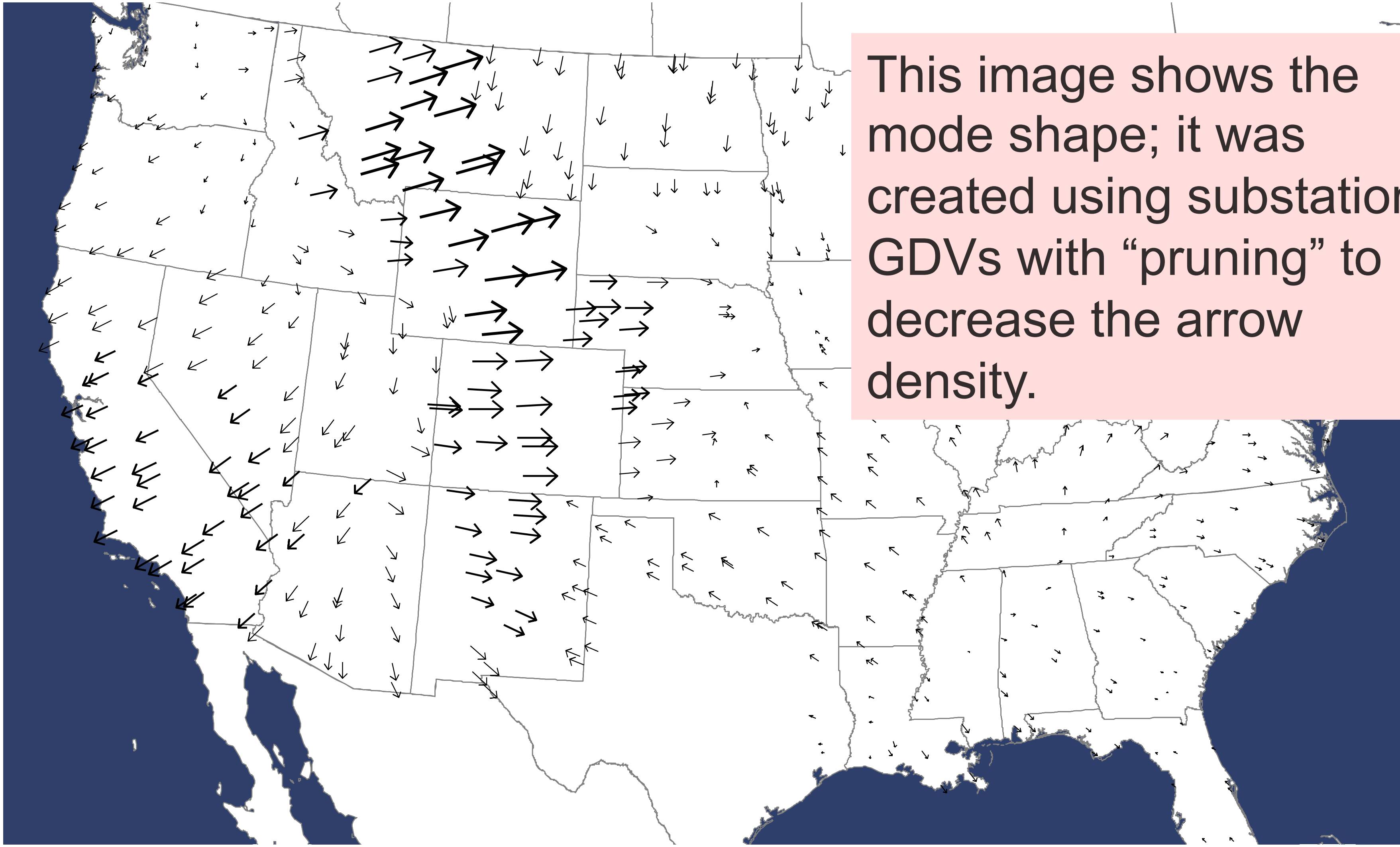
Real and Complex Modes - Editable to Change Initial Guesses

	Frequency (Hz)	Damping (%)	Largest Component in Mode, Unscaled	Name of Signal with Largest Component in Mode, Unscaled	Largest Component in Mode, Scaled	Na wi Co Mo
1	0.851	6.928	0.00108	Frequency Average \ Substation GLASGOW 2	1.185	Fre
2	0.763	2.621	0.00315	Frequency Average \ Substation GLASGOW 2	1.719	Fre
3	0.682	2.572	0.00085	Frequency Average \ Substation GLASGOW 2	0.733	Fre
4	0.615	5.196	0.00310	Frequency Average \ Substation POINT OF ROCKS	3.424	Fre
5	0.494	2.479	0.00289	Frequency Average \ Substation GLASGOW 2	1.981	Fre
6	0.410	9.246	0.00203	Frequency Average \ Substation EL PASO 37	3.131	Fre
7	0.319	2.564	0.00268	Frequency Average \ Substation COLSTRIP 4	2.046	Fre
8	0.000	100.000	0.00073	Frequency Average \ Substation COLSTRIP 4	4.183	Fre
9	0.028	-13.002	0.00030	Frequency Average \ Substation SANDWICH 4	1.520	Fre

The worst match out of 40,000 substation frequency signals (15 to 30 seconds)



Visualization of Lightly Damped 0.5 Hz Mode



This image shows the mode shape; it was created using substation GDVs with “pruning” to decrease the arrow density.

Conclusions

- We've reached the point in which there is too much data to handle most of it directly
  - Certainly the case with much time-varying data
- How data is transformed into actionable information is a crucial, yet often unemphasized, part of the software design process
- There is a need for continued research and development in this area
  - Synthetic power grid cases, including dynamics, are now emerging to provide input for this research

Some Recent Papers…

- T.J. Overbye, J. Wert, A.B. Birchfield, and J.D. Weber, “Wide-Area Electric Grid Visualization Using Pseudo-Geographic Mosaic Displays,” 2019 North American Power Symposium (NAPS), Wichita, KS, USA
- A.B. Birchfield and T.J. Overbye, “Mosaic Packing to Visualize Large-Scale Electric Grid Data,” accepted in *IEEE Open Access Journal of Power and Energy*, Jun. 2020
- A.B. Birchfield, Z. Mao, J. Weber, M. Davis, and T.J. Overbye, “An Interactive, Stand-Alone and Multi-User Power System Simulator for the PMU Time Frame,” in IEEE Texas Power and Energy Conference (TPEC) 2019
- A.B. Birchfield, T.J. Overbye, and K. R. Davis, “Educational Applications of Large Synthetic Power Grids”,” *IEEE Transaction on Power Systems*, vol. 34(1), pp. 765-772, Jan. 2019
- W. Trinh, K.S. Shetye, I. Idehen, and T.J. Overbye, “Iterative Matrix Pencil Method for Power System Modal Analysis,” Proc. 52nd Hawaii International Conference on System Sciences (HICSS), Jan. 2019
- A.B. Birchfield and T.J. Overbye, “Techniques for Drawing Geographic One-line Diagrams: Substation Spacing and Line Routing,” *IEEE Transaction on Power Systems*, vol. 33(6), pp. 7269-7276, Nov. 2018
- I. Idehen, B. Wang, K. Shetye, J. Weber, and T.J. Overbye, “Visualization of Large-Scale Electric Grid Oscillation Modes” in North American Power Symposium (NAPS), Nov. 2018
- A.B. Birchfield, T. Xu, K. Gegner, K.S. Shetye, and T.J. Overbye, “Grid Structural Characteristics as Validation Criteria for Synthetic Networks,” *IEEE Transactions on Power Systems* , vol. 32(4), pp. 3258-3265, Jul. 2017

All are available at overbye.engr.tamu.edu/publications

Thank You!

Questions?