Wide-Area Visualization of Electric Grid Time-Varying Information

Thomas J. Overbye
TEES Eminent Professor
Texas A&M University
Presentation at Purple Mountain Forum
Nanjing China, August 24, 2018

TEXAS A&M UNIVERSITY

Acknowledgments

- Work presented here has been supported by a variety of sources in the US, including PSERC, DOE, ARPA-E, NSF, EPRI, BPA, Illinois Center for a Smarter Electric Grid and PowerWorld. Their support is gratefully acknowledged!
- Slides also include contributions from TAMU and UIUC graduate student and engineers including Komal Shetye, Sudipta Dutta, Saurav Mohapatra, Trevor Hutchins, Adam Birchfield, Ti Xu, Kathleen Gegner, and Iyke Idehen
- Thanks for human factor aspects from Prof. Esa Rantanen, Rochester Institute of Technology

Overview

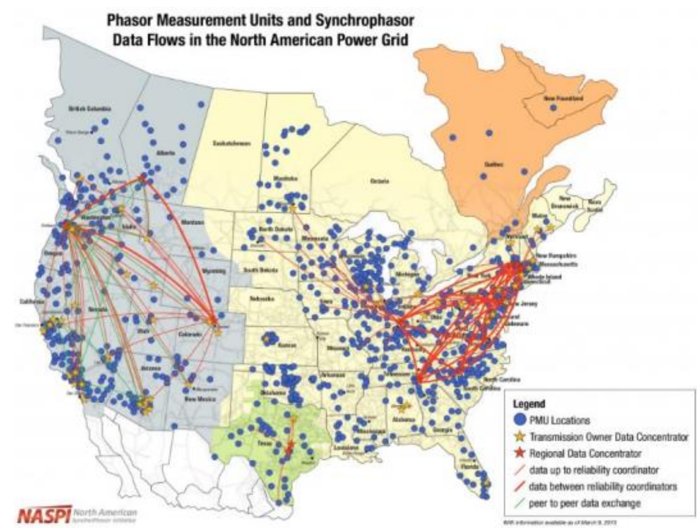
- Our modern society depends on reliable electricity; large blackouts can be catastrophic
- Interconnected electric grids worldwide are in a period of rapid transition. Examples include
  - Integration of large amounts of renewable generation
  - Changing load, including more electric vehicles
  - Customers having more choice in their electric service
  - Inclusion of new technologies for sensing and control, such as phasor measurement units with "big data"
- There are lots of opportunities for innovation, including in the area of visualization

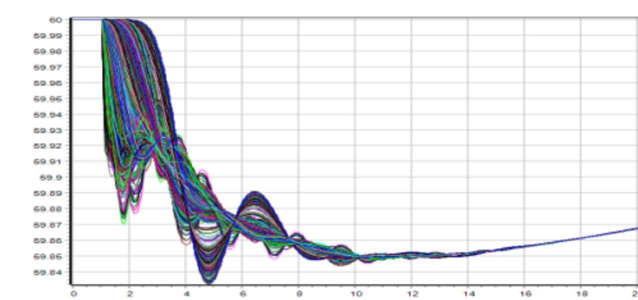
Overview, cont.

- Power system operations and planning are generating more data than ever
  - In operations thousands of PMUs are now deployed
  - In planning many thousand of studies are now routinely run, with a single transient stability run creating millions of values
- How data is transformed into actionable information is a crucial, yet often unemphasized, part of the software design process
- Presentation addresses some issues associated with dealing with this data

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

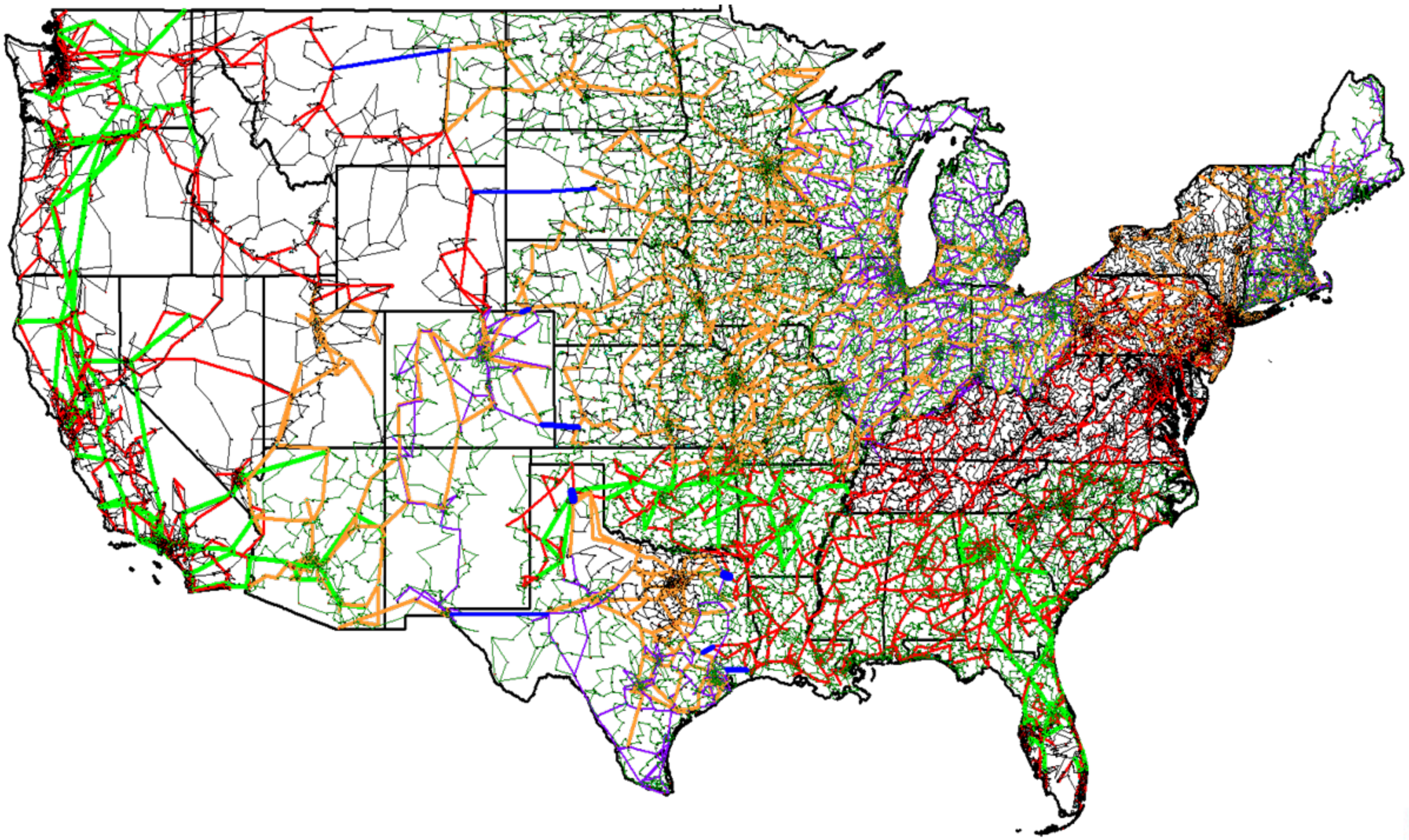




Examples of Power System "Big Data"

- A 100,000 bus grid solved hourly for one year generates 100K times 8760 = 876 million values
- Each hourly simulation may have 10,000 contingencies, giving 8.76 trillion bus values
- Each contingencies could also be run as a time domain simulation, which is sampled at PMU frequency (30 per second) for 30 seconds each gives about 8 quadrillion bus values

Example of 82,000 Bus Synthetic Grid



Example of Grid Dynamic Response

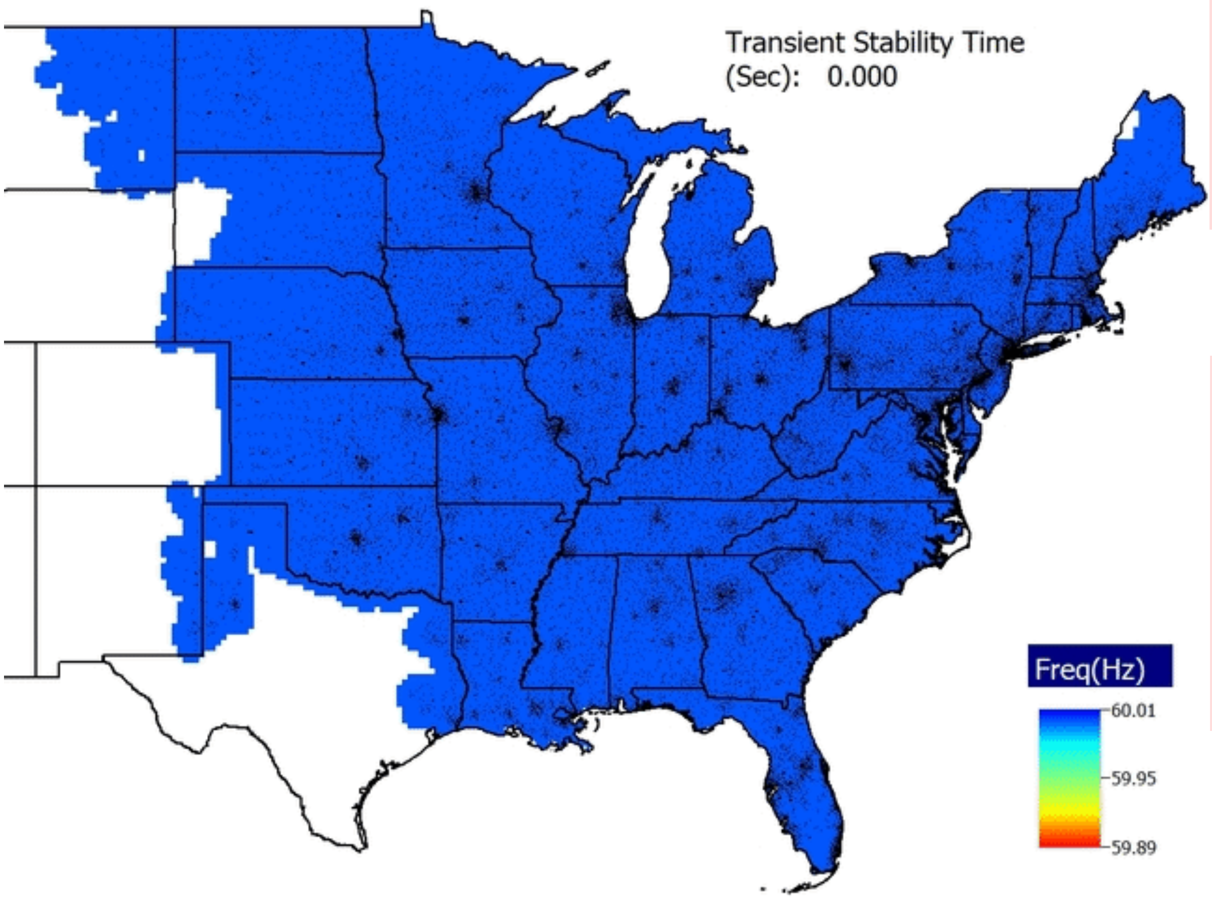


Image animates the frequency variation following a generator outage

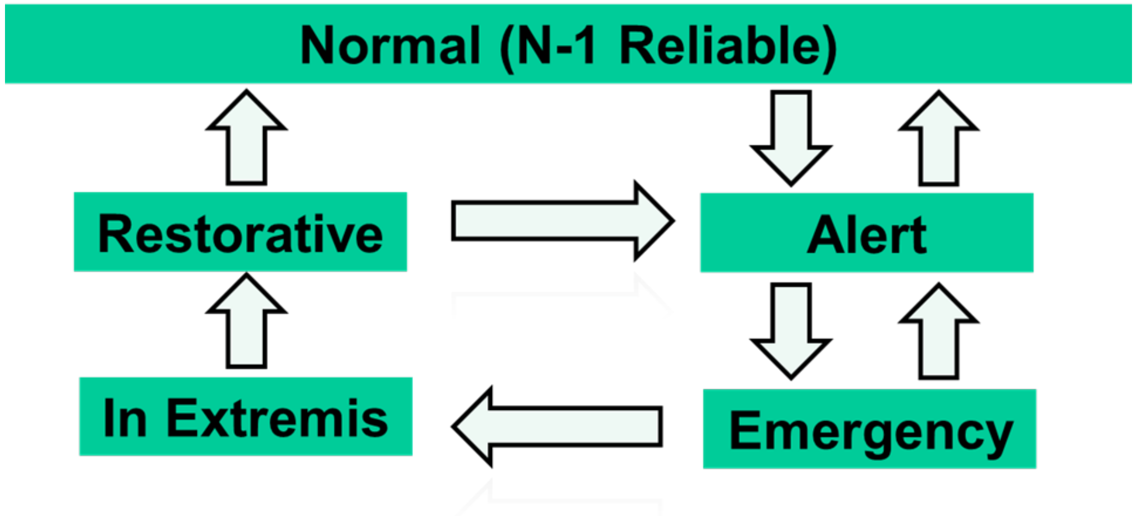
The models and all the data associated with this image is public

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 (like adding up bus flows)

Power System Operating States

- Effective data analysis and visualization for operations requires considering the different operating states



- Effective visualization is most needed for the more rare situations and for planning

Image Derived From L.H. Fink and K. Carlsen, Operating under stress and strain, *IEEE Spectrum*, March 1978, pp. 48-53

Synthetic Models and Visualization

- Access to actual power grid models is often restricted, and this can be a particular concern with data analysis and visualization since its purpose is provide insight into the model, including weaknesses
  - Models cannot be freely shared with other researchers, and even presenting results can be difficult
- Solution is to create entirely synthetic (fictitious) models the mimic characteristics of actual models
  - We are doing this on a US ARPA-E project, with all models containing geographic coordinates

Power System Visualization History: Time Varying Information

When computers were first used in system dispatch centers, they augmented the traditional analog systems. These systems were referred to as digital-directed analog control computers. As the digital computer became more reliable, it assumed full control.

Utility Control Room, 1960’s

PSE&G Control Center in 1988

Left Source: W. Stagg, M. Adibi, M. Laughton, J.E. Van Ness, A.J. Wood, “Thirty Years of Power Industry Computer Applications,” IEEE Computer Applications in Power, April 1994, pp. 43-49

Right Source: J.N. Wrubel, R. Hoffman, “The New Energy Management System at PSE&G,” IEEE Computer Applications in Power, July 1988, pp. 12-15.

Present: PJM Control Center: Electronic Strip-Charts

Image Source: http://tdworld.com/site-files/tdworld.com/files/imagecache/large_img/uploads/2013/07/pjmcontrolroom117.jpg

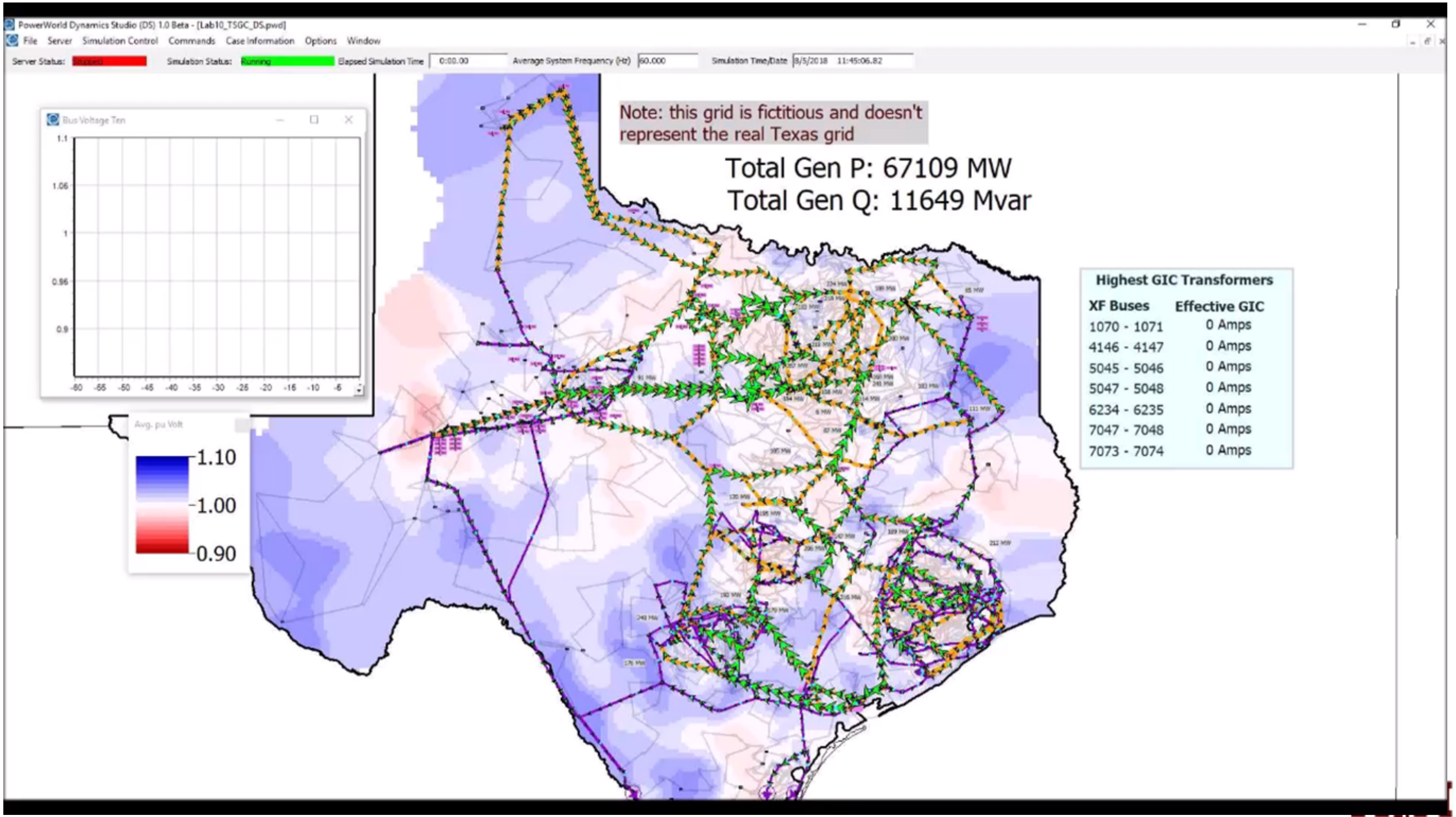
Blackouts and Operator Intervention

- Many large-scale blackouts have time scales of several minutes to a few dozen minutes
  - this time scale allows for operator intervention, but it must occur quickly to be effective (extreme emergency control)
- Operators can't respond effectively if they do not know what is going on– they need "situational awareness"

Extreme Emergency Control

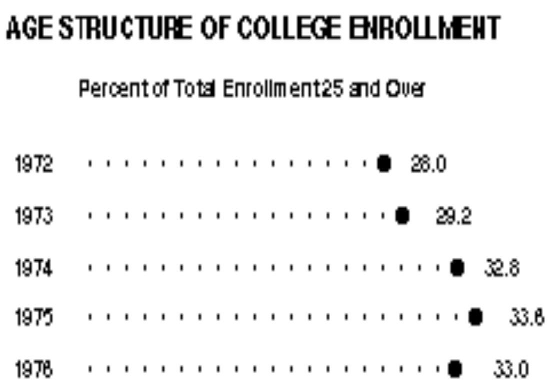
- How the control room environment might be different during such an event
  - advanced network analysis applications could be unavailable or overwhelmed
  - system state could be quite different, with unfamiliar flows and voltages
  - lots of alarms and phone calls
  - high level of stress for control room participants with many tasks requiring their attention
  - large number of decision makers might be present
- Designing software for extreme conditions is challenging since conditions seldom encountered

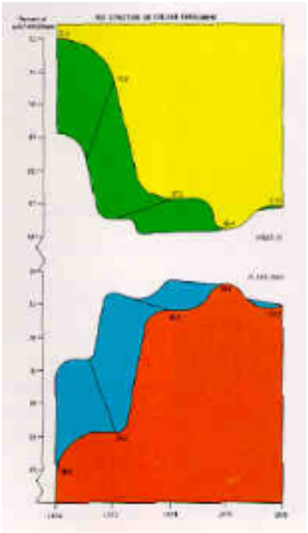
Demonstration of Extreme Situations Using Synthetic Grids



A Visualization Caution!

- 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



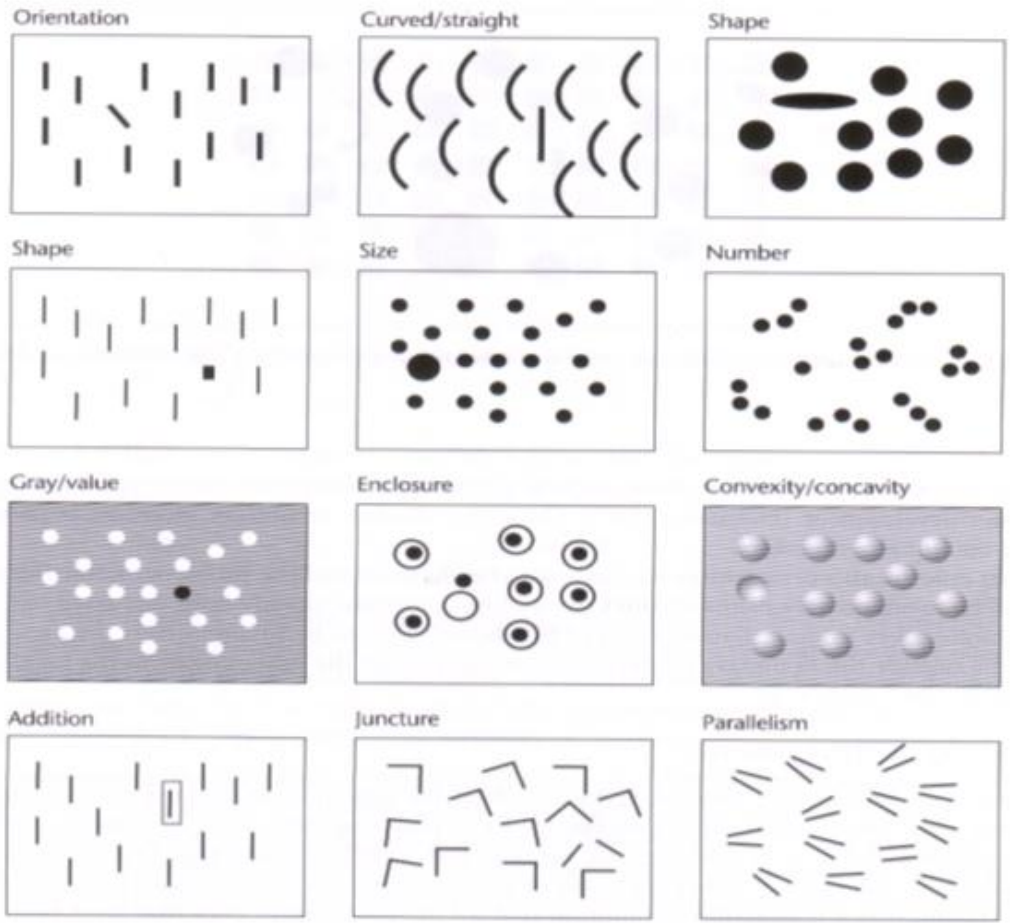


Source: E. Tufte, The Visual Display of Quantitative Information, Graphics Press, Cheshire, CT, 1983.

Visualization Background: Preattentive Processing

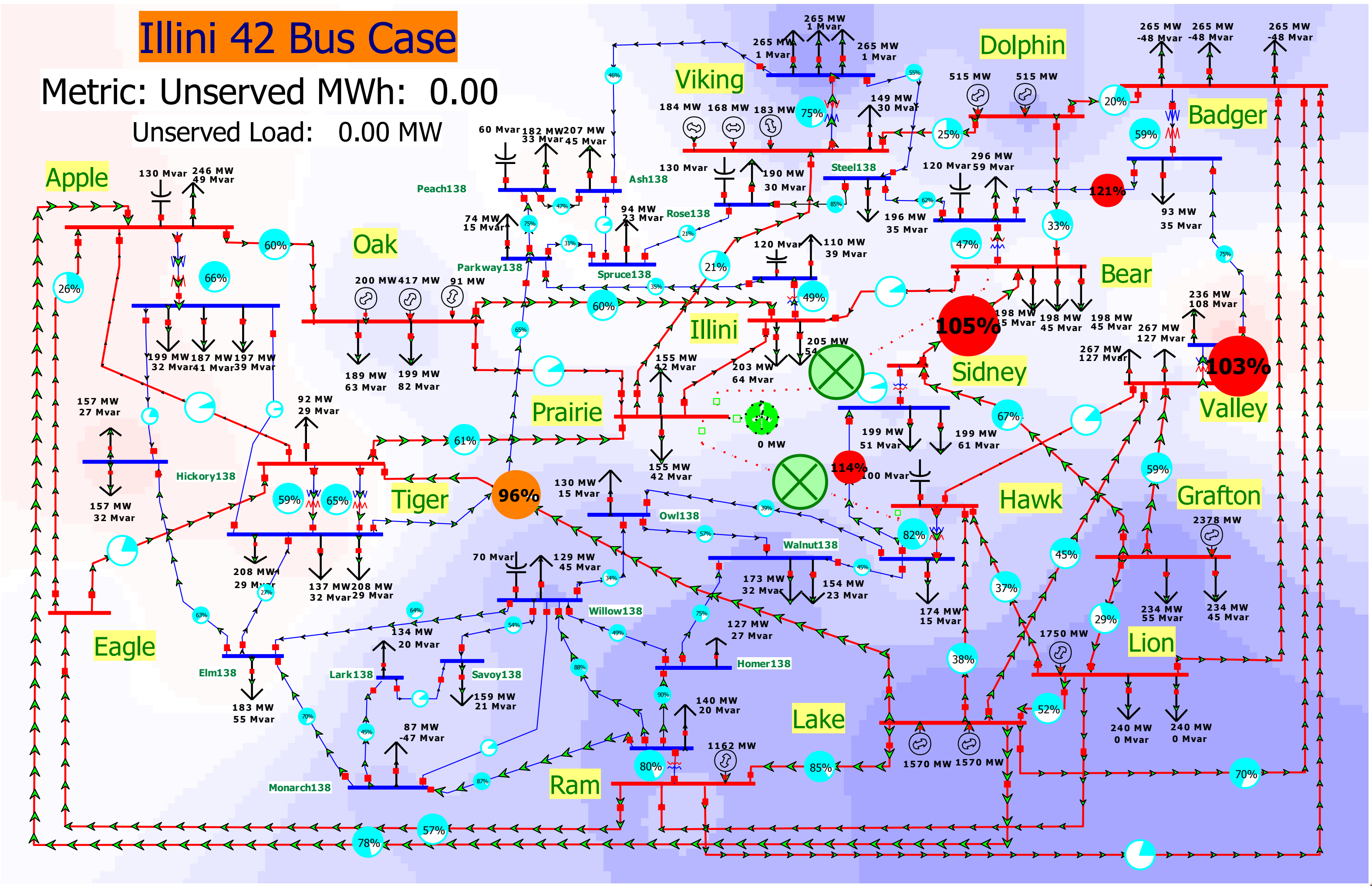
- Good reference book: Colin Ware, *Information Visualization: Perception for Design*, Third Edition, 2013
- When displaying large amounts of data, take advantage of preattentive cognitive processing
  - With preattentive processing the time spent to find a "target" is independent of the number of distractors
- Graphical features that are preattentively processed include the general categories of form, color, motion, spatial position

All are Preattentively Processed Except Juncture and Parallelism



Source: *Information Visualization* by Colin Ware, Fig 5.5

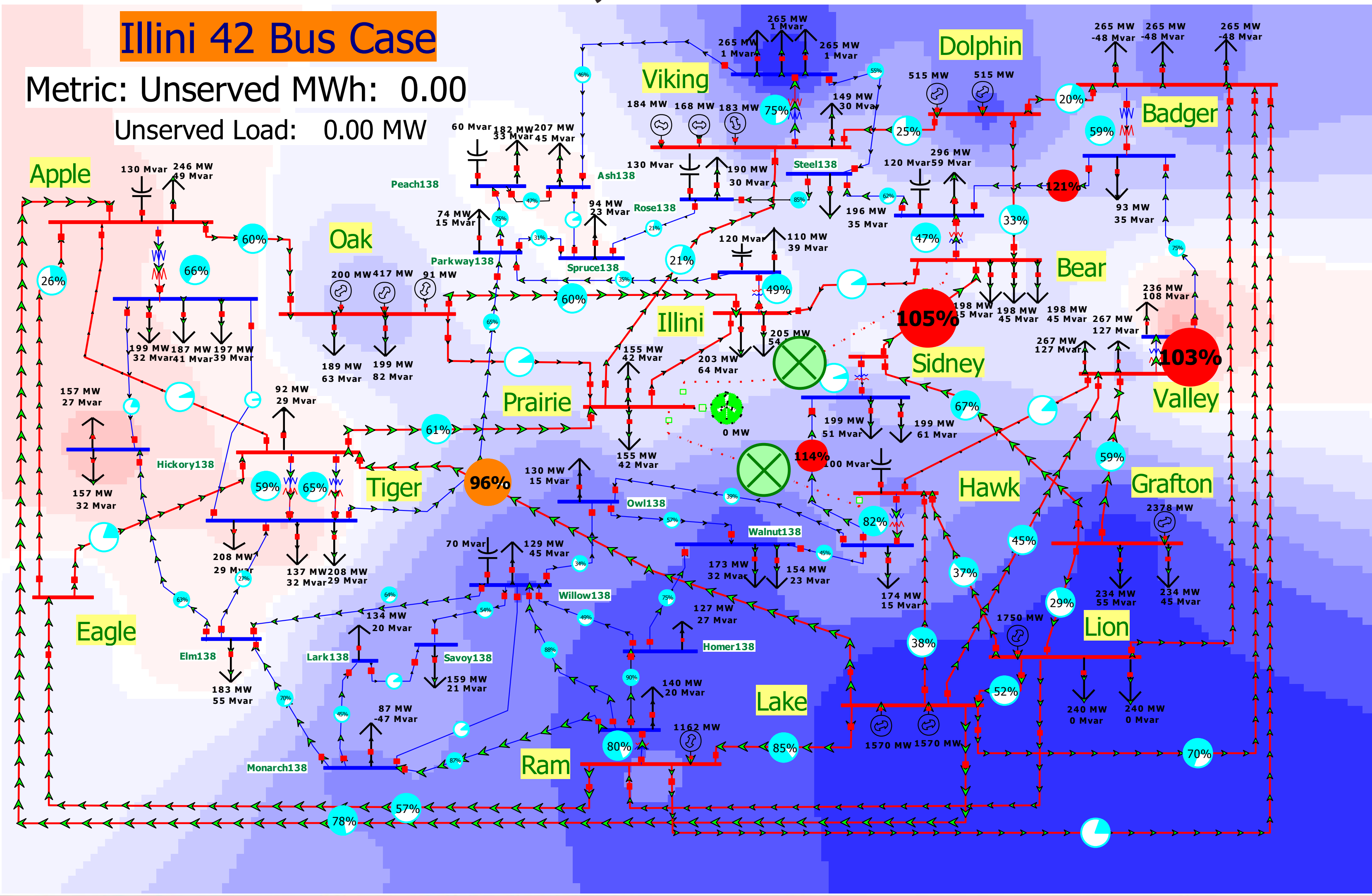
Preattentive Processing with Color & Size



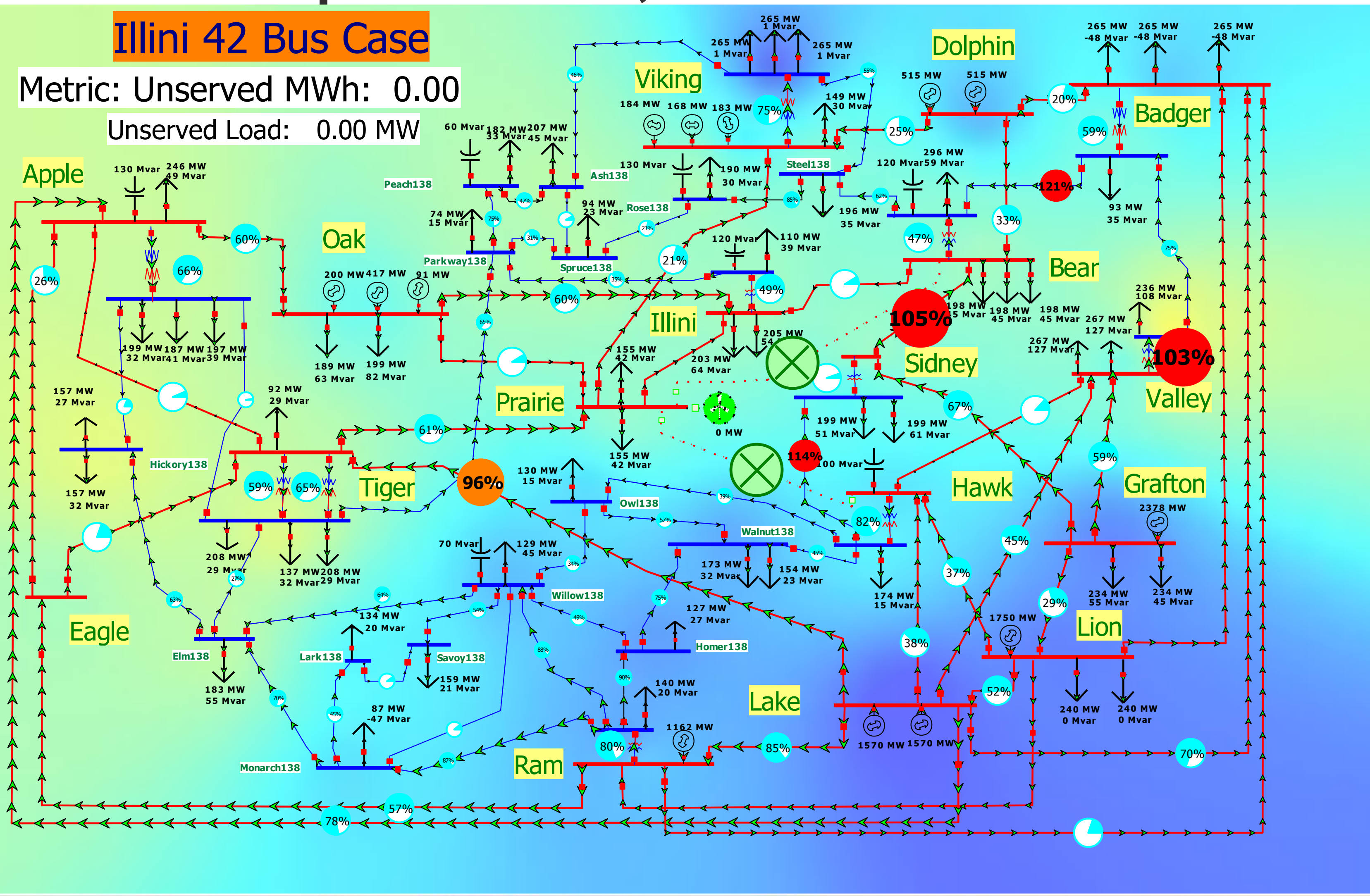
Use of Color

- 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
  - Multi-color scales (like a spectrum) have advantages (more steps) but also disadvantages (effectively comparing values) compared to bi-color sequences

Color Sequence Example: Blue/ Red, Discrete



Color Sequence Example: Spectrum, Continuous



Visual Working Memory and Change Blindness

- The visual working memory (what we retain about images) is limited to a small number of simple objects or patterns, perhaps 3 to 5.
- Because we remember so little, it is possible to make large changes to displays and people will generally not notice unless they are fixated upon it.
- Fast animation (without flicker) can help reduce this.

Change Blindness Example: With Flicker Hard to Detect

http://hps.elte.hu/Basler/Courses/Consciousness/Change%20Blindness/

Large Changes can be Hard to Detect with Local Disruptions

Source: http://nivea.psycho.univ-paris5.fr/ECS/ECS-CB.html

Change Blindness Comments

- Change blindness is most likely under high task load conditions with improbable events
- Less likely to occur when change is more salient (turning on a light is better than turning it off; changing “on” to “off” is not very salient).
- Changes in main field of vision easiest to detect
- Experts in domain are less likely to experience change blindness

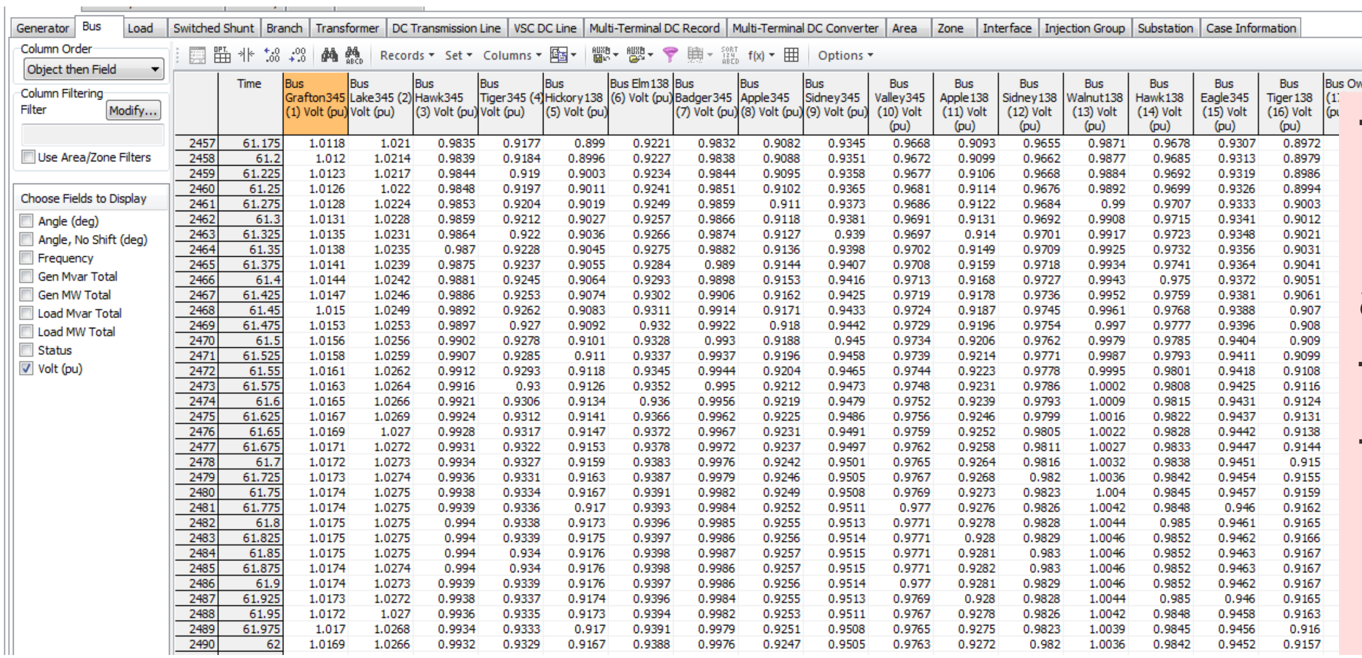
Reference: Wickens, Hollands, Banbury, Parasuraman, *Engineering Psychology and Human Performance*, 4th Edition, 2013

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!

Tabular Displays

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

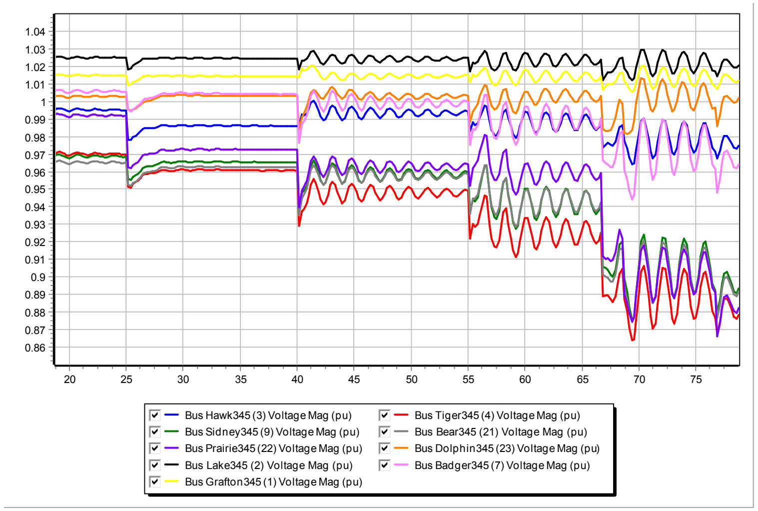


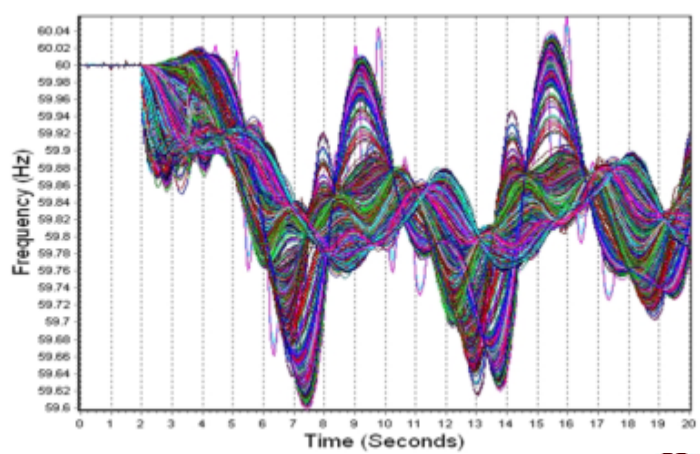
	Time	Bus Grafton345 (1) Volt (pu)	Bus Lake345 (2) Volt (pu)	Bus Hawk345 (3) Volt (pu)	Bus Tiger345 (4) Volt (pu)	Bus Hickory138 (5) Volt (pu)	Bus Elm138 (6) Volt (pu)	Bus Badger345 (7) Volt (pu)	Bus Apple345 (8) Volt (pu)	Bus Sidney345 (9) Volt (pu)	Bus Valley345 (10) Volt (pu)	Bus Apple138 (11) Volt (pu)	Bus Sidney138 (12) Volt (pu)	Bus Walnut138 (13) Volt (pu)	Bus Hawk138 (14) Volt (pu)	Bus Eagle345 (15) Volt (pu)	Bus Tiger138 (16) Volt (pu)	Bus Owl (1 (pu
2457	61.175	1.0118	1.021	0.9835	0.9177	0.899	0.9221	0.9832	0.9082	0.9345	0.9668	0.9093	0.9655	0.9871	0.9678	0.9307	0.8972	
2458	61.2	1.012	1.0214	0.9839	0.9184	0.8996	0.9227	0.9838	0.9088	0.9351	0.9672	0.9099	0.9662	0.9877	0.9685	0.9313	0.8979	
2459	61.225	1.0123	1.0217	0.9844	0.919	0.9003	0.9234	0.9844	0.9095	0.9358	0.9677	0.9106	0.9668	0.9884	0.9692	0.9319	0.8986	
2460	61.25	1.0126	1.022	0.9848	0.9197	0.9011	0.9241	0.9851	0.9102	0.9365	0.9681	0.9114	0.9676	0.9892	0.9699	0.9326	0.8994	
2461	61.275	1.0128	1.0224	0.9853	0.9204	0.9019	0.9249	0.9859	0.911	0.9373	0.9686	0.9122	0.9684	0.99	0.9707	0.9333	0.9003	
2462	61.3	1.0131	1.0228	0.9859	0.9212	0.9027	0.9257	0.9866	0.9118	0.9381	0.9691	0.9131	0.9692	0.9908	0.9715	0.9341	0.9012	
2463	61.325	1.0135	1.0231	0.9864	0.922	0.9036	0.9266	0.9874	0.9127	0.939	0.9697	0.914	0.9701	0.9917	0.9723	0.9348	0.9021	
2464	61.35	1.0138	1.0235	0.987	0.9228	0.9045	0.9275	0.9882	0.9136	0.9398	0.9702	0.9149	0.9709	0.9925	0.9732	0.9356	0.9031	
2465	61.375	1.0141	1.0239	0.9875	0.9237	0.9055	0.9284	0.989	0.9144	0.9407	0.9708	0.9159	0.9718	0.9934	0.9741	0.9364	0.9041	
2466	61.4	1.0144	1.0242	0.9881	0.9245	0.9064	0.9293	0.9898	0.9153	0.9416	0.9713	0.9168	0.9727	0.9943	0.975	0.9372	0.9051	
2467	61.425	1.0147	1.0246	0.9886	0.9253	0.9074	0.9302	0.9906	0.9162	0.9425	0.9719	0.9178	0.9736	0.9952	0.9759	0.9381	0.9061	
2468	61.45	1.015	1.0249	0.9892	0.9262	0.9083	0.9311	0.9914	0.9171	0.9433	0.9724	0.9187	0.9745	0.9961	0.9768	0.9388	0.907	
2469	61.475	1.0153	1.0253	0.9897	0.927	0.9092	0.932	0.9922	0.918	0.9442	0.9729	0.9196	0.9754	0.997	0.9777	0.9396	0.908	
2470	61.5	1.0156	1.0256	0.9902	0.9278	0.9101	0.9328	0.993	0.9188	0.945	0.9734	0.9206	0.9762	0.9979	0.9785	0.9404	0.909	
2471	61.525	1.0158	1.0259	0.9907	0.9285	0.911	0.9337	0.9937	0.9196	0.9458	0.9739	0.9214	0.9771	0.9987	0.9793	0.9411	0.9099	
2472	61.55	1.0161	1.0262	0.9912	0.9293	0.9118	0.9345	0.9944	0.9204	0.9465	0.9744	0.9223	0.9778	0.9995	0.9801	0.9418	0.9108	
2473	61.575	1.0163	1.0264	0.9916	0.93	0.9126	0.9352	0.995	0.9212	0.9473	0.9748	0.9231	0.9786	1.0002	0.9808	0.9425	0.9116	
2474	61.6	1.0165	1.0266	0.9921	0.9306	0.9134	0.936	0.9956	0.9219	0.9479	0.9752	0.9239	0.9793	1.0009	0.9815	0.9431	0.9124	
2475	61.625	1.0167	1.0269	0.9924	0.9312	0.9141	0.9366	0.9962	0.9225	0.9486	0.9756	0.9246	0.9799	1.0016	0.9822	0.9437	0.9131	
2476	61.65	1.0169	1.027	0.9928	0.9317	0.9147	0.9372	0.9967	0.9231	0.9491	0.9759	0.9252	0.9805	1.0022	0.9828	0.9442	0.9138	
2477	61.675	1.0171	1.0272	0.9931	0.9322	0.9153	0.9378	0.9972	0.9237	0.9497	0.9762	0.9258	0.9811	1.0027	0.9833	0.9447	0.9144	
2478	61.7	1.0172	1.0273	0.9934	0.9327	0.9159	0.9383	0.9976	0.9242	0.9501	0.9765	0.9264	0.9816	1.0032	0.9838	0.9451	0.915	
2479	61.725	1.0173	1.0274	0.9936	0.9331	0.9163	0.9387	0.9979	0.9246	0.9505	0.9767	0.9268	0.982	1.0036	0.9842	0.9454	0.9155	
2480	61.75	1.0174	1.0275	0.9938	0.9334	0.9167	0.9391	0.9982	0.9249	0.9508	0.9769	0.9273	0.9823	1.004	0.9845	0.9457	0.9159	
2481	61.775	1.0174	1.0275	0.9939	0.9336	0.917	0.9393	0.9984	0.9252	0.9511	0.977	0.9276	0.9826	1.0042	0.9848	0.946	0.9162	
2482	61.8	1.0175	1.0275	0.994	0.9338	0.9173	0.9396	0.9985	0.9255	0.9513	0.9771	0.9278	0.9828	1.0044	0.985	0.9461	0.9165	
2483	61.825	1.0175	1.0275	0.994	0.9339	0.9175	0.9397	0.9986	0.9256	0.9514	0.9771	0.928	0.9829	1.0046	0.9852	0.9462	0.9166	
2484	61.85	1.0175	1.0275	0.994	0.934	0.9176	0.9398	0.9987	0.9257	0.9515	0.9771	0.9281	0.983	1.0046	0.9852	0.9463	0.9167	
2485	61.875	1.0174	1.0274	0.994	0.934	0.9176	0.9398	0.9986	0.9257	0.9515	0.9771	0.9282	0.983	1.0046	0.9852	0.9463	0.9167	
2486	61.9	1.0174	1.0273	0.9939	0.9339	0.9176	0.9397	0.9986	0.9256	0.9514	0.977	0.9281	0.9829	1.0046	0.9852	0.9462	0.9167	
2487	61.925	1.0173	1.0272	0.9938	0.9337	0.9174	0.9396	0.9984	0.9255	0.9513	0.9769	0.928	0.9828	1.0044	0.985	0.946	0.9165	
2488	61.95	1.0172	1.027	0.9936	0.9335	0.9173	0.9394	0.9982	0.9253	0.9511	0.9767	0.9278	0.9826	1.0042	0.9848	0.9458	0.9163	
2489	61.975	1.017	1.0268	0.9934	0.9333	0.917	0.9391	0.9979	0.9251	0.9508	0.9765	0.9275	0.9823	1.0039	0.9845	0.9456	0.916	
2490	62	1.0169	1.0266	0.9932	0.9329	0.9167	0.9388	0.9976	0.9247	0.9505	0.9763	0.9272	0.982	1.0036	0.9842	0.9452	0.9157	

The information access cost for the task at hand is key!

Time-based graphs

- Graphs can be quite helpful for showing exact values if no more than about ten individual signals are shown
  - In larger sets outliers may be missed
- Showing more values can be helpful in identifying response envelope
  - Graph at bottom left shows 2400 signals





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

Animation Loops: SCADA vs. PMUs

- A potential visualization change is how much future displays are visualized at PMU rates (30 times per second) versus SCADA rates (every 4-12 seconds)

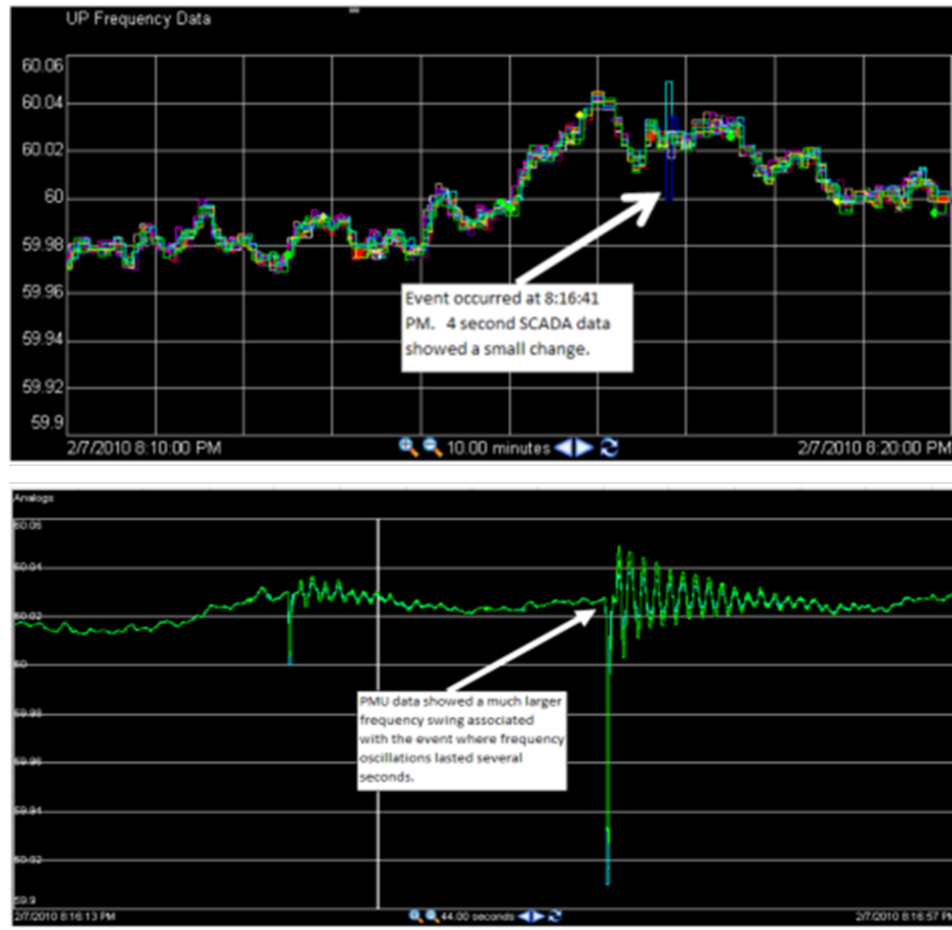


Image Source: Jay Giri (Alstom Grid), "Control Center EMS Solutions for the Grid of the Future," EPCC, June 2013

Data Analysis Algorithms

- Usually there is too much data to make sense of without some type of analysis
- Several terms are used to denote the idea of discovering insight from data:
  - Statistics, data mining, knowledge discovery, data analytics, machine learning, and big data
- Large field, so I'll just present a few examples

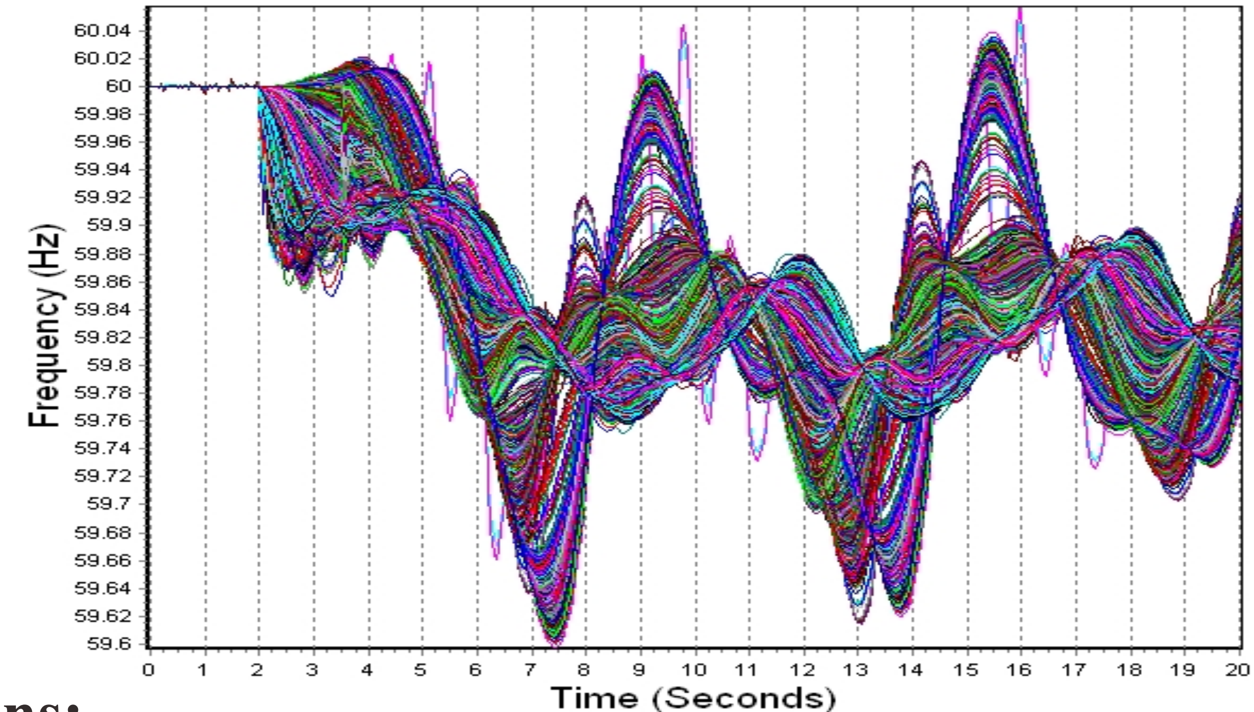
Figure 1.2 The visualization process.

Image source: Colin Ware, *Information Visualization*, Third Edition, 2013

Clustering Example: Transient Stability, PMU, or SCADA Analysis

- A single transient stability solution can generate large amounts of output data
- In real-time a similar situation occurs with PMU data, or on a longer time frame with SCADA
- How much this data needs to be considered is application dependent
  - In operations the concern may just be OK or Not OK
  - In planning more detailed analysis may be required. Issue is how to determine if the results are "correct"?
- Clustering is an example of unsupervised machine learning to make data manageable

Frequency Graph of Data for 2400 Generators



Questions:

- Is the system responding as expected?
- How to separate out the patterns?
- How to incorporate geographic information in the visualization?

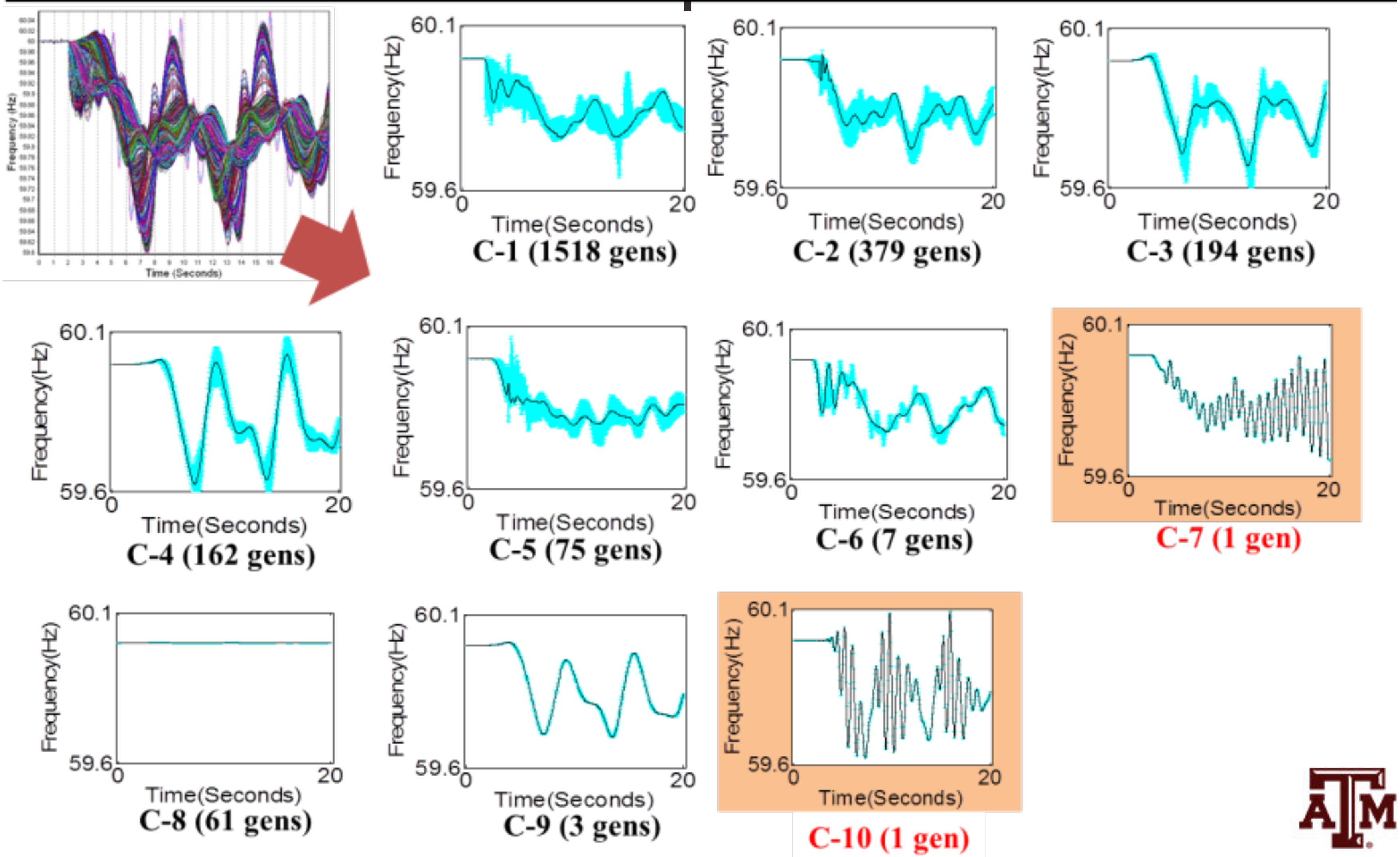
Solution: Apply Data Mining Clustering Techniques

- Clustering is the process of grouping a set of objects so similar objects are together, and dissimilar objects are not together
- There is no perfect clustering method or even a single definition for what constitutes a cluster

Clustering Algorithms

- There are a variety of clustering algorithms. Two common algorithms are
  - K-Means
    - The number of clusters must be specified
    - Very fast and simple in practice
    - Different initial clusters may lead to different results
  - QT: Quality Threshold
    - Form an unknown number of potentially large clusters that meet a “quality standard” which is a specified threshold cluster diameter
    - Requires more computation

Clustering Applied to Results, Ten Distinct Responses Identified

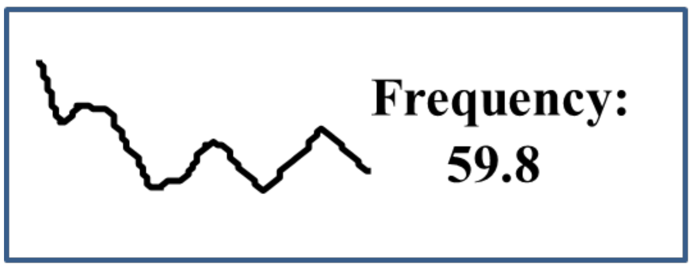


Results Combined with Visualization with Spark-Lines

- Spark-lines (from E. Tufte, *Beautiful Evidence*, 2006) are “intense, word-sized graphics”

Frequency:
59.8

Display of value only



Display of value and trend

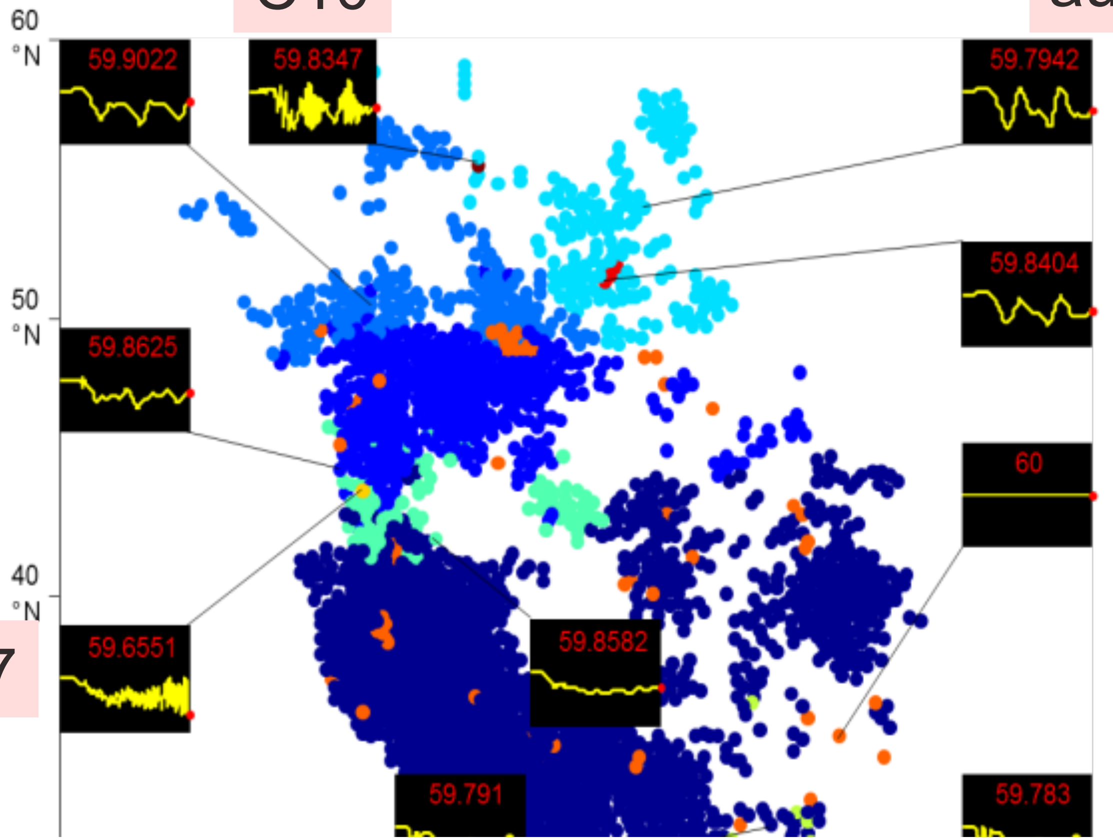
Display of value and trend highlighting present value

Spark-line plot

2400 Generator Results Visualized in a Geographic Context

Outliers are detected automatically

C10

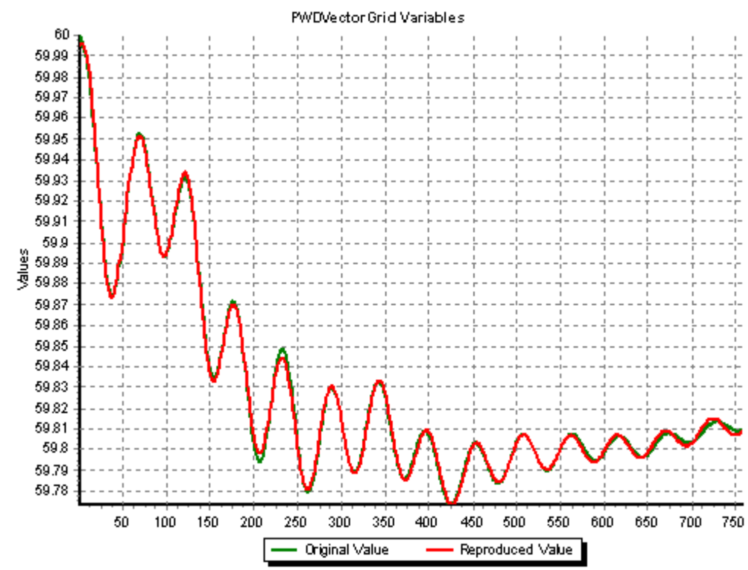


C7

- 10 distinct frequency responses identified
- Visualized on actual geographic location with “spark-lines”
- Different color dots = generators of a cluster

Signal Based Ringdown Modal Decomposition

- 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
- A number of different techniques have been proposed to do this for power systems, starting with Prony analysis in the late 1980's

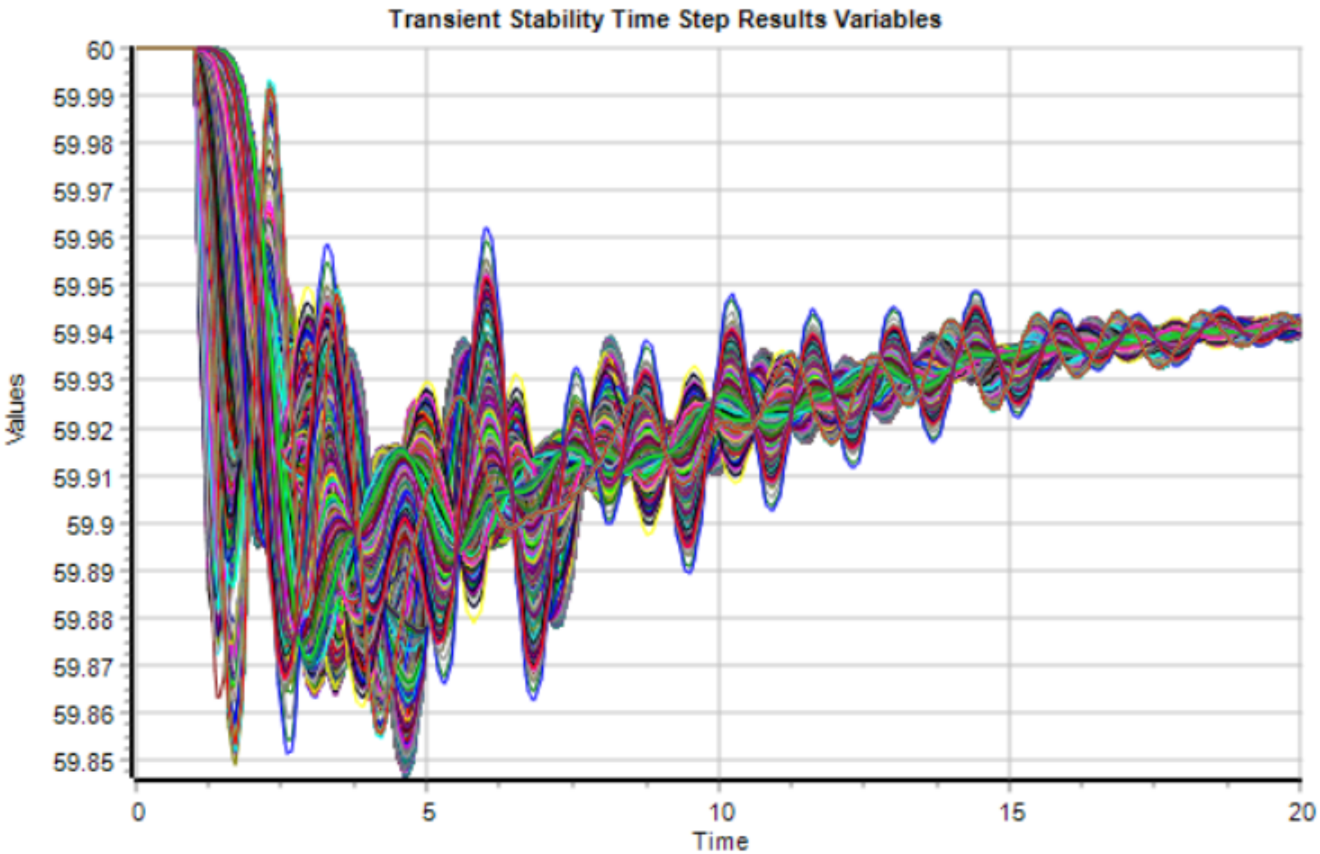


Actual Input | Sampled Input | Fast Fourier Transform Results | Modal Results | Original and Reproduce

	Damping (%)	Frequency (Hz)	Magnitude Scaled by SD	Magnitude, Unscaled	Angle (Deg)
1	2.822	0.766	1.481	0.052	-58.06
2	3.865	0.691	0.368	0.013	150.29
3	11.348	0.325	0.715	0.025	109.78
4	-15.196	0.032	0.898	0.031	116.53
5	100.000	0.000	4.203	0.147	0.00
6	15.546	0.203	1.189	0.041	144.79

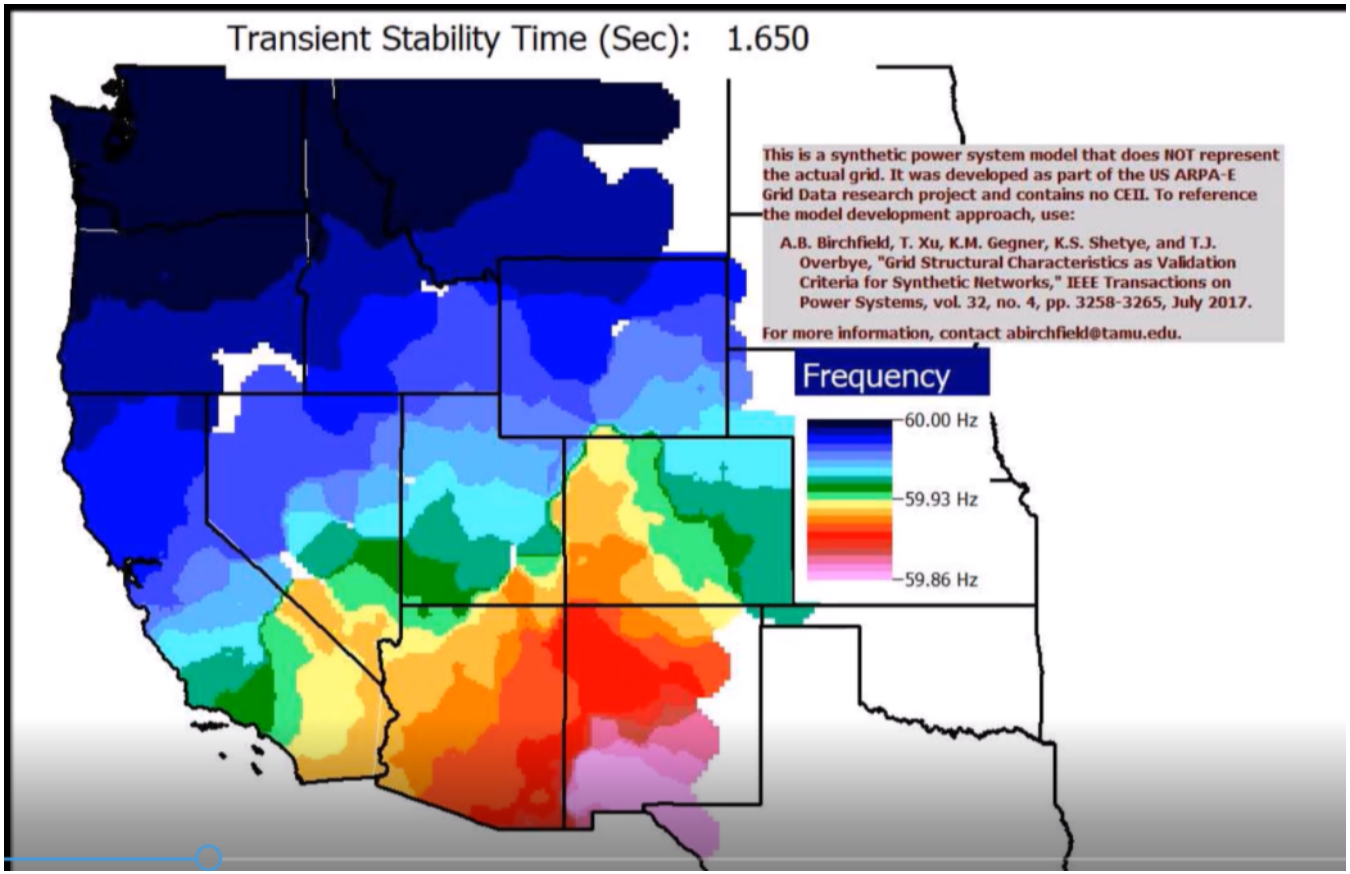
10K Example Transient Stability Results: Generator Outage

Substation Frequency Results for All Substations



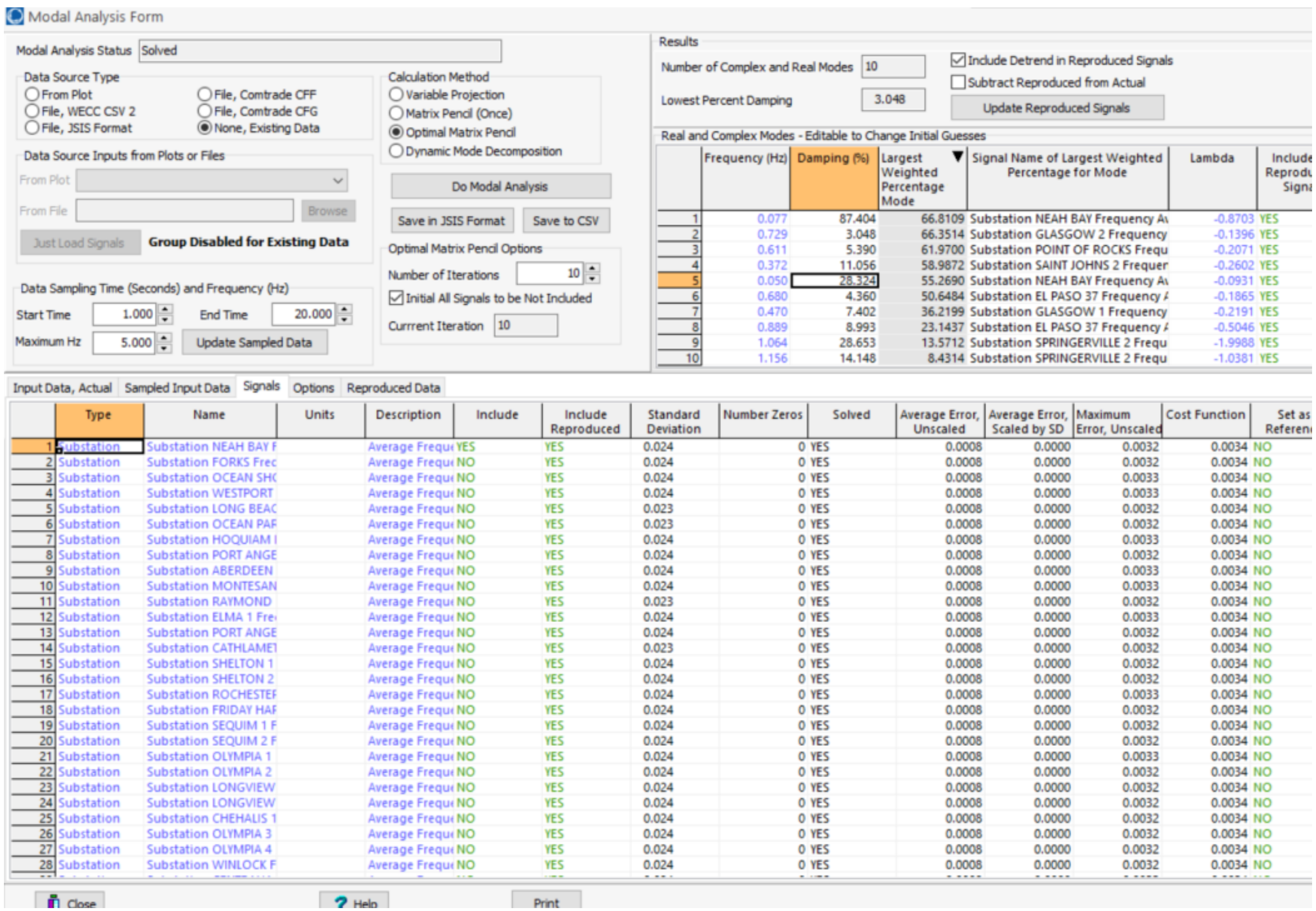
10K Example Transient Stability Results: Generator Outage

Movie Shows Spatial Variation in Frequency

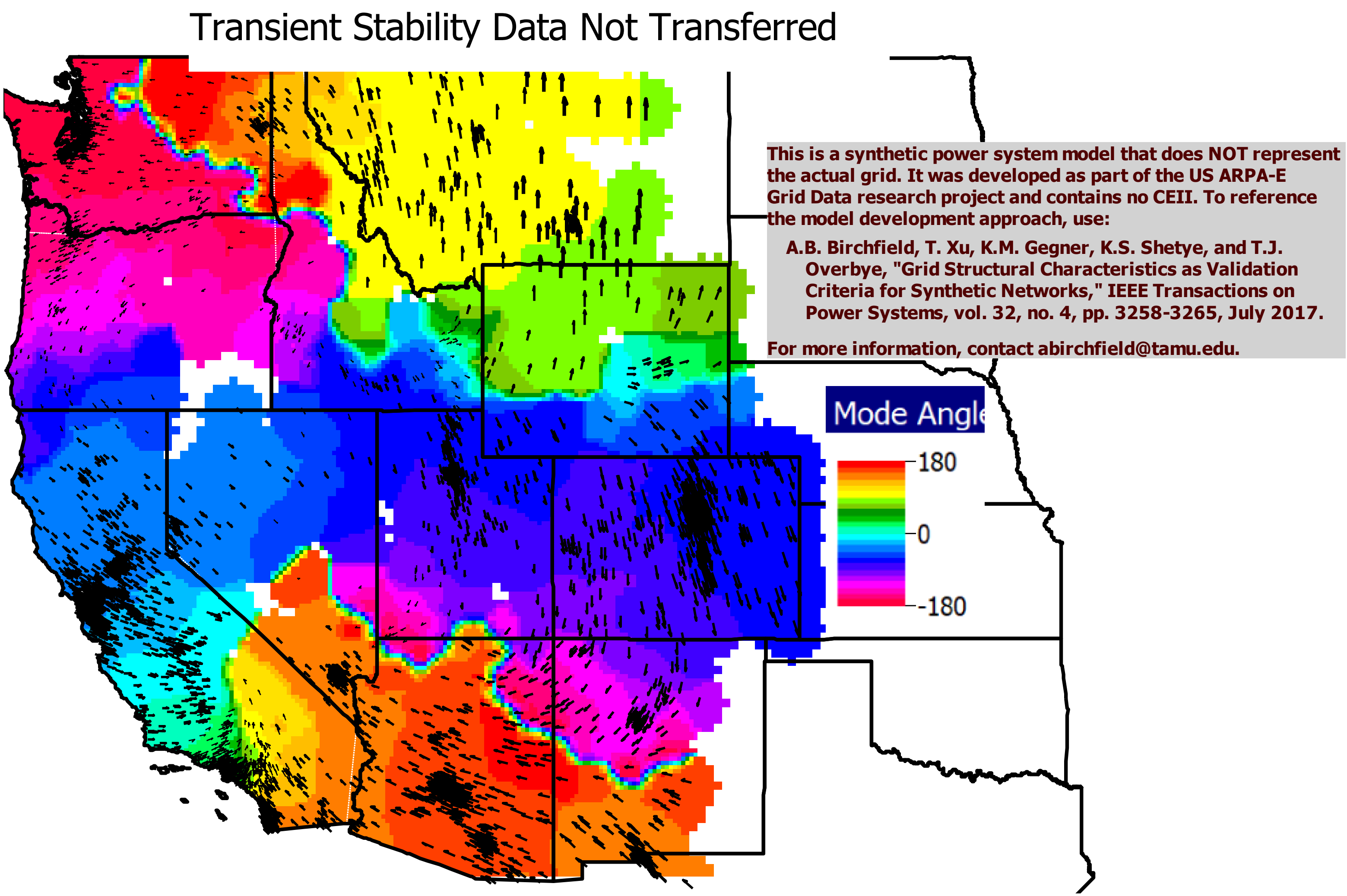


10K Example Transient Stability Modal Analysis of Results

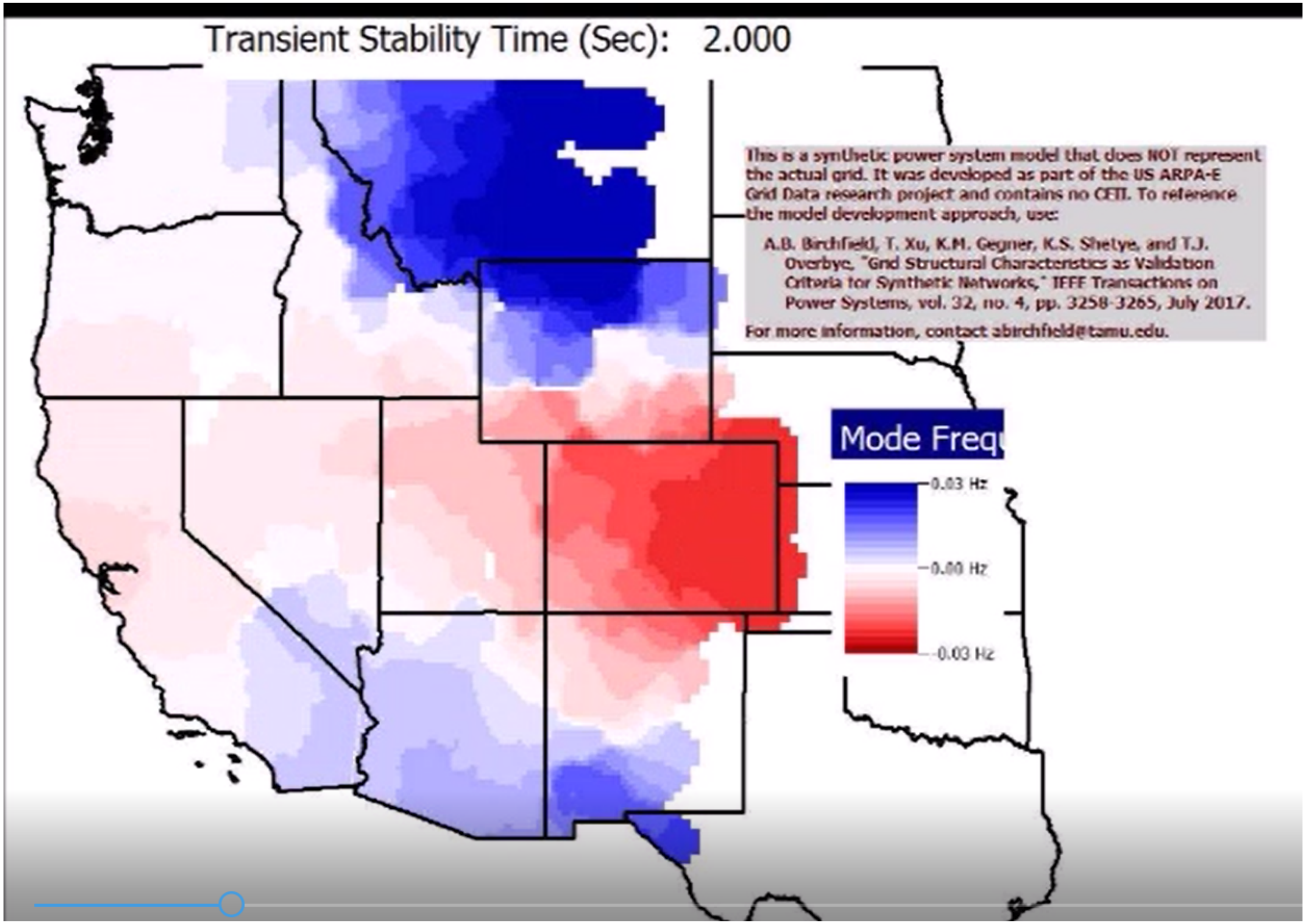
Results Can Be Grouped into Associated Modes



Visualization of the Lightly Damped 0.73 Hz Mode



Visualization of the Lightly Damped 0.73 Hz Mode



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

Thank You!

Questions?

TEXAS A&M UNIVERSITY