12-2012

Data Mining and Machine Learning Applications of Wide-Area Measurement Data in Electric Power Systems

Penn Norris Markham
markham@utk.edu

To the Graduate Council:

I am submitting herewith a dissertation written by Penn Norris Markham entitled "Data Mining and Machine Learning Applications of Wide-Area Measurement Data in Electric Power Systems." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Electrical Engineering.

Yilu Liu, Major Professor

We have read this dissertation and recommend its acceptance:

Lynne E. Parker, Leon M. Tolbert, Joshua S. Fu

Accepted for the Council:

Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

Data Mining and Machine Learning Applications of Wide-Area Measurement Data in Electric Power Systems

A Dissertation Presented for
the Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Penn Norris Markham
December 2012

This dissertation is dedicated in loving memory of Lauren M. Armistead, M.D.

Acknowledgements

This work would not have been possible without the support, encouragement, and assistance of many different people.

First, I would like to express my gratitude and appreciation to my advisor, Dr. Yilu Liu, for guiding me throughout my Ph.D. studies and funding this research. She has not only been an excellent mentor, but a great friend as well. In addition, I wish to thank the members of my committee, Dr. Lynne Parker, Dr. Leon Tolbert, and Dr. Joshua Fu for their advice, comments, and suggestions.

Special thanks go to Mr. Robert D. Bottoms and the Tennessee Valley Authority, who provided power system models and technical guidance for the Eastern Interconnection study in Chapter 7. In addition, Mr. John Stovall at Oak Ridge National Laboratory mentored me for two summers as part of the Grid Innovation Leaders Fellowship. His knowledge and experience were vital in planning and executing the research for the EI study.

There isn't enough space here to list all of my wonderful colleagues in the Power Information Technology Lab and the various ways they have helped me in this endeavor. I am so thankful for their assistance and friendship throughout the past four and a half years. I have enjoyed every minute of this adventure, mostly because of them. Thanks for putting up with me, and know that I will miss you all very much.

Two staff members in the EECS department were particularly helpful during my time here at the University of Tennessee, and they deserve special mention. Ms. Dana Bryson made sure that I stayed on track to graduate by always being able to answer my questions about policies and procedures. She listened to me kvetch more times than I can count, and I am deeply grateful for her friendship and support. I would be remiss if I didn't thank Mr. Markus Iturriaga for expertly handling so many of the information technology-related functions for our group, which allowed me to focus instead on my own work. Not only is Markus an excellent IT administrator, he's also a great friend.

Finally, I would like to thank my family for their love and support.

Abstract

Wide-area measurement systems (WAMS) are quickly becoming an important part of modern power system operation. By utilizing the Global Positioning System, WAMS offer highly accurate time-synchronized measurements that can reveal previously unobtainable insights into the grid's status. An example WAMS is the Frequency Monitoring Network (FNET), which utilizes a large number of Internet-connected low-cost Frequency Disturbance Recorders (FDRs) that are installed at the distribution level.

The large amounts of data collected by FNET and other WAMS present unique opportunities for data mining and machine learning applications, yet these techniques have only recently been applied in this domain. The research presented here explores some additional applications that may prove useful once WAMS are fully integrated into the power system. Chapter 1 provides a brief overview of the FNET system that supplies the data used for this research. Chapter 2 reviews recent research efforts in the application of data mining and machine learning techniques to wide-area measurement data. In Chapter 3, patterns in frequency extrema in the Eastern and Western Interconnections are explored using cluster analysis. In Chapter 4, an artificial neural network (ANN)-based classifier is presented that can reliably distinguish between different types of power system disturbances based solely on their frequency signatures. Chapter 5 presents a technique for constructing electromechanical transient speed maps for large power systems using FNET data from previously detected events. Chapter 6 describes an object-oriented software framework useful for developing FNET data analysis applications.

In the United States, recent environmental regulations will likely result in the removal of nearly 30 GW of oil and coal-fired generation from the grid, mostly in the Eastern Interconnection (EI). The effects of this transition on voltage stability and transmission line flows have previously not been studied from a system-wide perspective. Chapter 7 discusses the results of power flow studies designed to simulate the evolution of the EI over the next few years as traditional generation sources are replaced with greener ones such as natural gas and wind.

Conclusions, a summary of the main contributions of this work, and a discussion of possible future research topics are given in Chapter 8.

Table of Contents

List of Figures

List of Attachments

1. Frequency Monitoring Network (FNET)

Structure

Originally developed at Virginia Tech in 2003, the Frequency Monitoring Network (FNET) is a wide-area phasor measurement system that collects power system data using low-cost, high-accuracy Frequency Disturbance Recorders (FDRs). The FDR (Fig. 1.1) can be thought of as a single phase phasor measurement unit (PMU) that is designed to take voltage waveform inputs at ordinary 120-V electrical outlets, rather than at the much higher transmission-level voltages used by traditional PMUs. Since the FDR uses distribution level voltages, it can be installed virtually anywhere. In fact, many FDRs are installed in private residences, offices, and schools. A map of the current FDR locations is shown in Fig. 1.2.

Fig. 1.1: Frequency Disturbance Recorder

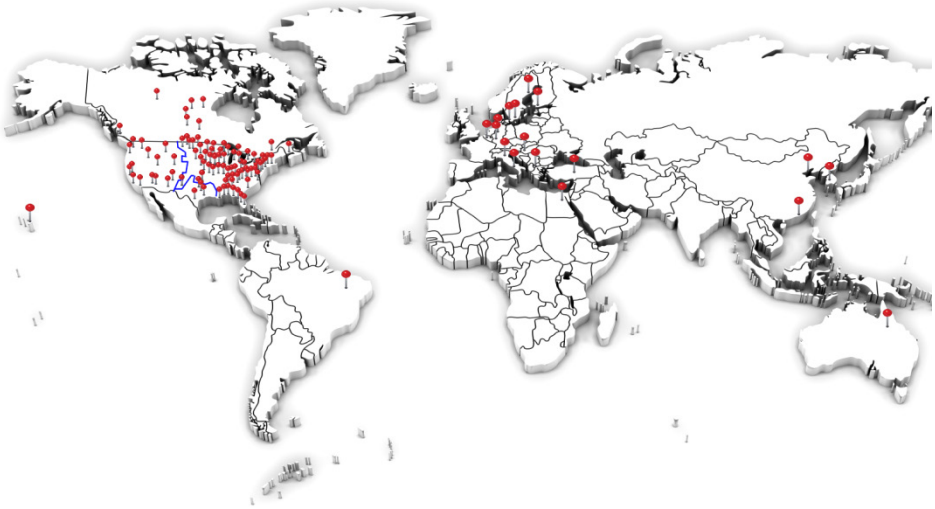


Fig. 1.2: Locations of currently installed FDRs as of March 2012

By making use of the precise timing information provided by the Global Positioning System, the FDR is able to compute the frequency and absolute phase angle of the voltage signal very accurately at 100-ms intervals. These measurements are then timestamped and transmitted to a phasor data

concentrator (PDC) at the University of Tennessee, Knoxville, where they are recorded and archived (Fig. 1.3). Since the FNET system went online in 2004, more than 6 TB of data have been collected from FDRs located within the United States and around the world. The data are used for a variety of applications, including event detection and location, oscillation detection, visualization, and forensic authentication of digital evidence [1-8]. A thorough description of the FNET system can be found in [9-12]. A list of FDRs currently deployed in the Eastern Interconnection can be found in Appendix A.

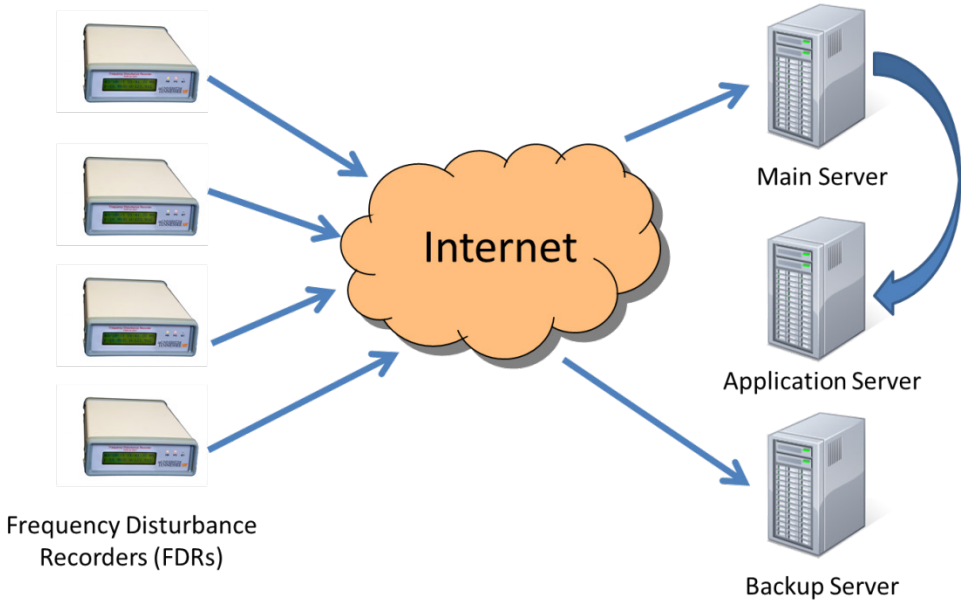


Fig. 1.3: FNET system architecture.

2. Literature Review

Data Mining and Machine Learning Applications of WAMS Data

Since the development of the first Supervisory Control and Data Acquisition (SCADA)-based wide-area measurement systems (WAMS) in the 1960's, system operators have sought to automate the process of monitoring and controlling the power system. When computers were very expensive, this was done in a centralized fashion, with measurements from transducers being sent to a single SCADA master, which then made decisions based upon that information and communicated control actions to devices in the field. The decision algorithms were usually relatively simple. For example, if a bus voltage was found to be low, the system would correct it by adjusting transformer taps or switching the appropriate capacitor banks into service. As microprocessors have dropped in cost, some of this control functionality has devolved to the remote devices themselves, eliminating much of the reliance on a single computer [13]. Still, SCADA systems operate over relatively long time intervals; measurements are collected every few seconds, and only then can control actions be determined and executed. This time frame is sufficient for correcting a variety of system issues, but is far too long for others.

Phasor measurement units (PMUs) capable of measuring and calculating a large number of quantities at sub-second intervals represent the latest evolution of power system instrumentation technology. Although PMUs have existed since the 1980's, only recently have they been installed throughout the power system in significant numbers. Enabled by advances in computing power and storage technology, the resulting explosion of data has quickly become overwhelming as utilities and researchers struggle with how to efficiently transmit, use, and store these measurements. The problem has often been likened by those in the industry to drinking out of a fire hose, particularly since the reporting rates of modern PMUs exceed 30 records per second, compared with one record every few seconds in SCADA systems. Data mining and machine learning techniques have been developed to deal with large quantities of high-dimensional data, and some of these have been applied to wide-area measurements collected from electric power systems [14]. This section provides an overview of recent developments in this area, which can generally be divided into two categories: classification of system phenomena and stability assessment.

Classification of System Phenomena

Research into automated classification of power system events began well before phasor measurement units were in widespread use. Before that time, Intelligent Electronic Devices (IEDs) such as Digital Fault Recorders (DFRs) and microprocessor-based digital relays were essentially the only

means to record and analyze time-domain measurements obtained from substation transducers. Much of the early work done in this area focused on methods of detection and classification of faults that didn't require performing extensive pre-processing or the explicit calculation of phasor quantities or symmetrical components. Although not explicitly labeled as such, Kezunovic *et al.* utilized a type of decision tree to determine the type of fault (e.g., phase-to-phase, phase-to-ground, three-phase, etc.) using actual DFR data, although it was not trained in the classical sense [15]. Later, he used neural networks to perform the same function on simulated current and voltage signals, achieving greater than 90% accuracy [16]. A few years after that, Poeltl and Fröhlich computed phasor quantities from simulated measurements that were then fed into a neural network to classify the type of fault. Their technique was notable because it could perform the phasor computation and classification within five milliseconds, less than a quarter of a cycle on 60-Hz power systems [17]. Most recently, the Tennessee Valley Authority used a nearest-neighbor approach to find different types of power system disturbances in a large (25 TB) database of PMU data. Due to the size of the dataset, the algorithm was implemented using the MapReduce programming model and Apache Hadoop Distributed File System, which are popular tools for this type of analysis [18].

Stability Assessment

The literature contains several examples of machine learning techniques being applied to synchrophasor data to create to assess both voltage and transient stability. In [19, 20], Bernabeu *et al.* used the Classification and Regression Tree (CART) algorithm on simulated PMU measurements from a 4000-bus model of the California power grid to create a decision tree capable of classifying the system as either stressed or safe. Depending on which classification was chosen, the relay protection scheme in use could then be shifted towards dependability or security. The classifier itself was shown to be highly accurate – 99% of the scenarios tested were categorized correctly. An interesting by-product of this technique was that the tree revealed which PMU locations were actually necessary to assess the state of the system.

In [21], Kamwa *et al.* applied several different machine learning techniques to the voltage stability problem in order to compare the performance of black-box models (e.g., artificial neural networks (ANN), support vector machines (SVM),and random forests) with more transparent ones, such as decision trees. Using more than 60,000 cases derived from both real and simulated data, they concluded that the black-box models were significantly better than those whose structure was more easily interpreted. However, the authors also pointed out that a more transparent model with

reasonable accuracy is in many cases preferable to a highly accurate one, depending on the perspective of the user.

A variety of machine learning techniques have been applied to the transient stability problem as well. He, Zhang, and Vittal described a decision tree-based dynamic security assessment (DSA) framework used to classify the system as being secure or insecure based upon simulated PMU data [22]. Building on earlier work by Sun and Diao [23-25], they advocated boosting multiple simple decision trees as opposed to monolithic ones in order to improve accuracy and reduce the computational complexity of the updating and training process. In addition, they introduced continual updating of the trees as new cases became available. Finally, principal component analysis (PCA) was applied to significantly reduce the dimensionality of the input data. The resulting classifier was found to be highly accurate and continued to perform well even after additional contingencies were added. In 2008, a Power Systems Energy Research Center report described the creation of a decision tree-based classifier built using the CART algorithm that provided good results with PMU data taken from an actual system [26]. Later, Hashiesh et al. developed a stability predictor for the Egyptian power system using the derivatives of simulated generator bus voltage magnitude and angle as inputs to an ANN with two hidden layers, and were able achieve 91% accuracy [27]. In [28], Lezama applied the magnitudes of simulated voltage phasors to an SVM-based classifier, which was found to work well on a variety of different power systems.

Environmental Regulation Impacts on the Eastern Interconnection

Coal-fired generators provide much of the base load generating capacity in the United States due to their operational requirements and low marginal costs. Because these plants tend to be larger in capacity than other sources, they provide a significant amount of system inertia, which plays an important role in the system's dynamic response. The U.S. Environmental Protection Agency (EPA) recently finalized the Mercury and Air Toxics Standards (MATS) and the Cross-State Air Pollution Rule (CSAPR), which are regulations designed to reduce power plant emissions such as mercury, NO_x, SO_2, and ozone [29, 30]. Assuming these rules pass judicial review, as much as 30 GW of generation capacity (mainly coal and oil-fired units) will be taken offline within the next few years [31].

The North American Electric Reliability Corporation (NERC) has performed some resource adequacy analyses based on the projected deactivations, and found that there is a "significant potential impact to reliability" if MATS and CSAPR are implemented in their present forms [32, 33]. Studies conducted by the U.S. Department of Energy found that although there should be sufficient resource adequacy, "retirements of power plants or other factors could lead to grid reliability challenges in some

cases" [34]. To date, however, there do not appear to be any published studies on the possible steady-state or dynamic implications of these regulations. While it is likely that individual utilities have studied the effects of removing generators within their own service territories, the results of these studies are typically not made public. Thus, there is no clear overall picture of how the Eastern Interconnection will be affected by these regulations.

Electromechanical Speed Map Development

Power system disturbances propagate through an interconnection at speeds much less than that of light due to electrical inertia in the system, impedance, governor settings, and the amount of spinning reserve [35-38]. Thorp *et al*. modeled this phenomenon in [39] using a nonlinear partial differential equation corresponding to a discrete system, and found that their results closely matched phase propagation speeds observed in real power systems. This approach, however, requires that an accurate dynamics model of the system be known *a priori*, which is unrealistic since the grid's topology is constantly changing as lines are switched out of service, loads are disconnected, and generators are dispatched. Even if the topology is known, it is often difficult to obtain correct model parameters needed by time-domain simulation programs. Efforts by Kook to create speed maps using simulated data yielded mixed results, with estimates of speed propagation at a given location varying considerably from one event to another [35]. In [40], Gardner proposed that measurement-based speed maps could be created using PMU or FDR data, which could prove extremely useful for event triangulation algorithms such as those used by FNET. Most recently, Backhaus and Liu used techniques borrowed from the field of seismology to estimate the Green's functions for a small number of locations in the grid [41]. Unlike earlier methods that used the time delay of arrival (TDOA) during transient events to estimate the propagation speed from the disturbance to the sensor, their technique used a nearest-neighbor approach involving *ambient* FNET data obtained when the system was more or less in a steady state.

3. Analysis of Frequency Extrema in the Eastern and Western Interconnections

Introduction

The power system frequency is an important indicator of the grid's health and stability [42]. Changes in frequency reflect mismatches between generation and load. As the system becomes more heavily loaded, generators will slow down unless additional power can be supplied. Similarly, too much generation with too little load results in an increase of the system frequency. These changes are usually quite small and thus do not seriously affect the operation of the grid. Larger frequency deviations can damage machinery and cause generators to trip offline for their own protection, which can then lead to islanding, and in extreme cases, blackouts. Thus, the frequency extrema may provide a means to determine when the grid is at its most vulnerable state.

With the advent of wide-area phasor measurement systems, it is now possible to analyze the historical frequency characteristics of the grid. This chapter describes the techniques used to search for the frequency extrema within the FNET data and discusses some of the challenges associated with this type of analysis. Finally, the results of this study are presented.

Features of FNET Measurement Data

Phasor measurement data tends to be voluminous by its nature. By recording 10 data points per second from over 120 active FDRs installed throughout the world, the FNET system generates roughly 6 GB of data each day and more than 1.2 TB each year. Thus, an algorithm performing data-mining tasks on this large volume of data must be very time-efficient.

The FNET data are stored in Microsoft Access Database MDB files, which impose a 2-GB size limit [43]. Due to the volume of phasor measurement data being recorded, the FNET server application creates several database files each day to store the measurements. In order to perform any type of long-term analysis, the files must be read individually while taking into account the fact that no file contains an entire day's worth of data. MATLAB was chosen as the analysis platform for this study since its database toolbox can read directly from MDB files using freely available Open Database Connectivity (ODBC) drivers.

Because the FDR computes the frequency from the distribution-level voltage, which can be easily influenced by random behaviors in the surrounding loads, there tends to be a fair amount of noise in the computed frequency data, even after bandpass filtering of the input signal [44]. Additionally, random computational and/or sensor errors can introduce spikes in the computed values. Both of these factors make it impossible to simply choose the smallest or largest frequency value recorded by an FDR,

since such a value may not be a true extremum. To eliminate these values in the extrema results, the algorithm described here uses statistical techniques to test the suspected extrema for validity.

Description of Algorithm

The algorithm used in this study was first introduced in [45], but is described here for convenience. A flowchart of the algorithm is given in Fig. 3.1.

First, the program selects the daily extrema candidates for each FDR. In this step, the program reads each FDR's daily data files and selects the first 100 extrema in order of magnitude, excluding the obviously out-of-range values that occasionally appear. At first, it appeared that an element-by-element comparison might be required to find the extrema, however this approach proved to be too slow given the large volume of measurement data. Due to the fact that the JET database engine used by Access is already optimized to perform certain types of queries, the algorithm is designed to simply query the database to locate the possible local minimums or maximums, and then use additional program logic to determine their validity in the following steps. Thus, when searching for local minimums, it finds the 100 lowest frequency values and stores them in the candidate list in increasing order. Likewise, when searching for maximums, the list is stored in decreasing order. This approach has proven to be much faster than an element-by-element comparison.

Next, each suspected extremum is tested for validity. Two important assumptions are used for this process. The first is that the power system frequency does not change drastically (i.e., more than a few millihertz) over a short (eight-second) time interval, which is true under steady-state conditions. The second is that the small variations that do occur are normally distributed about some average value.

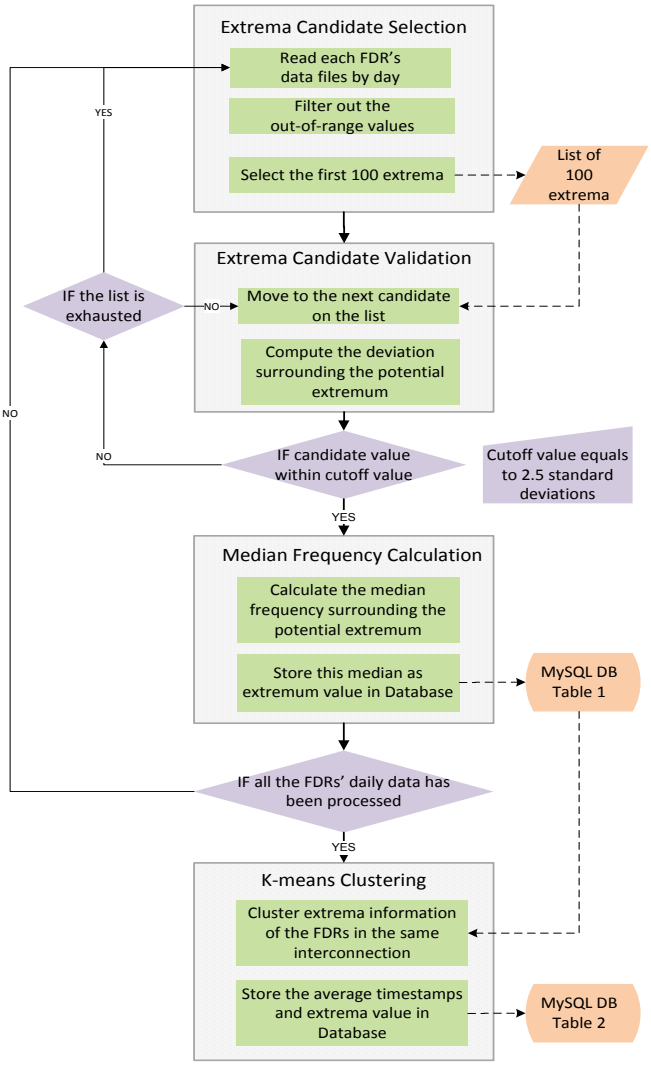


Fig. 3.1: Flowchart of frequency extrema analysis algorithm

During this step, the program computes the average and standard deviation of the frequency within the eight-second period surrounding the potential extremum. Then the standard error of the suspect value is compared with a predetermined cutoff value to check its validity. If the standard error of the suspect value is less than the cutoff, it is considered valid and the program goes directly to the next step. Thus, the cutoff value must be chosen carefully. For a normally distributed random variable, it is unlikely that a valid measurement would be more than a few standard deviations away from the mean. Analysis of several real cases showed that erroneous data points are usually more than 10 standard deviations away from the mean, making them fairly easy to detect. It should be noted that the choice of the standard deviation cutoff leaves some room for subjectivity. A smaller cutoff can cause truly valid points to be ignored (thus slowing down the program by requiring it to examine more

candidates), while a larger one will register more false positives. The cutoff value in this algorithm was set to be 2.5 standard deviations, which was chosen as a compromise between speed and accuracy.

Once a suspected extremum is found to be invalid, the program moves on to the next value in the candidate list and restarts the validity checking procedure. This repeats until the initial list of 100 extrema is exhausted. If no valid extremum can be found by then, no result is reported for this FDR. This could happen if the FDR has recorded an unusually large number of erroneous measurements that day, or if the suspected extremum is just slightly greater than 2.5 standard deviations from the mean. Using this algorithm, it takes roughly 12 minutes to analyze a single day's worth of FDR data for a particular interconnection.

After an extremum point has been validated, the program then calculates the median frequency of the two-second period surrounding the suspect value and reports this as the extremum. (Because the FDR provides data at 100-ms intervals, this leads to a 21-point median.) The goal of this step is to lessen the effects of noise in the raw data. It then stores the results in a MySQL database for further analysis.

In the last step, *k*-means clustering is used to determine the valid extrema results for each interconnection for a specific day, based upon the timestamps reported by the FDRs deployed in a particular interconnection. Because of the inherent characteristics of the frequency data and the algorithm, not all FDRs will necessarily "agree" with one another on the time of the maximum or minimum for a particular day. For example, given the data for a certain day, the algorithm might report that several FDRs have approximately the same time (within a few seconds) for an extremum, but a few others might have completely different times (perhaps hours apart). This could happen if some FDRs are not able to report data back to the FNET server during the time when the extremum occurs, such as during a network failure, or when GPS synchronization is lost. Given the synchronous nature of the grid and the fact that islanding events are extremely rare, it would be highly unlikely for some parts of the interconnection to experience extrema while others do not.

Clustering of the initial results' timestamps yields one or more sets of extrema values. The largest cluster is chosen to represent the true extrema information. Using this cluster, the algorithm calculates an average timestamp and extremum value for each interconnection for each day and then records the results in a separate database table. More information on *k*-means clustering can be found in [46] and [47].

Results and Observations

MATLAB programs were developed to implement the algorithm and run to analyze FNET data collected from 2005 to 2011. Additional programs were written to query the result database to extract information for data analysis.

A. Hourly Analysis

One motivation of this research is to investigate if frequency extrema are more likely to appear during certain periods of the day than others. This can be determined by counting the number of extrema occurring within each hour for a particular year categorized by different interconnections and extrema type (minimum or maximum). The results for minimums in the Eastern Interconnection are shown in Fig. 3.2. Note that Universal Coordinated Time (UTC) is used throughout this study.

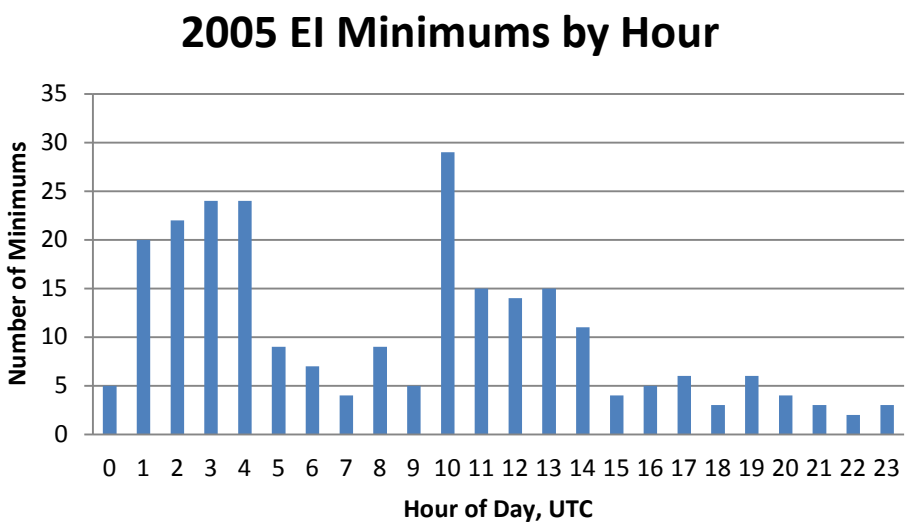
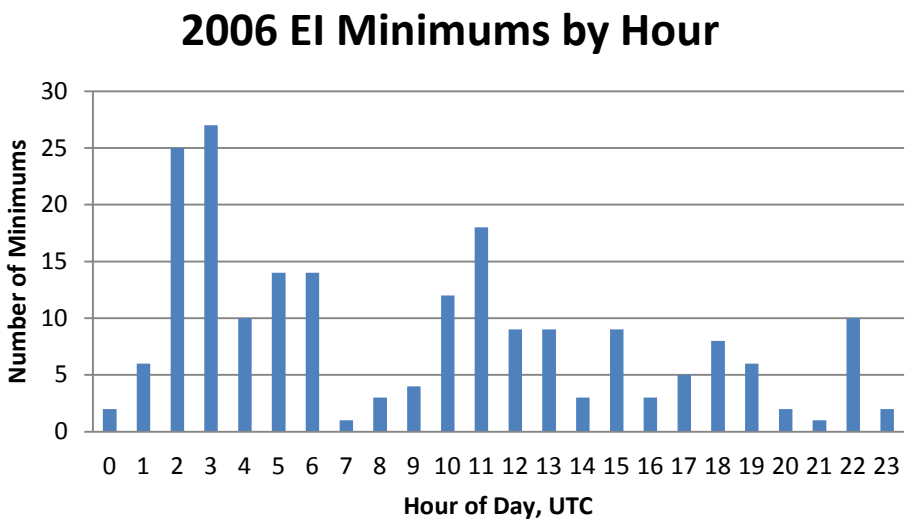


Fig. 3.2: Hourly distribution of frequency minimums within the Eastern Interconnection



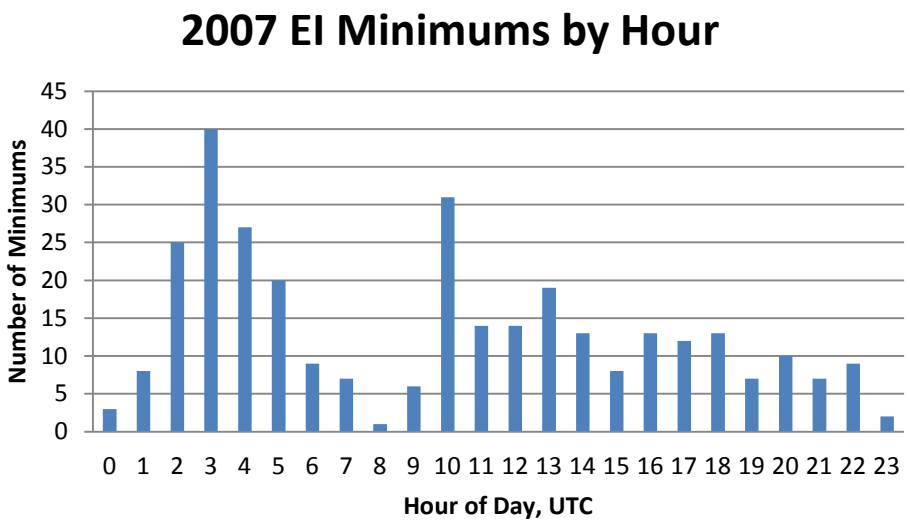
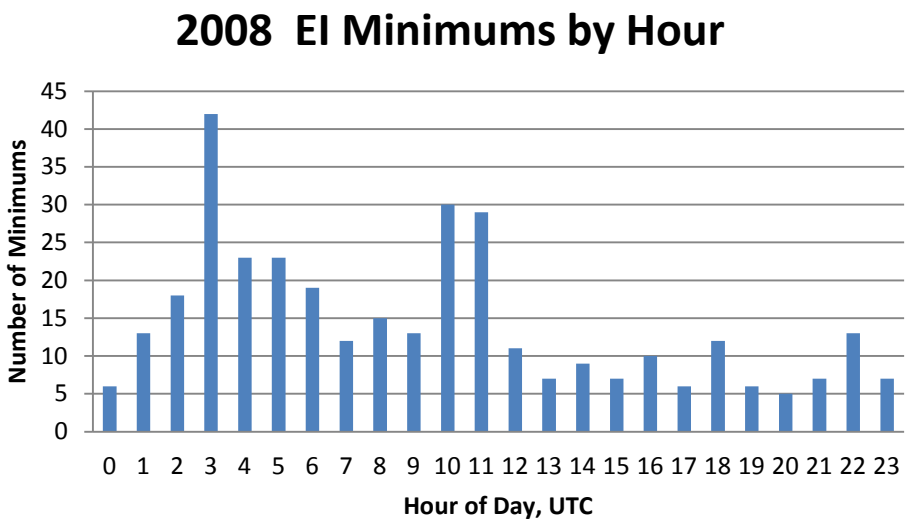


Fig. 3.2: Continued.



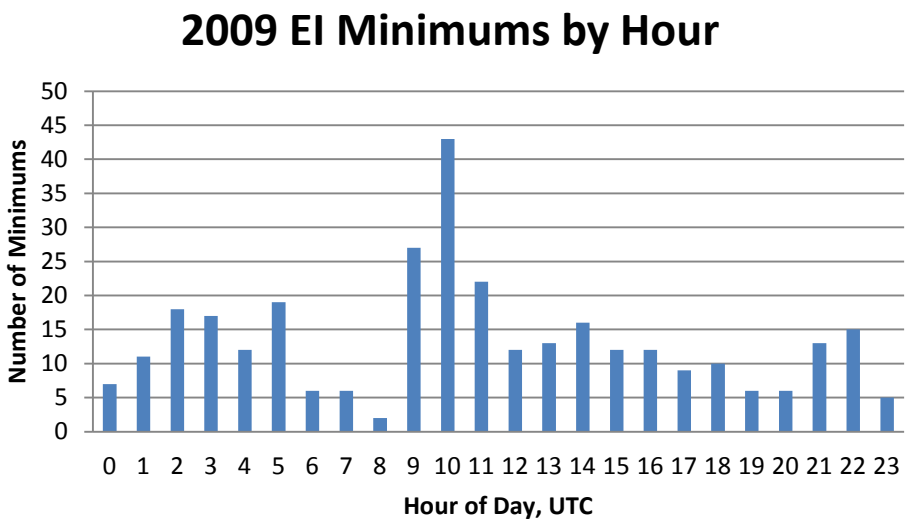
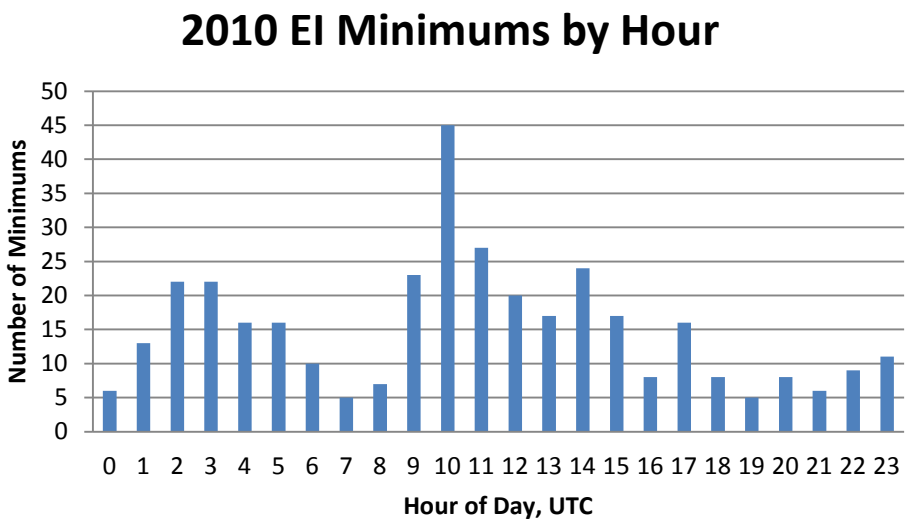


Fig. 3.2: Continued.



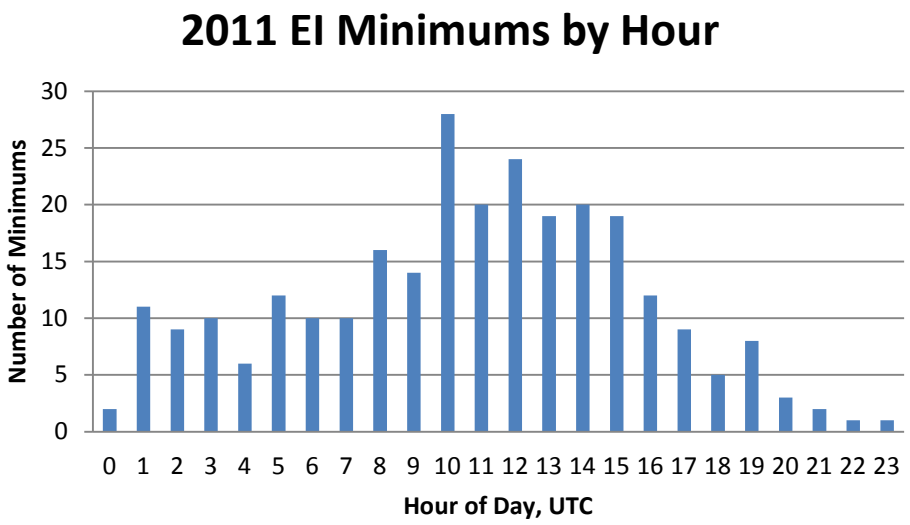


Fig. 3.2: Continued.

As shown in Fig. 3.2, the hourly distribution of frequency minimums appears to have some similarities from year to year within the EI. There are two noticeable spikes around 3:00 and 10:00 UTC for most of the years, with the 10:00 spike being present in all years. In most cases, the minimums appear to follow a roughly bimodal distribution. Interestingly, the hourly distributions of frequency minima in the EI do not seem to correspond strongly with the hourly distributions of generator trips in that interconnection as one might expect, though a peak around 3:00 UTC can be observed in both (Fig. 3.3).

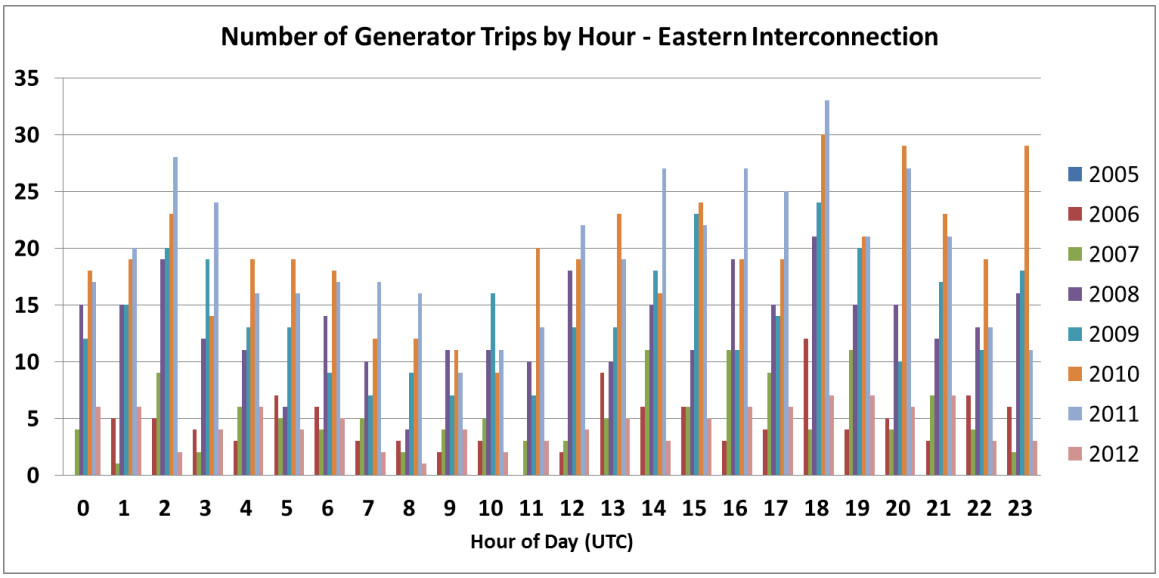


Fig. 3.3: Number of Generator Trips by Hour - Eastern Interconnection

The hourly distribution of frequency maximums in the EI are shown in Fig. 3.4.

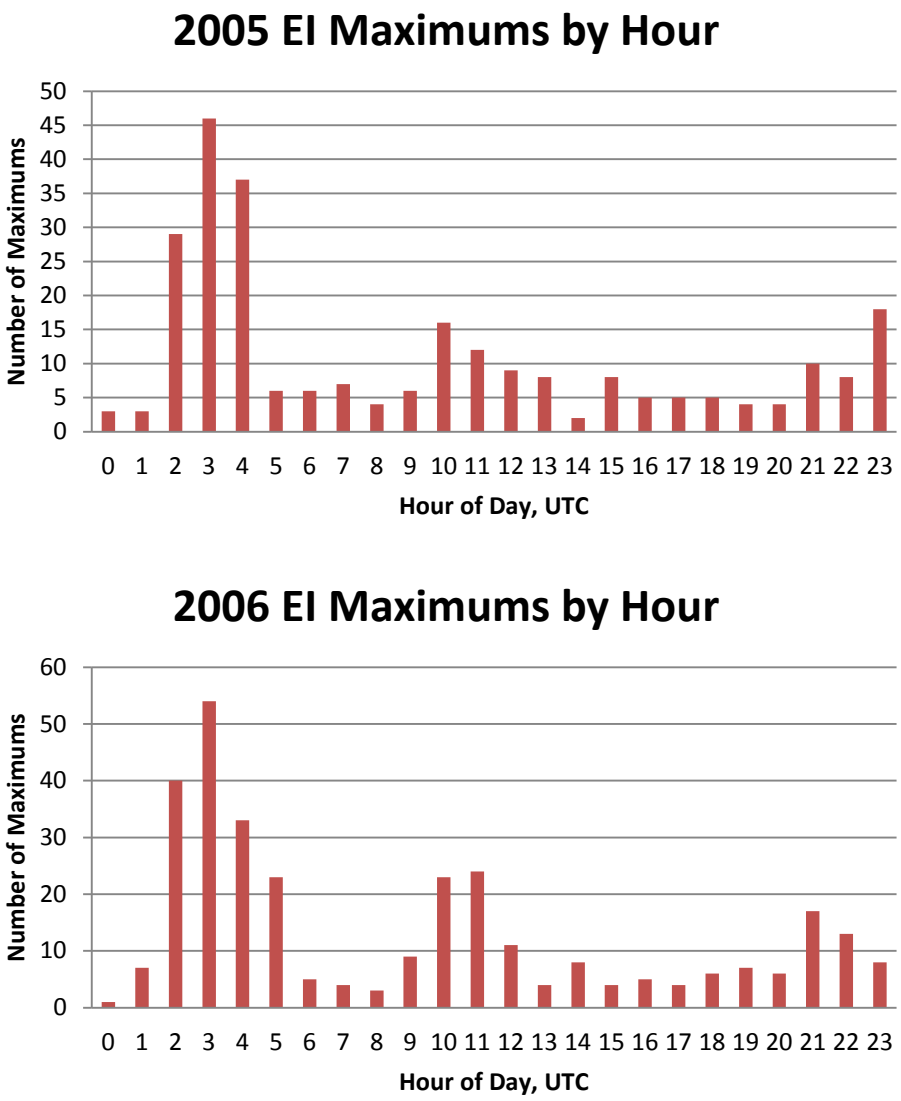
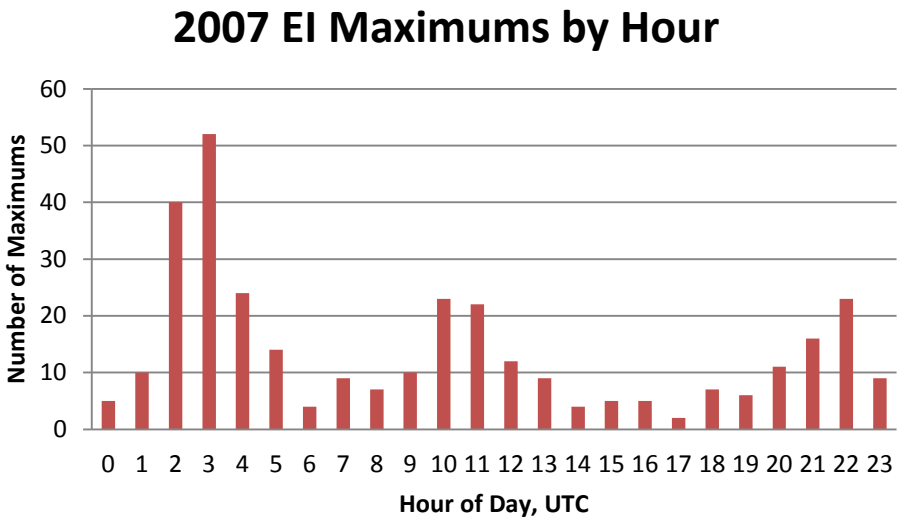


Fig. 3.4: Hourly distribution of frequency maximums within the Eastern Interconnection



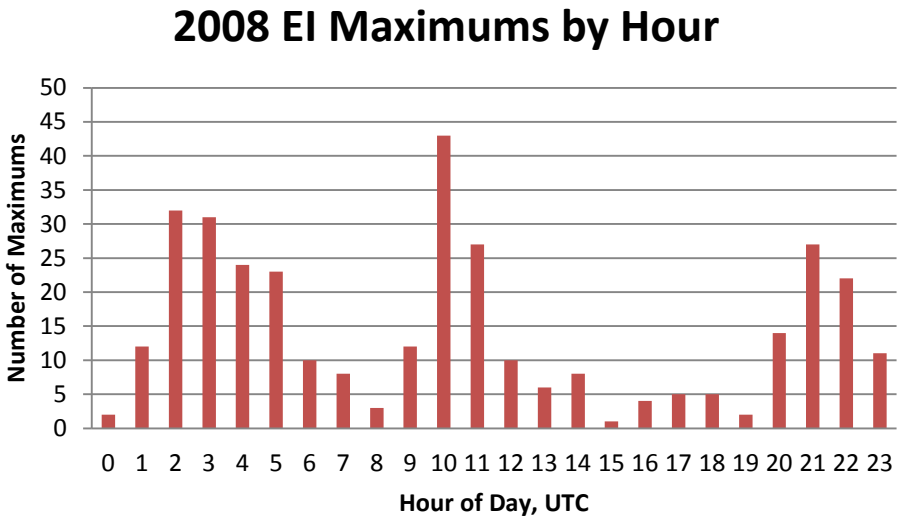
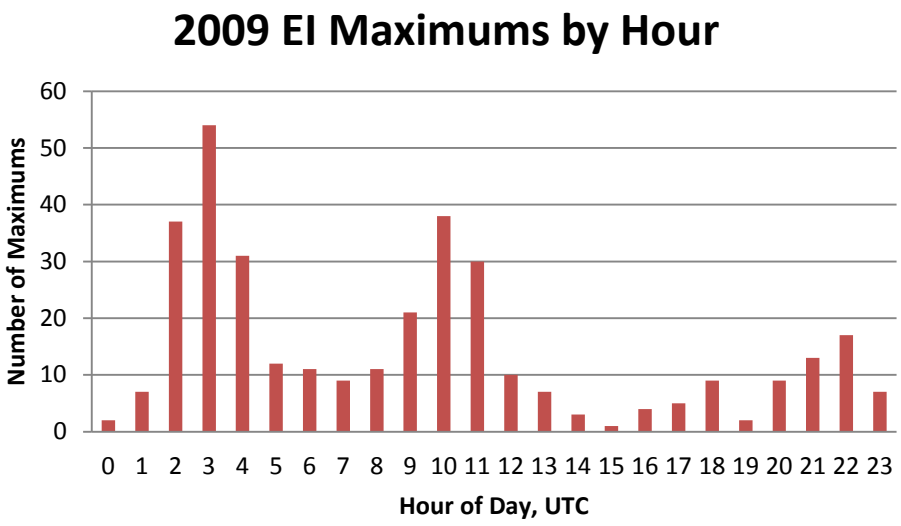


Fig. 3.4: Continued.



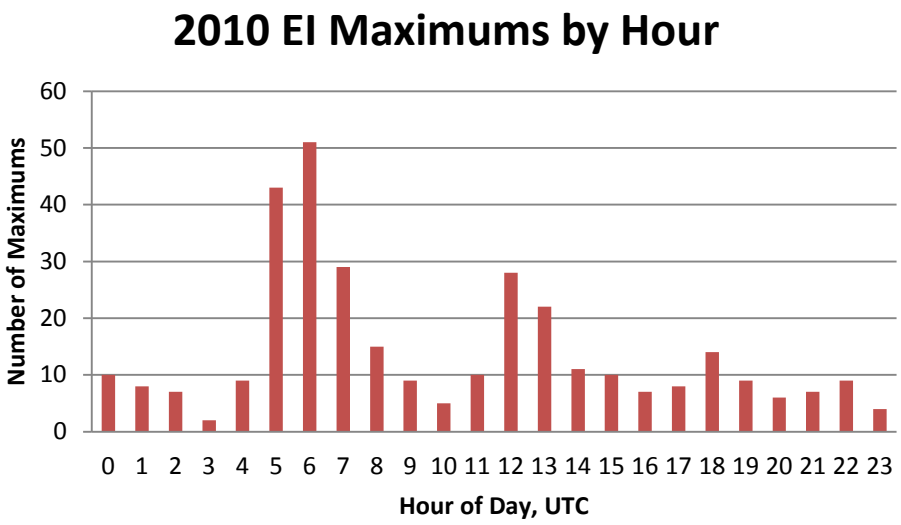


Fig. 3.4: Continued.

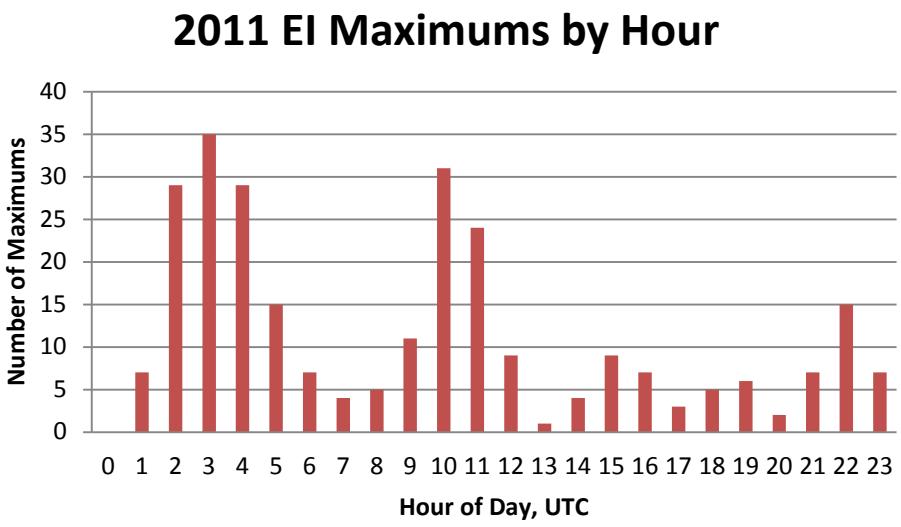


Fig. 3.4: Continued.

One interesting observation is that the hourly distribution of maximums within the EI shows a similar pattern to that of the minimums, particularly in the peaks around 03:00 and 10:00 UTC. However, besides those two dominant spikes, the hourly distribution of maximums also has one more relatively small peak at the end of the UTC day. Additionally, the 2010 EI maximums are shifted by several hours from their usual positions. The hourly distribution of load shedding/pumped storage disconnection events in the EI is shown in Fig. 3.5. Here, there is an obvious peak around 10:00 UTC, which does seem to correspond with one of the peaks in the frequency maxima distribution.

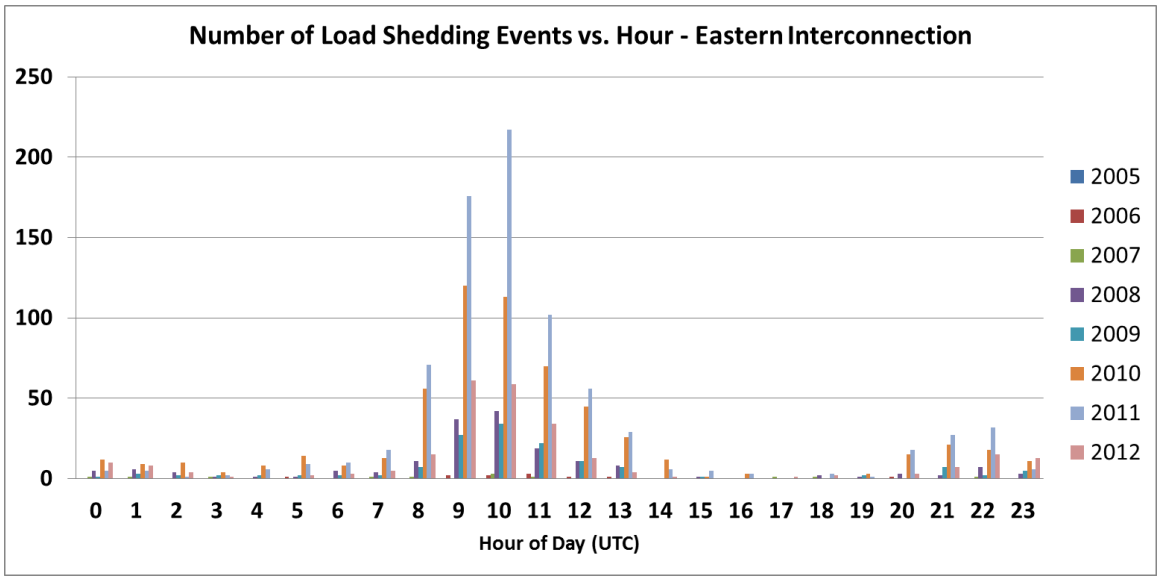


Fig. 3.5: Hourly distribution of load shedding/pumped storage disconnection events in the Eastern Interconnection

The hourly distributions for minimums within the Western Interconnection (WECC) are shown in Fig. 3.6.

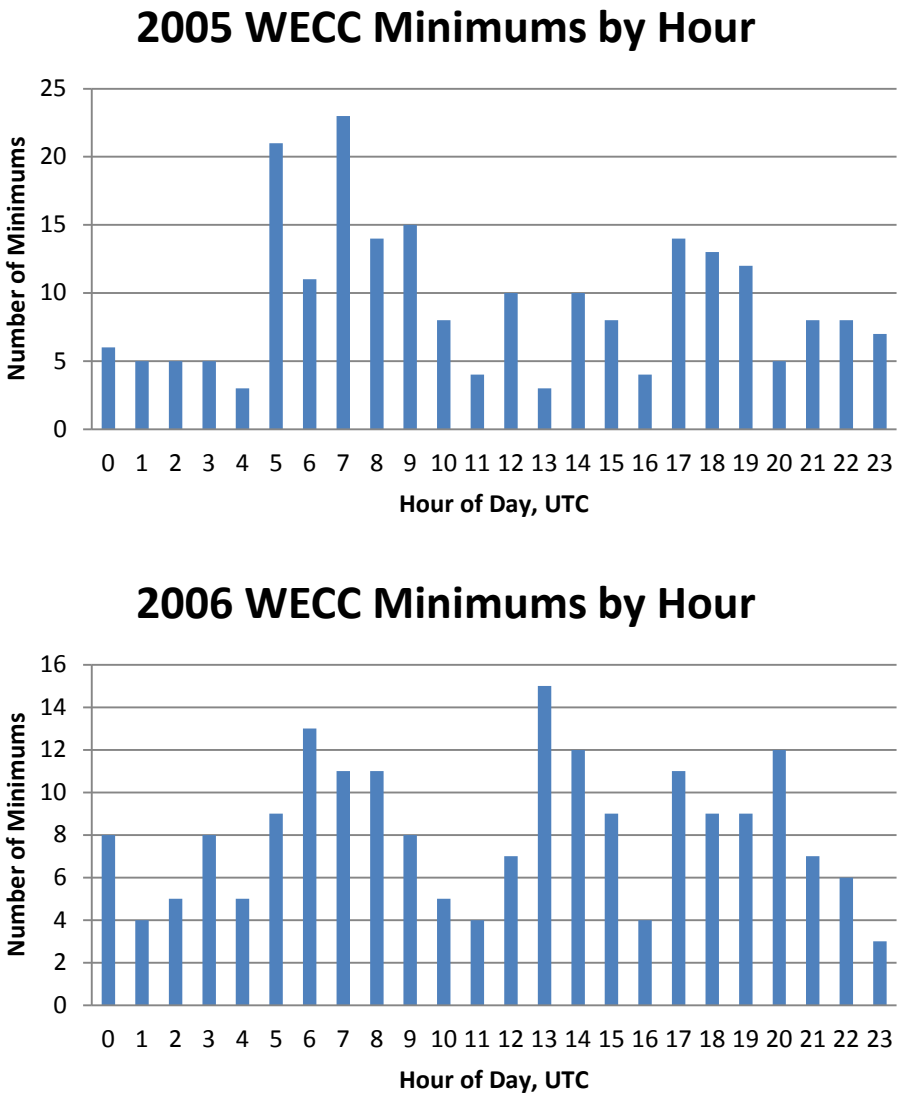
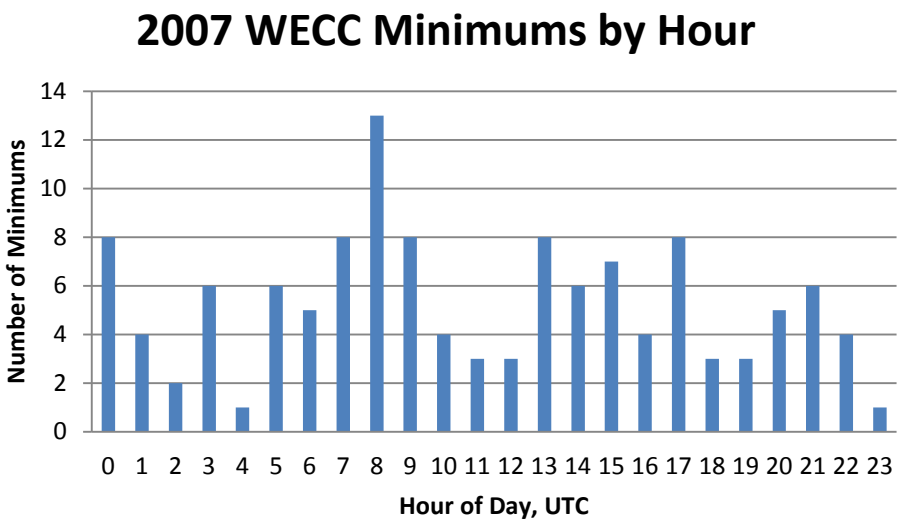


Fig. 3.6: Hourly distribution of frequency minimums within the Western Interconnection



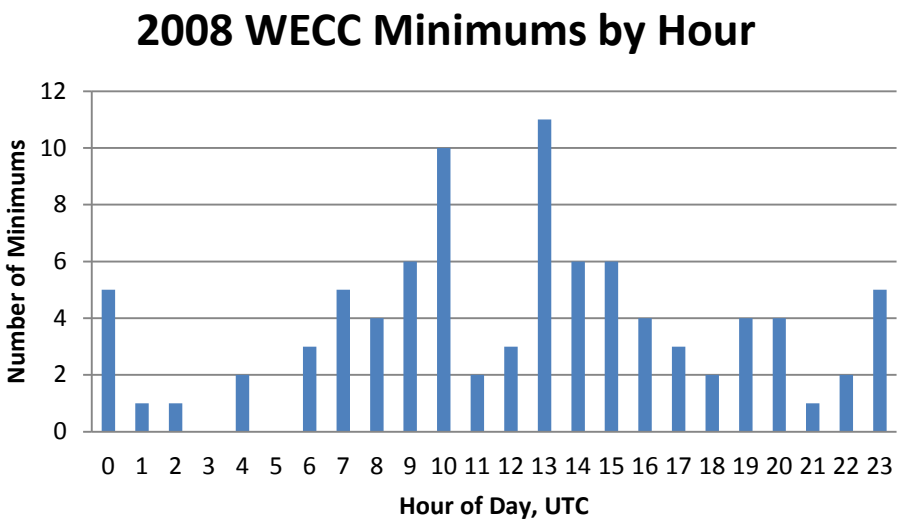
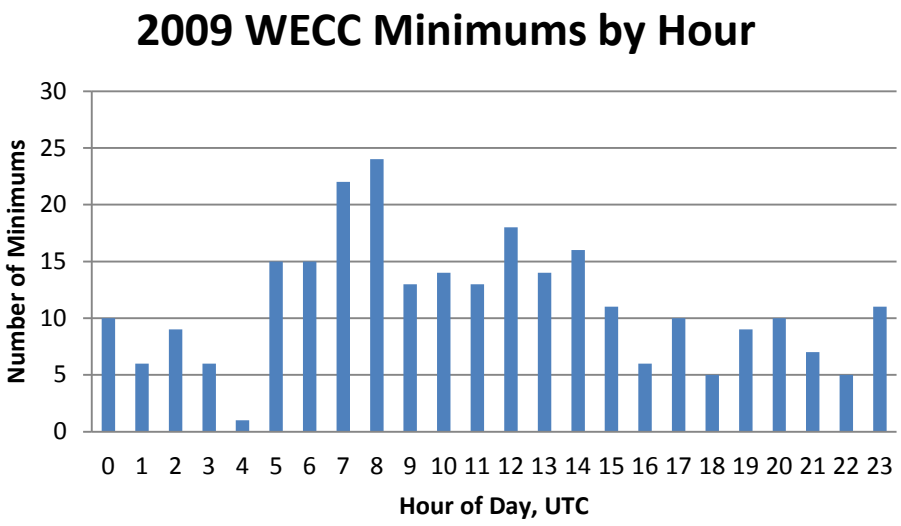


Fig. 3.6: Continued.



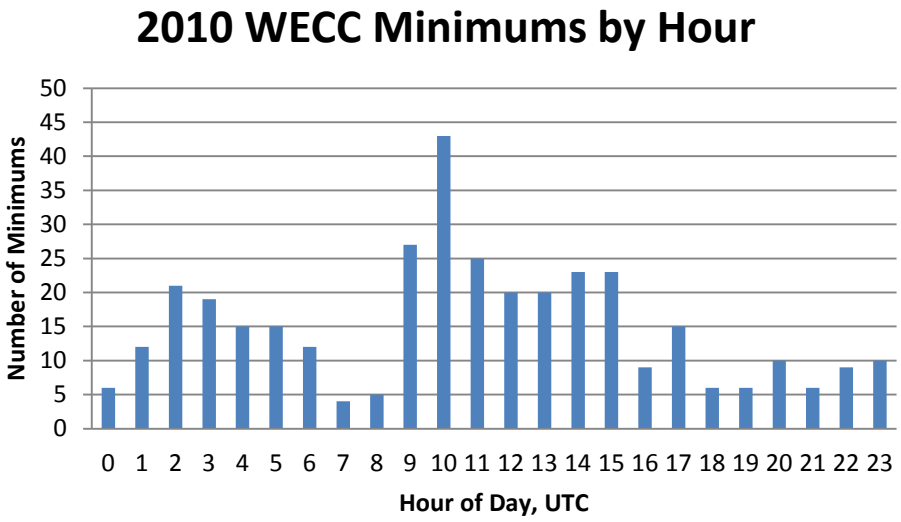


Fig. 3.6: Continued.

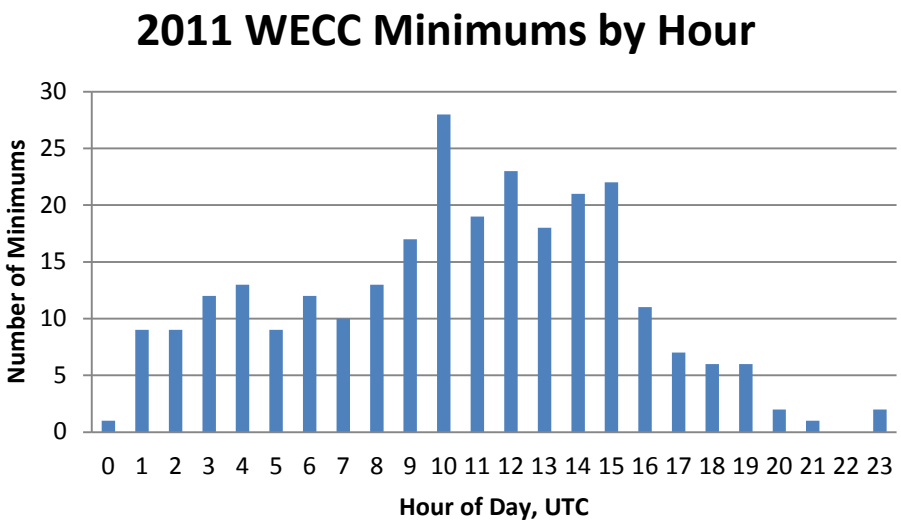


Fig. 3.6: Continued.

For the WECC, although some cases do show a spike around 10:00 UTC, the hourly distribution of minimums does not appear to follow a similar trend over the years as it does in the EI. In a similar fashion, the hourly distribution of generator trips appears to be quite random from year to year in the WECC (Fig. 3.7).

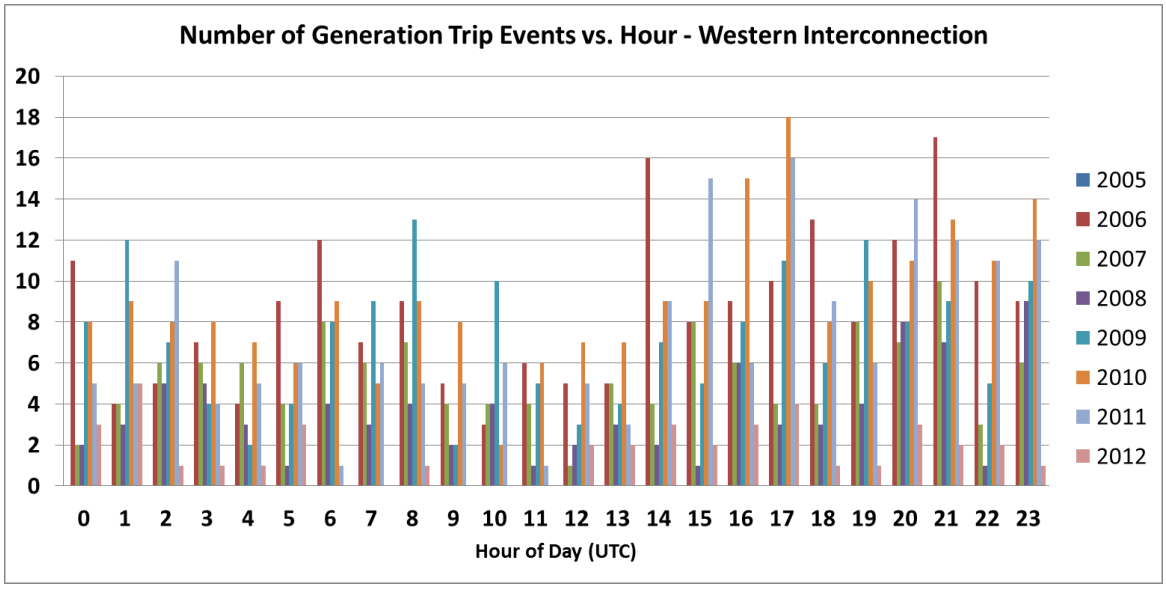


Fig. 3.7: Hourly distribution of generation trip events in the Western Interconnection

The hourly distributions of maximums within WECC from 2005-2011 are shown in Fig. 3.8.

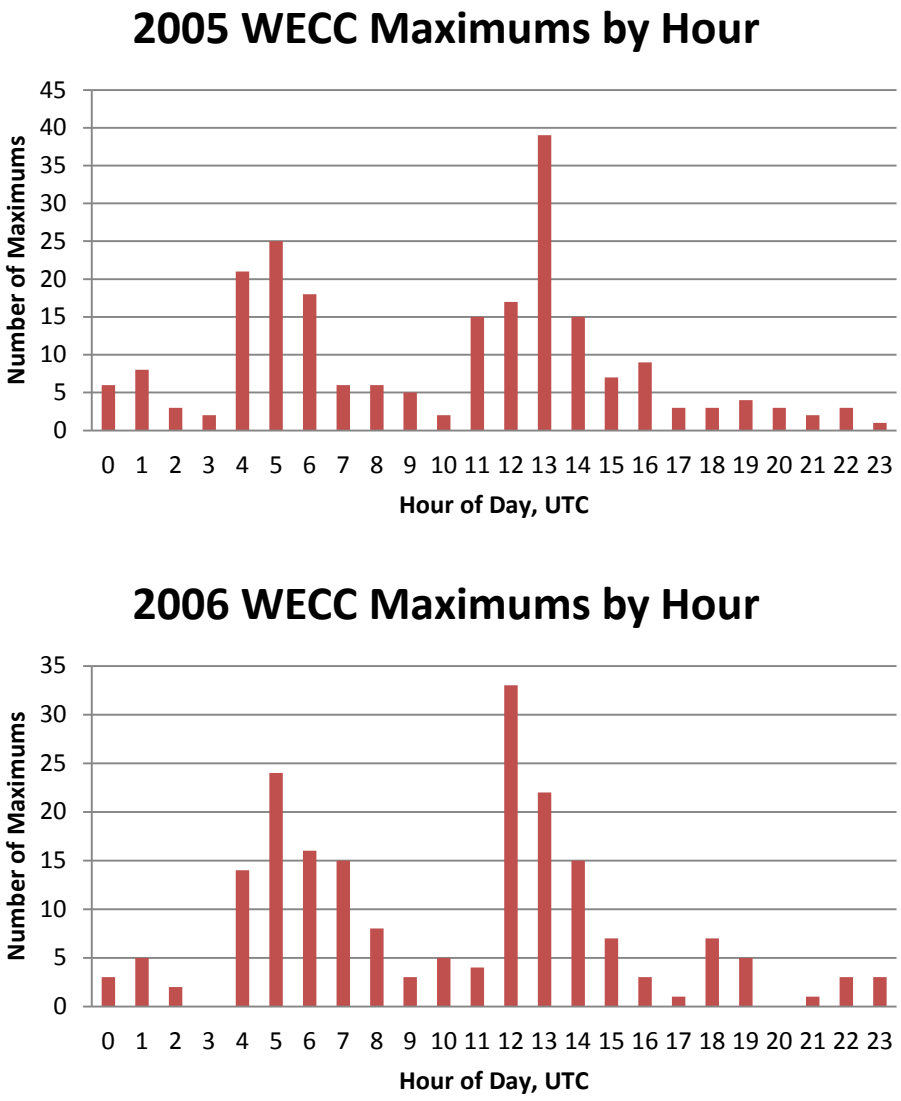
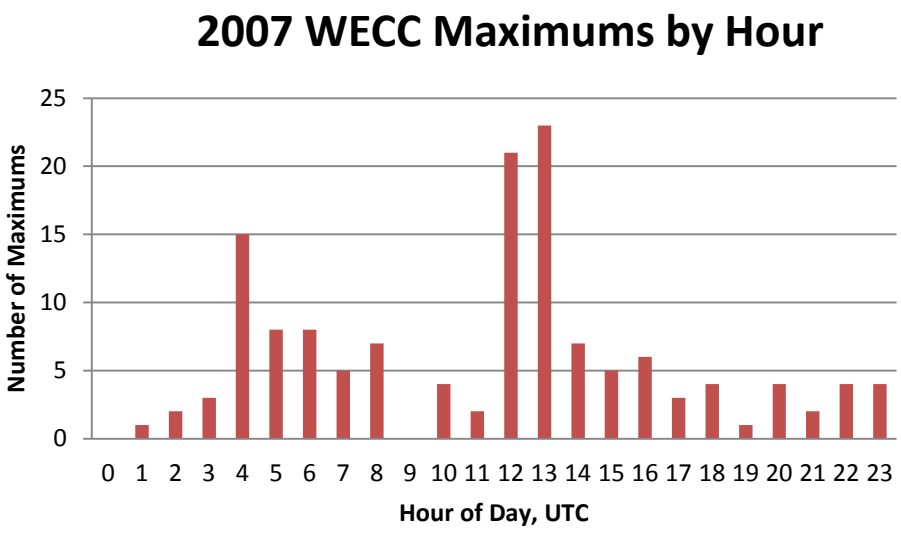


Fig. 3.8: Hourly distribution of frequency maximums within the Western Interconnection



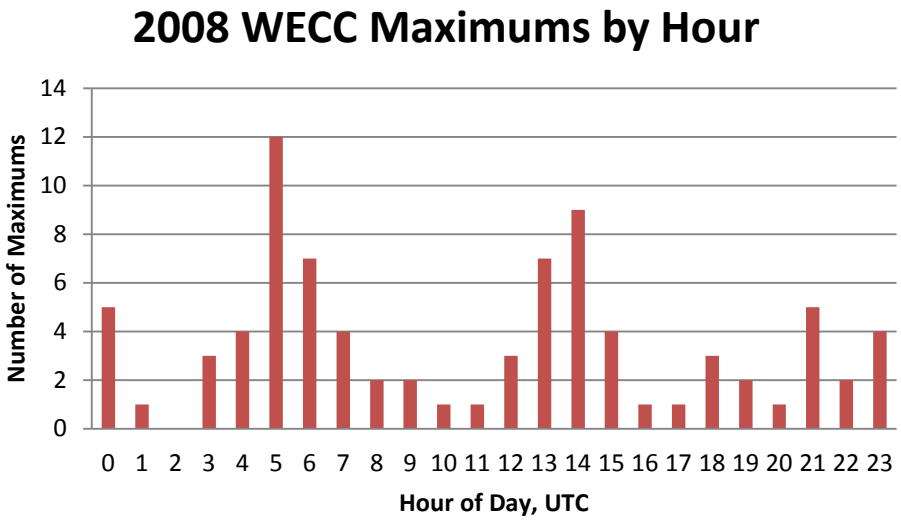
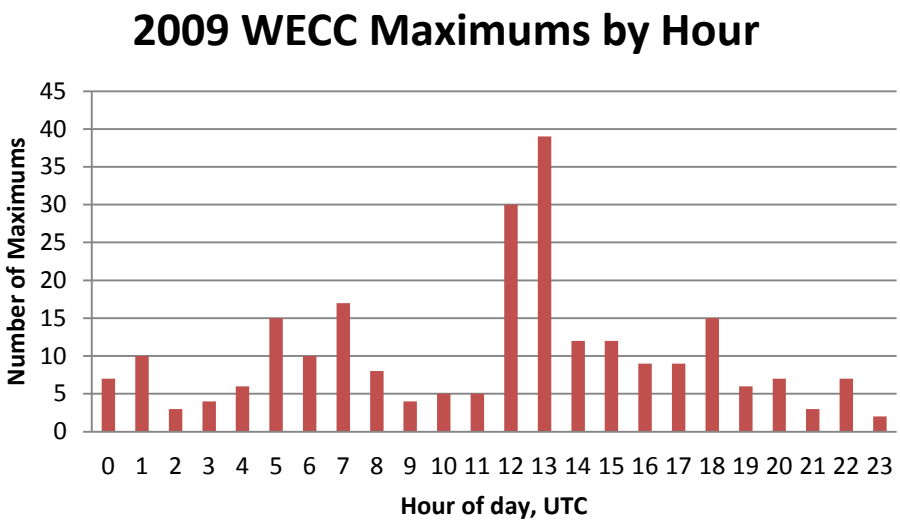


Fig. 3.8: Continued.



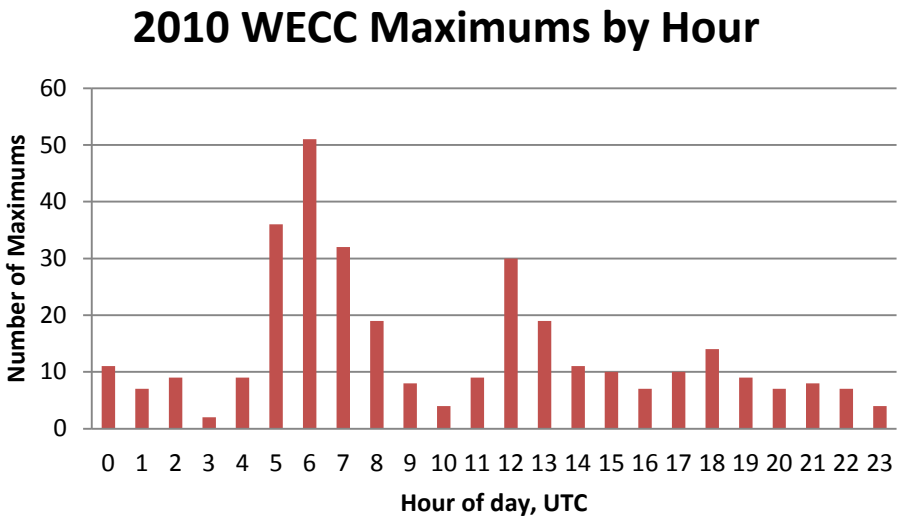


Fig. 3.8: Continued.

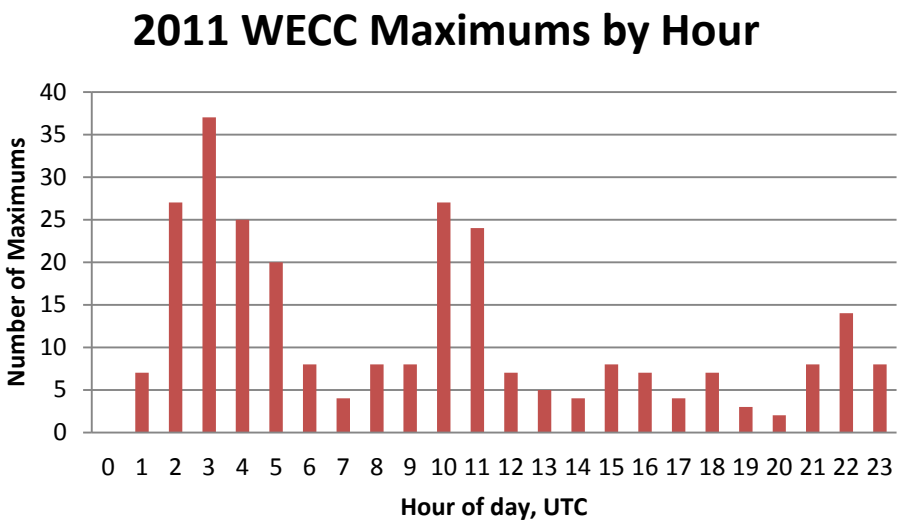


Fig. 3.8: Continued.

Some common characteristics can be observed in certain years, most notably the two obvious peaks fluctuating around 05:00 and 13:00 UTC from year to year. The 2011 data appear to show a one-hour shift from the previous years; the reasons for this remain unclear, but different system operating procedures could be a possible cause. A peak in the number of load shedding/pumped storage disconnection events is also observed around 13:00 UTC as shown in Fig. 3.9.

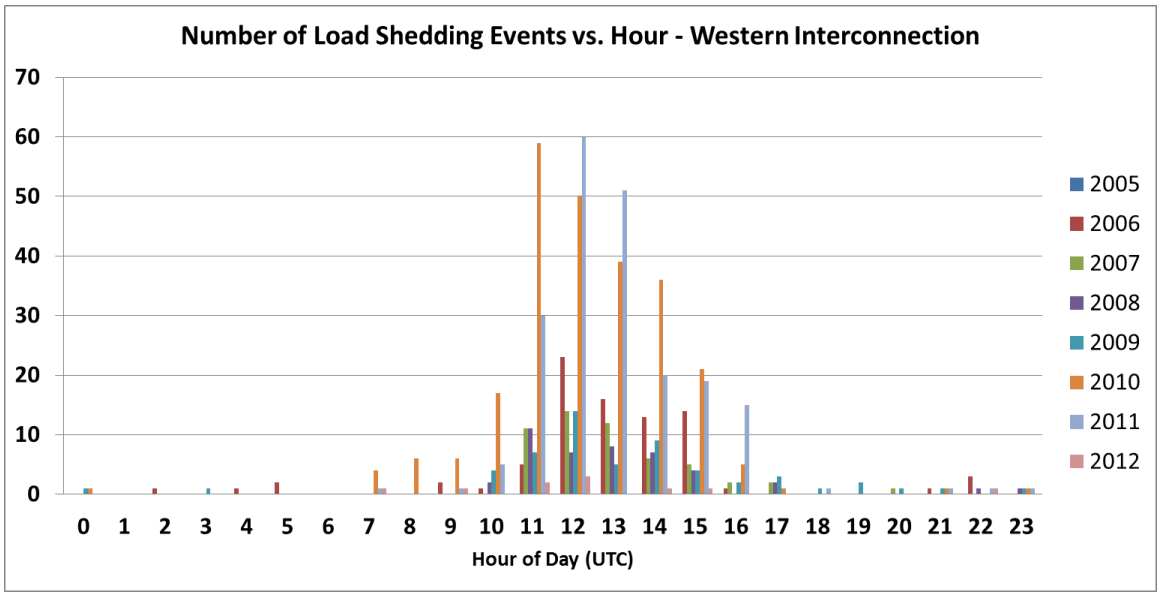


Fig. 3.9: Hourly distribution of load shedding/pumped storage disconnection events in the Western Interconnection

B. Daily Analysis

A second goal of this research is to investigate if particular days of the week experience more severe extrema than others. To do this, the average extrema values for each day of the week were calculated and are shown in Fig. 3.10 and Fig. 3.11.

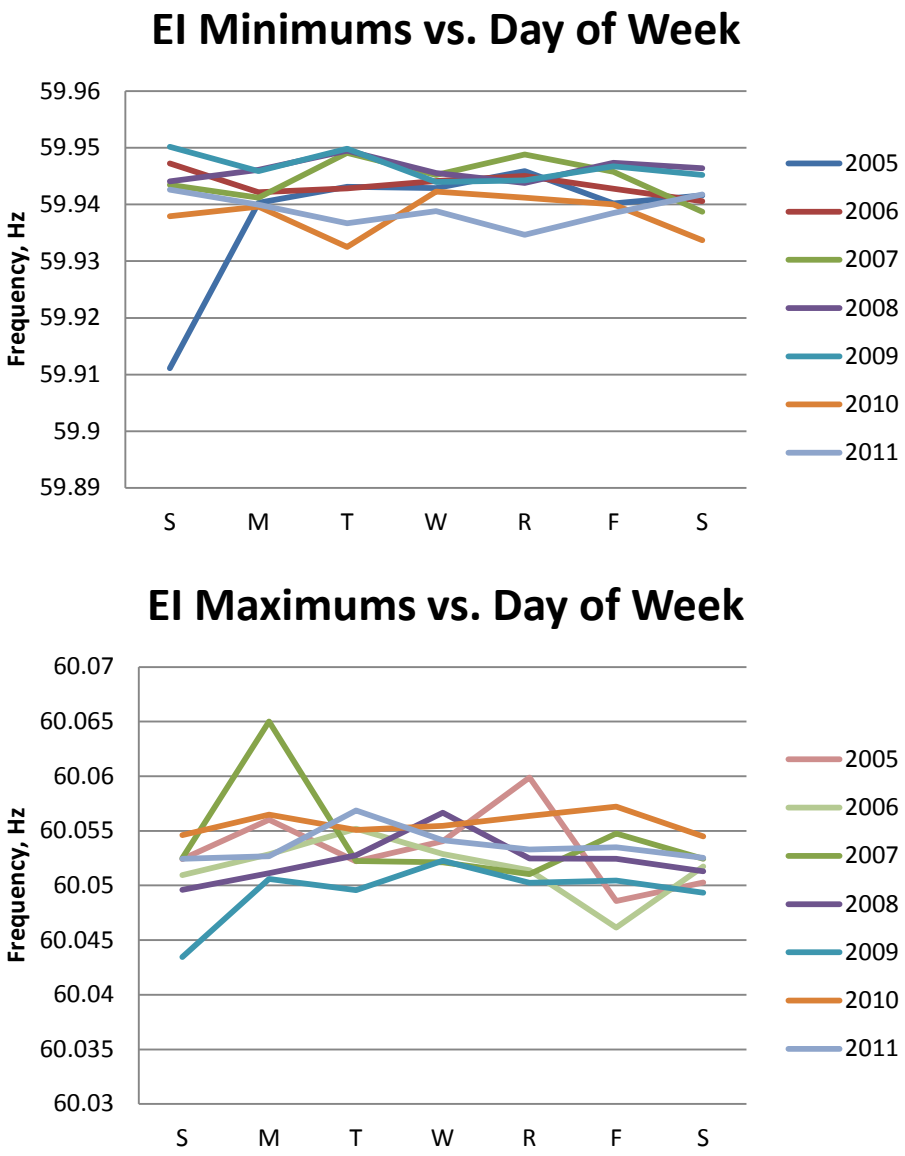
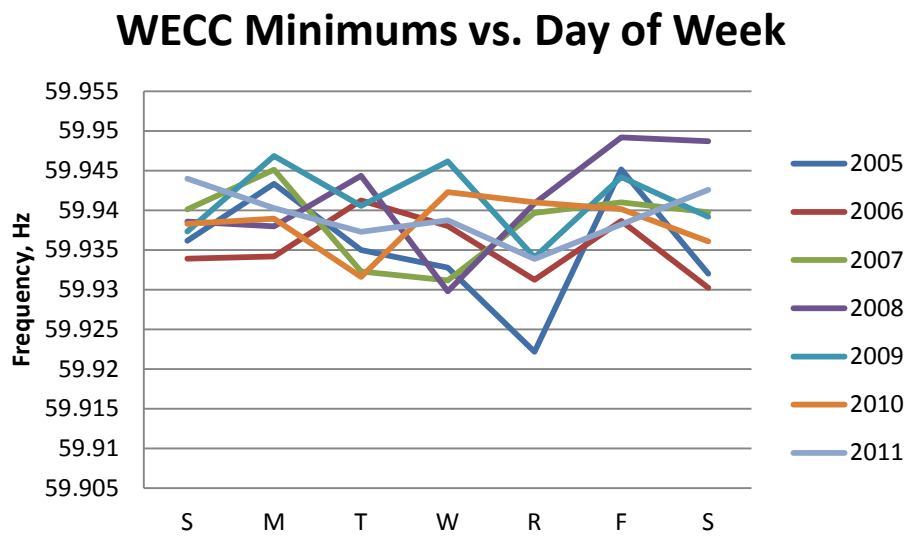


Fig. 3.10: Average daily values of frequency extrema within the Eastern Interconnection, 2007-2011



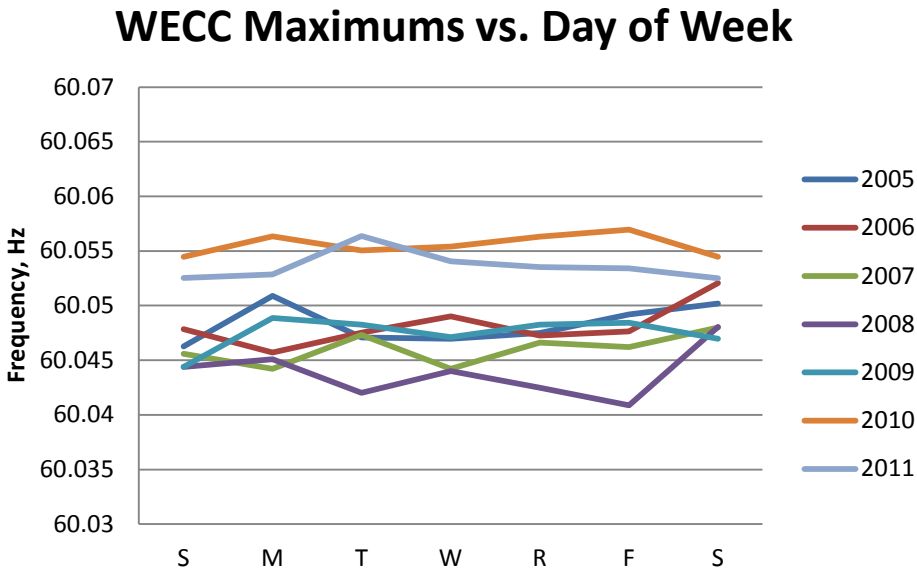


Fig. 3.11: Average daily values of frequency extrema within the Western Interconnection, 2007-2011

It can be observed from Fig. 3.10 and Fig. 3.11 that there does not seem to be any correlation between the day of the week and the magnitude of either kind of extremum in either interconnection. It is also interesting to note that the day-of-week minimums in the EI tended to be lower for 2010 and 2011 than in previous years, while the recent WECC maximums tended to be higher.

C. Weekly Analysis

Similar to the day-of-week analysis, the results were further examined to see if the magnitude of the frequency extrema is affected by the week of the year. These results are shown in Fig. 3.12 and

Fig. 3.13. Some series of the plots have missing data points, probably due to data corruption or failure of the clustering algorithm to converge.

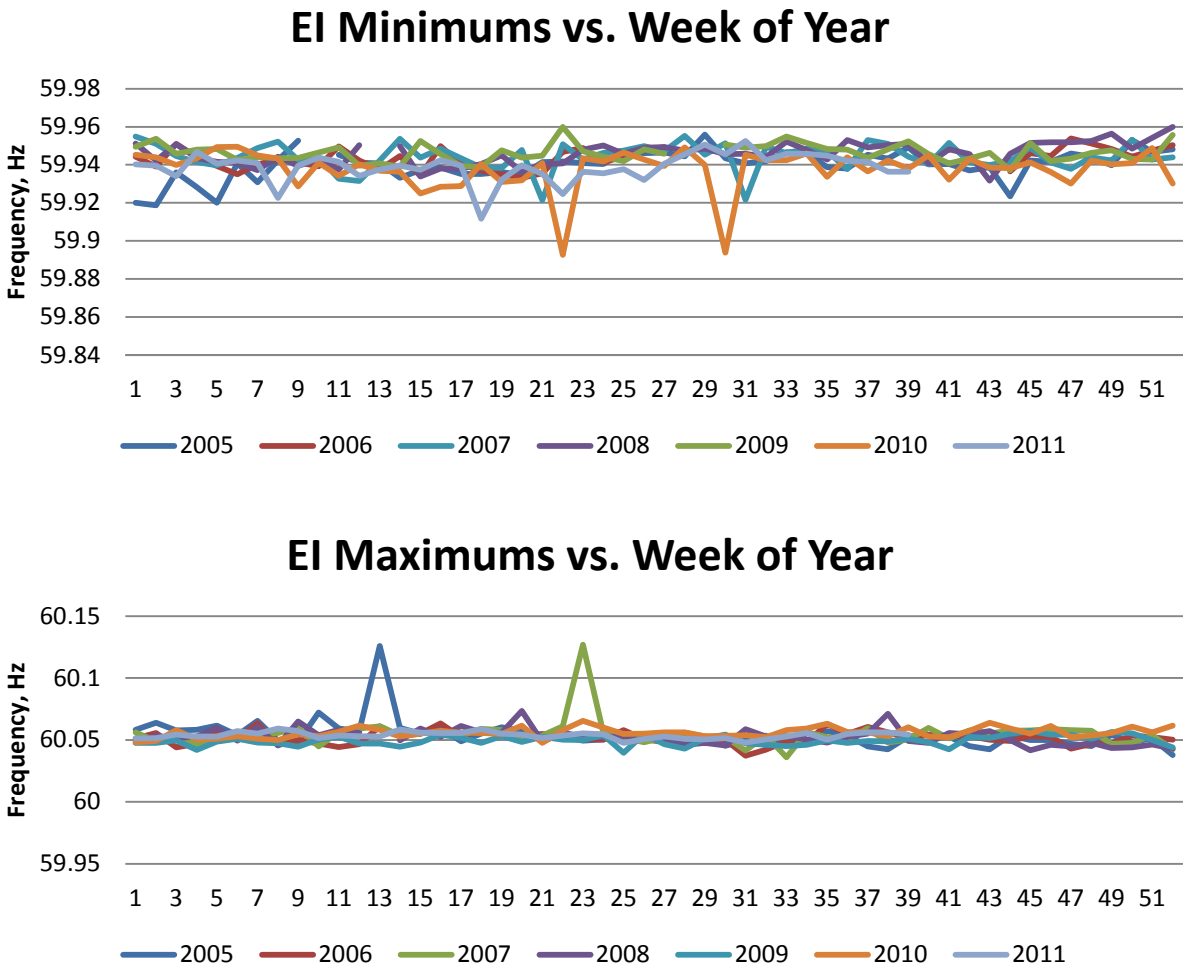


Fig. 3.12: Average weekly values of frequency extrema within the Eastern Interconnection, 2007-2011

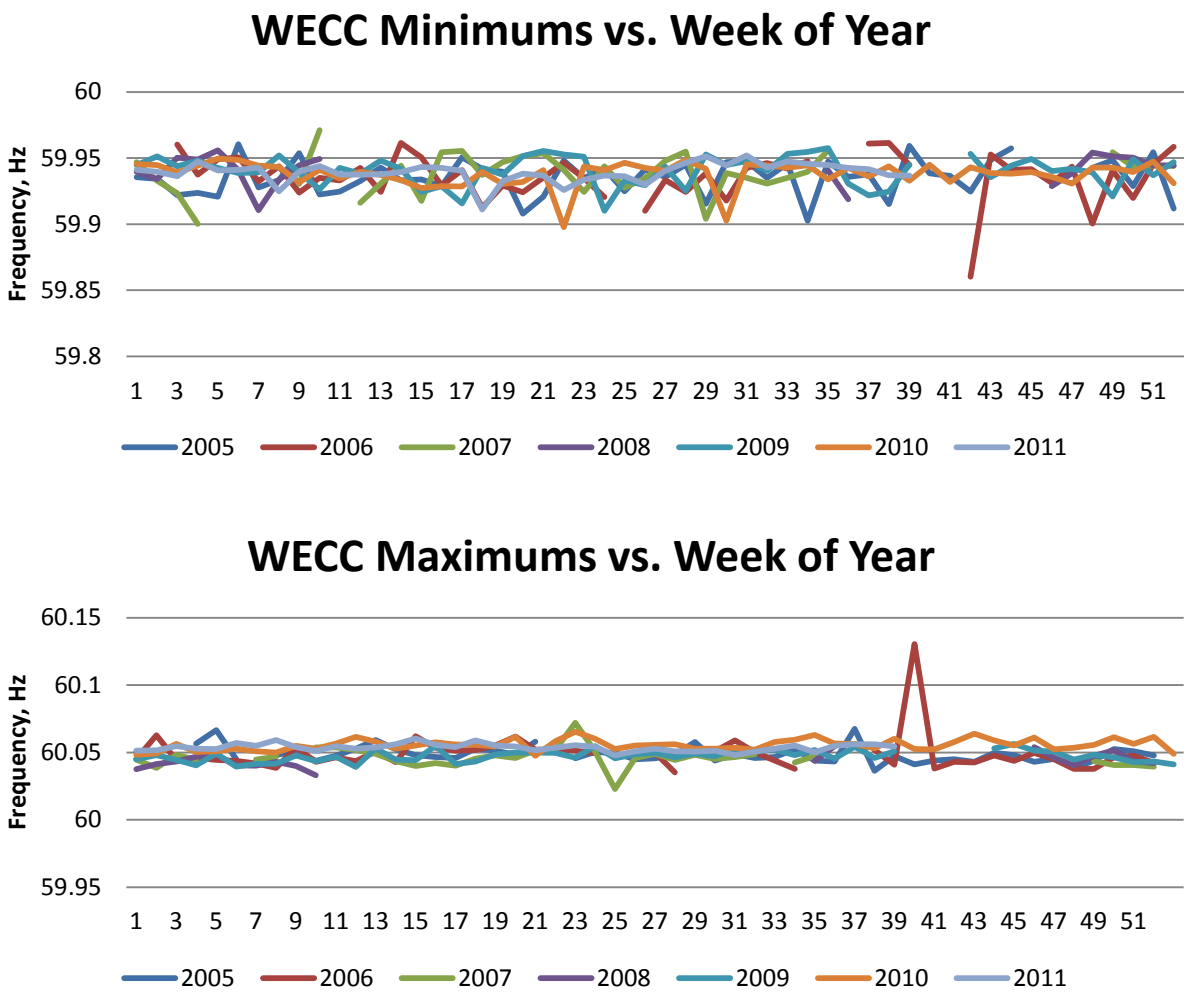


Fig. 3.13: Average weekly values of frequency extrema within the Western Interconnection, 2007-2011

Based on the observation of these plots, it does not appear that the magnitude of the frequency extrema depends on the week of the year.

D. Monthly Analysis

Seasonal variation in the frequency response has been noted previously in [48]. Thus, it is interesting to examine how the extrema change throughout the year. For this analysis, the average extrema values for each month were compared across several years (Fig. 3.14 and Fig. 3.15). Although some segments do seem to parallel each other from year to year, it is difficult to conclude with certainty that there is a relationship between the average extrema magnitudes and the month of the year.

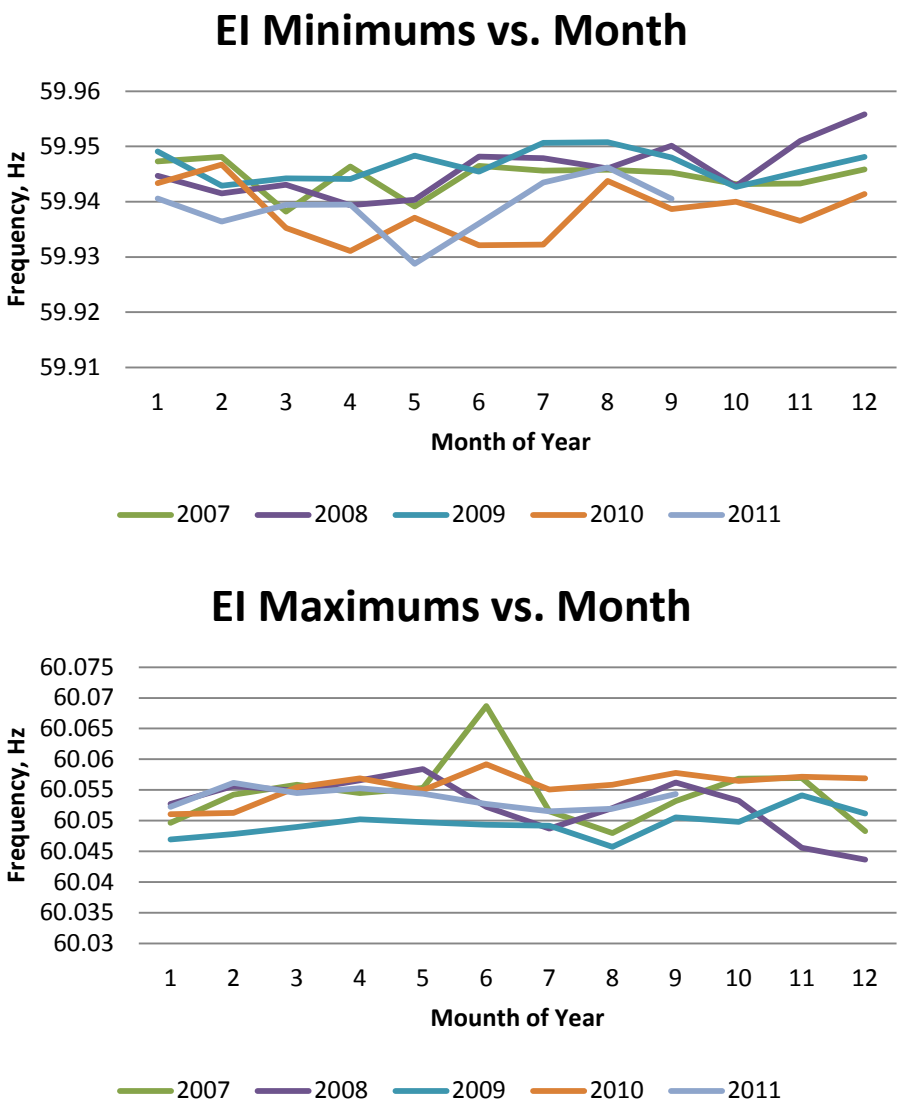


Fig. 3.14: Average monthly values of frequency extrema within the Eastern Interconnection, 2007-2011

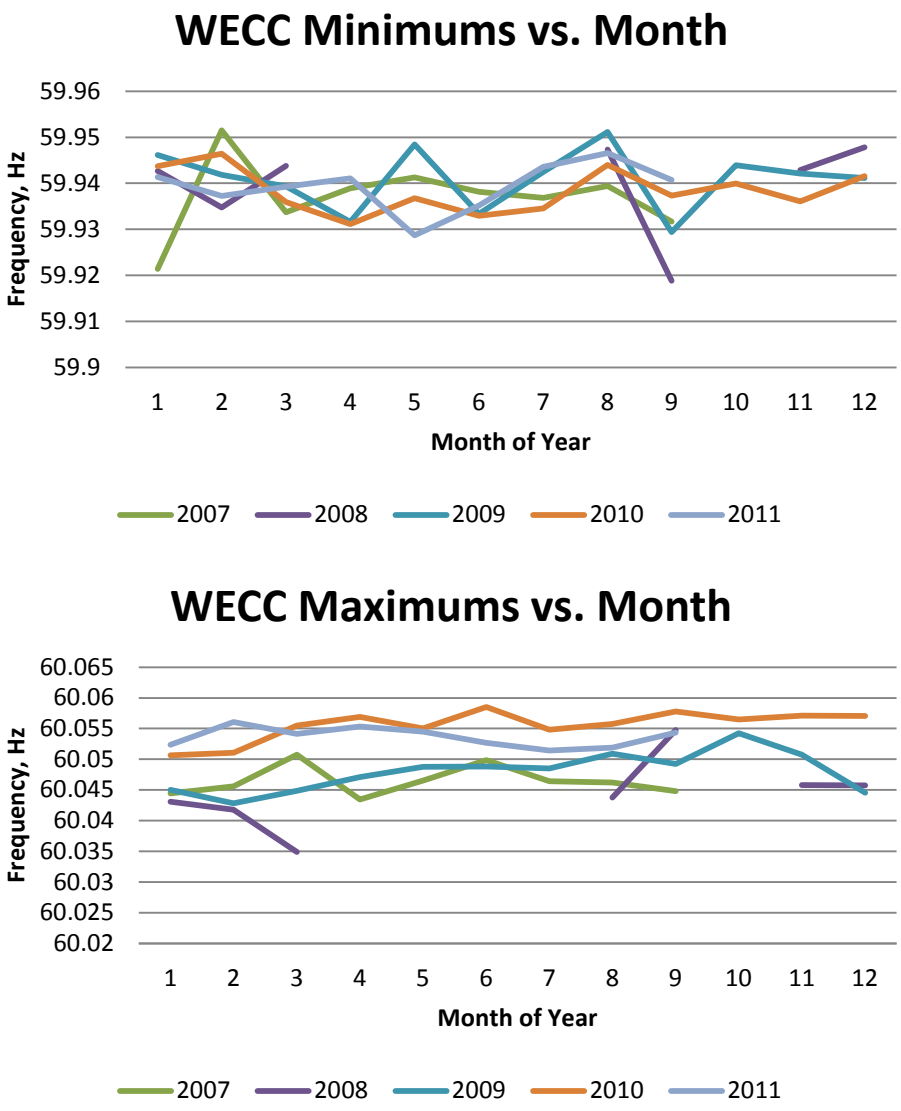


Fig. 3.15: Average monthly values of frequency extrema within the Western Interconnection, 2007-2011

E. Yearly Analysis

The analysis of the yearly average extrema may indicate the overall performance of the power grid over time. A smaller absolute value between the extrema and the nominal frequency value implies better frequency control. The results are shown in Fig. 3.16.

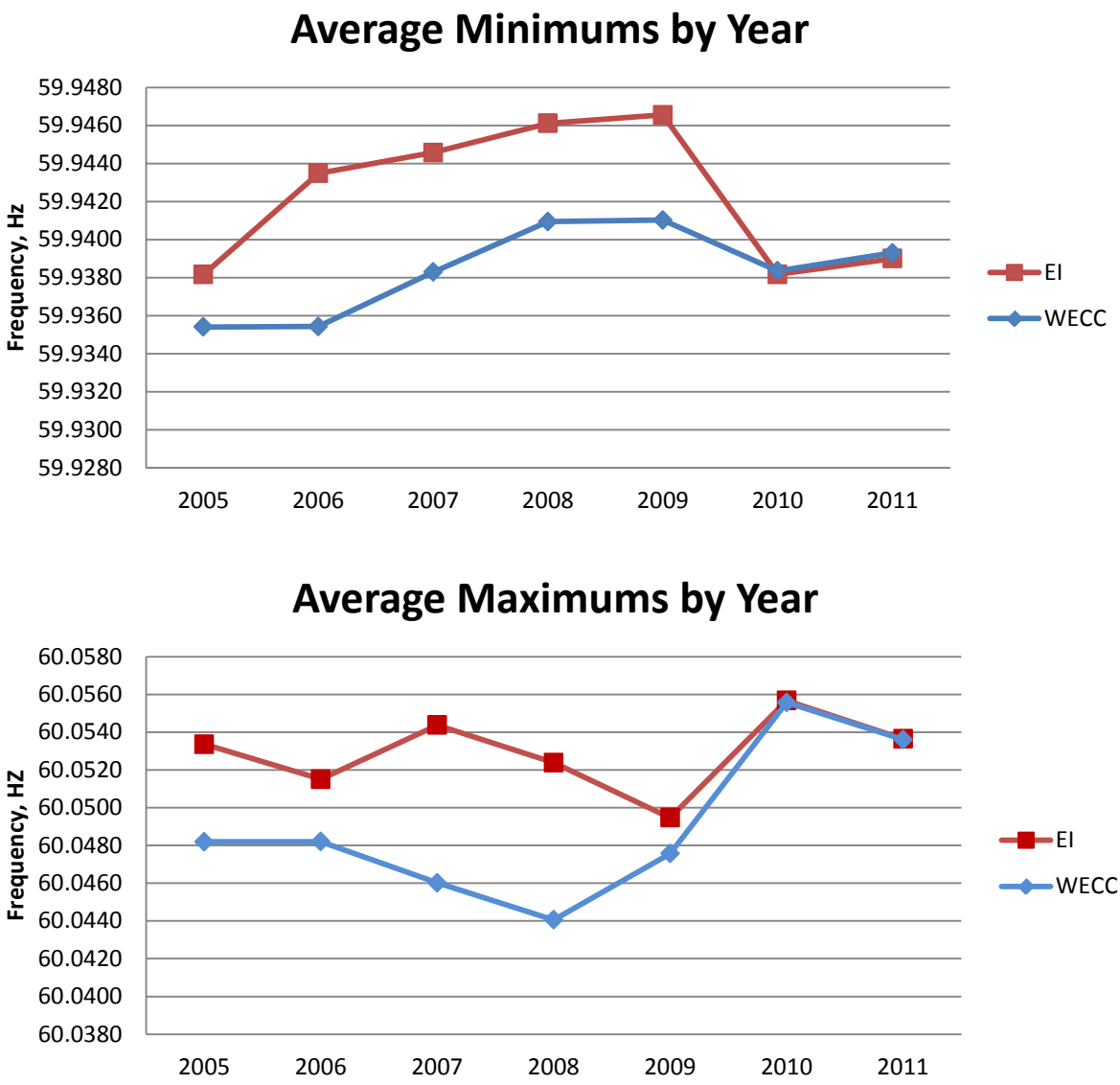


Fig. 3.16: Average yearly values of frequency extrema within the Eastern and Western Interconnections, 2007-2011

Fig. 3.16 shows that the average minimums in both the EI and WECC increased from 2007 to 2009, but dropped down to almost the same value in 2010 before increasing slightly in 2011. Likewise, the second plot in Fig. 3.16 shows that the average maximums in the EI decreased from 2007 to 2009 before increasing in 2010 and declining slightly in 2011. The WECC maximums followed a similar trend except for the change from 2008 to 2009. The values converge for 2010 and 2011 in a similar fashion.

Fig. 3.17 shows the standard deviations of the extrema over time. Interestingly, the deviations of both interconnections tend to track each other, though the reason for this remains unclear.

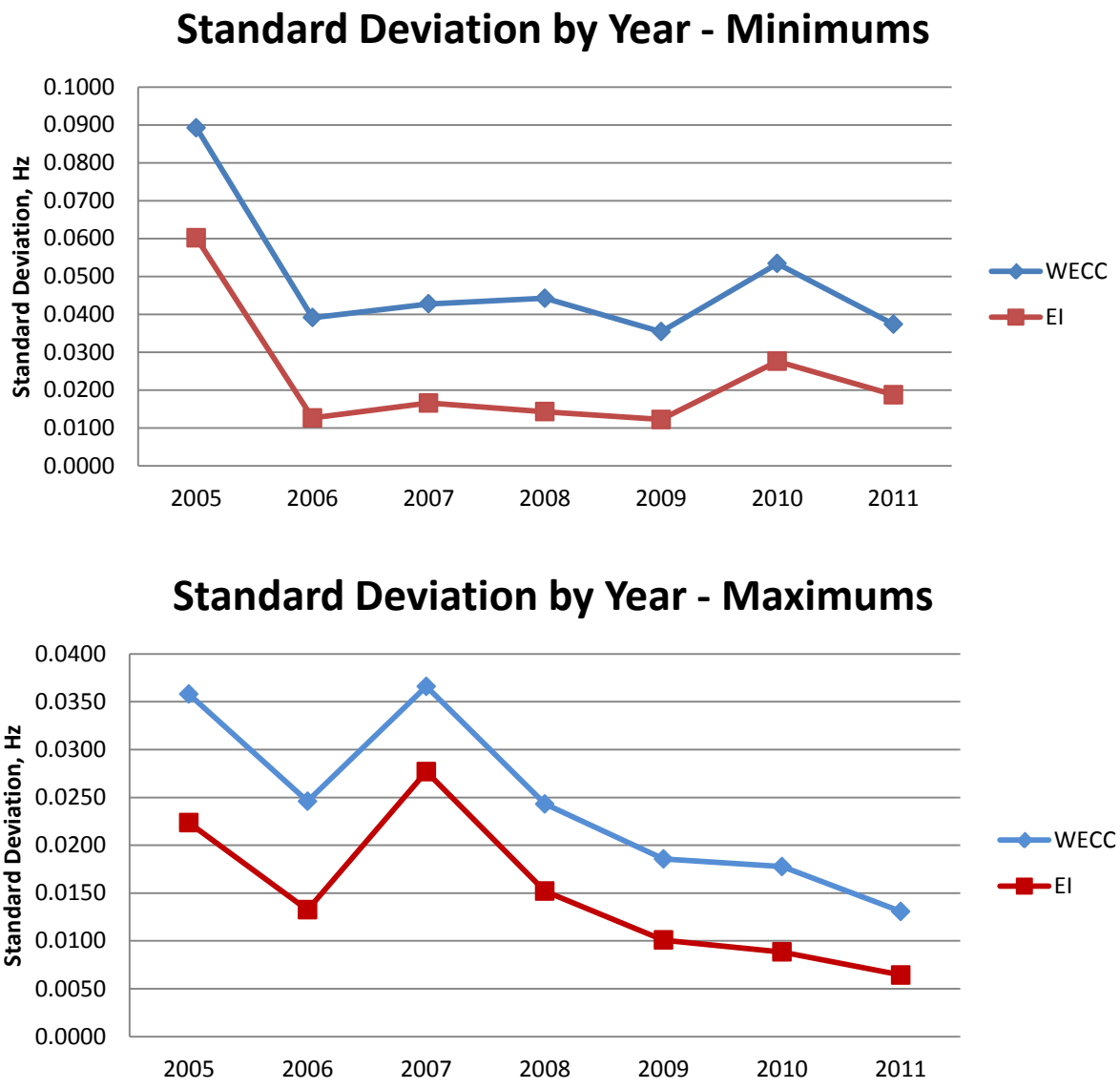


Fig. 3.17: Standard deviations of extrema for the Eastern and Western Interconnections

Conclusions

Using an algorithm developed to detect and validate the extrema, EI and WECC frequency data from the past seven years were analyzed. The results of this analysis were then studied to make comparisons from year to year. On an hourly basis, the frequency minimums in the EI follow a roughly bimodal distribution in most of the cases with two peaks often appearing around 03:00 and 10:00 UTC. The maximums show a similar pattern with the addition of one small peak at the end of the day. One explanation for the first peak (3:00 UTC) is as follows. This time typically corresponds to 22:00 Eastern

Time, which is when a large number of peaker plants are taken offline at the end of each day. The loss of these plants removes some amount of inertia from the system, which would necessarily weaken the frequency response and result in larger variations in frequency. Other sources [49, 50] suggest that the extrema result from mismatches created when generators ramp up and down during these time intervals at rates much faster than that of the load. It is likely a combination of these factors that results in the regular patterns observed here.

The hourly distribution of minimums within the WECC does not seem to show an obvious pattern like in the EI, though the maximums do tend to have a tri-modal distribution as in the EI. It appears that some of the extrema can indeed be attributed to generator trip and load shedding/pumped storage disconnection events in these interconnections. Other extrema that occur at regular intervals can most likely be explained by operational factors such as those described above.

The average magnitudes of the frequency extrema were also studied with respect to the day of the week and week of the year in which they occurred to investigate possible correlations. From the plots, there does not appear to be a definite relationship between these variables in either interconnection.

Analysis of the yearly extrema averages for the EI and the WECC shows that minimums have increased since 2007 in both interconnections and arrived at the same value for 2010 and 2011. Interestingly, the maximums of the two interconnections also merge in 2010 and 2011 after fluctuating in different ways from 2007 to 2009. The standard deviations of the extrema in both interconnections have continued to follow similar patterns, a trend that has been observed since 2005.

Frequency measurement data collected by FNET allows the historical behavior of frequency extrema within a power system to be studied. These extrema can indicate when the system is operating furthest from its nominal frequency, and provide insight into how the power grid operates across a variety of time scales. This information could prove useful to utilities and system operators, as well as market designers and regulators.

4. Artificial Neural Network-Based Classifier for Power System Events

Introduction

The Frequency Monitoring Network (FNET) receives data from more than 100 Frequency Disturbance Recorders located in power grids around the world. These data are processed in real time by the FNET application server, which includes a variety of different situational awareness modules. Event detection and classification are perhaps the two most important applications of the FNET system. Currently, the FNET event classification program relies upon empirically-derived models of system behavior to categorize different disturbances. While these perform reasonably well, they are somewhat difficult to develop since each trigger must be manually "tuned" for each event type and power grid. It is therefore desirable to obtain a classifier that eliminates this trial-and-error-based configuration.

Recent interest in the application of artificial neural networks (ANN) to phasor measurement data has resulted in several potential uses being discovered, including voltage stability assessment, protection, and transient stability prediction [27, 51, 52]. This chapter presents several ANN-based classifiers for power system disturbances that can reliably identify events in multiple interconnections. The ANN is chosen because neural networks can perform pattern recognition even when the input data is noisy or incomplete, which is sometimes the case with FNET frequency measurements [53, 54]. Beginning with an introduction to power system events and their frequency signatures, this chapter describes the existing triggering system and its deficiencies. It then discusses the process of selecting training cases, extracting the data, training, and testing the neural networks. Finally, the minimum signal length required to reliably classify a disturbance using a particular ANN-based classifier is determined.

Power System Events

At its most basic level, an electric power system is composed of generators, transformers, transmission lines, and loads (e.g., motors, computers, televisions, etc.). The power system can be thought of as an extremely large machine with hundreds of thousands of parts, all working in synchronism. In the United States, the power system is divided into the Eastern (EI), Western (WECC), and Texas (ERCOT) interconnections, which are not synchronized with each other.

As with any machine, these parts occasionally experience failures or malfunctions. Even normal operations sometimes result in unusual behaviors. Because the system is designed to be resilient, most of these phenomena go unnoticed by consumers. In some cases, however, they can lead to widespread power outages ("blackouts") or voltage problems ("brownouts").

Power system disturbances can be classified into several different categories, which are described below.

Generation Trip

Mismatches between generation and load result in changes in the system frequency, which is a consequence of Newton's 2nd Law of Motion for rotating masses. In power systems, this relationship is governed by the swing equation:

$$\frac{2H}{\omega_0}\frac{d\omega(t)}{dt} = P_{mech} - P_{elec}$$

Here, H is the normalized inertia constant in seconds, ω_0 is the nominal system frequency in rad/s, P_{mech} is the per-unit mechanical power applied to the generator by the prime mover, and P_{elec} is the per-unit electrical load being supplied. As a consequence of the swing equation, when generation exceeds load, generators will accelerate. Conversely, a surplus of load causes the generators to slow down. Normally, the mismatch is small enough that control systems are able to keep the machines very close to their nominal speed. This is done by modulating the amount of mechanical power applied to the generator.

In some cases, generators must be removed from the system very quickly, usually for their own protection. For a large generator, this can lead to a dramatic decline in frequency since it takes several minutes for additional generators to come online and make up for the lost capacity. Examples of generator trips in the Eastern (EI) and Western (WECC) Interconnections are shown in Fig. 4.1 (a) and (b), respectively.

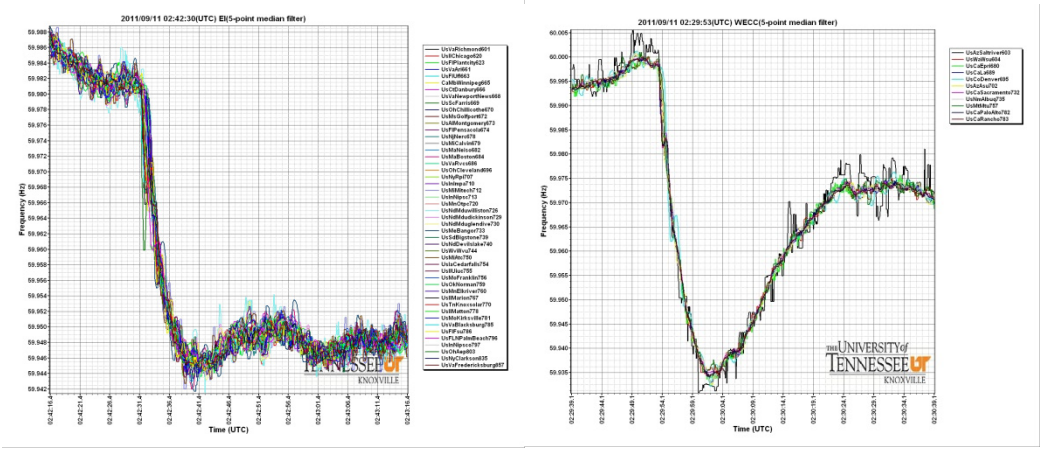


Fig. 4.1: (a) Frequency drop after generator trip in the Eastern Interconnection, (b) Frequency drop after generator trip in the Western Interconnection.

From these plots, we can observe that while the frequency declines rapidly in both cases, it recovers more slowly in the EI. This is due to the topology of the system, the amount of inertia, and the

control schemes being used by system operators. Thus, generation trips result in similar, though slightly different frequency signatures from one interconnection to another.

Load Shedding/Pumped Storage Disconnection

If the system frequency declines too much, utilities may resort to load shedding, which is the temporary disconnection of large numbers of customers. Additionally, disconnection of pumped storage power plants can also appear as load shedding. Rapid removal of load increases the system frequency since the generators are producing more power than the system can consume. Examples of load shedding events for the EI and WECC are shown in Fig. 4.2 (a) and (b), respectively.

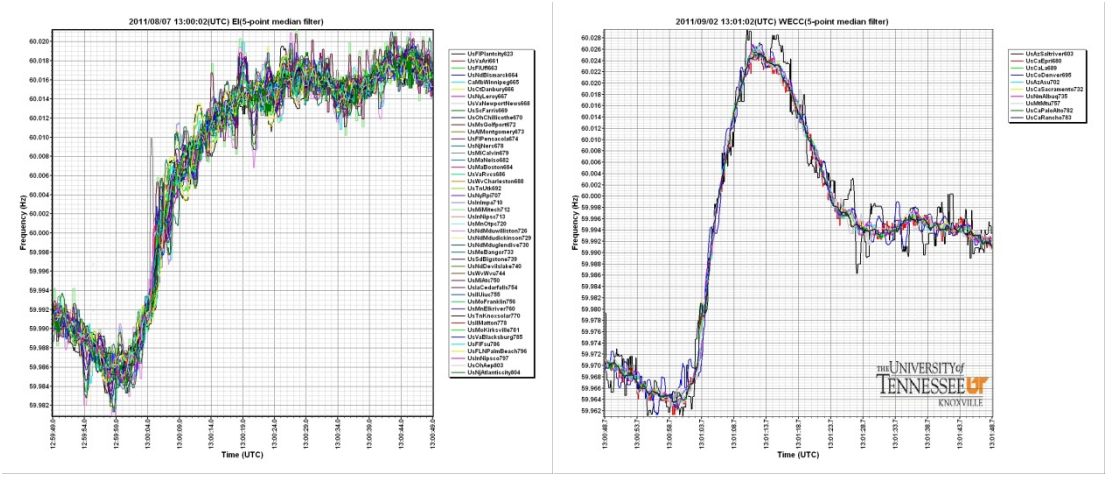


Fig. 4.2: (a) Frequency increase after load shedding in the Eastern Interconnection, (b) Frequency increase after load shedding in the Western Interconnection.

Again, while the general trend is similar, the two interconnections respond differently after the initial frequency increase.

Oscillation

Generator trips, faults, and load shedding events can result in system oscillations as the generators try to redistribute power amongst themselves, although it should be pointed out that some oscillations have no clear initiating event. If these oscillations are not properly damped, blackouts and islanding can occur as parts of the system disconnect to avoid damage. An example of an oscillation is shown in Fig. 4.3.

Line Trip

Transmission line switching operations create unique transients in the frequency measurements. These transients are typically not observed system-wide; perhaps only one or two

nearby FDRs will “see” this type of event when it occurs. Currently, FNET does not automatically detect line trips.

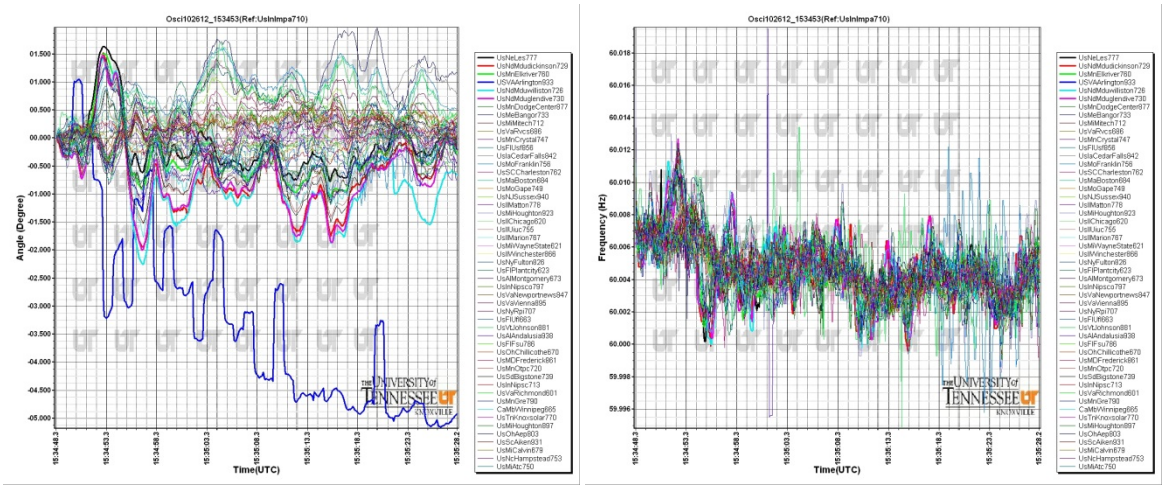


Fig. 4.3: (a) Oscillation in phase angle, (b) Oscillation in frequency.

Description of Existing Triggering Mechanism

FDR data are transmitted over the Internet to the main FNET server, which time-aligns and archives the measurements. It then forwards the data to the application server, which hosts the programs that detect, classify, and triangulate disturbances in near real-time.

Event triggers are implemented for each monitored interconnection individually using conditional logic. For example, to detect a generation trip or load shedding event, the trigger will examine the time derivative of the frequency and compare it to a known threshold value. If the derivative exceeds this threshold (and a sufficient number of FDRs agree), an event is declared. Oscillation detection is considerably more difficult to implement in this fashion because oscillations vary widely in their frequency signature. Also, some events are concurrent (e.g., line trip during oscillation) or happen in quick succession (e.g., line trip leading to oscillation).

Although the source code for the triggers is virtually the same for each interconnection, there exist some differences due to the nature of the system being studied. This is because different interconnections have slightly different responses to similar events. Thus, each additional trigger adds a significant amount (≈400 lines) of code that must be maintained, a factor that has limited the number of interconnections that the FNET system can monitor.

Research Tasks

Successful training of a neural network usually requires large numbers of previously classified example cases. Fortunately, FNET has been detecting disturbances since 2006, and approximately 9,800

generation trip/load shedding events and about 8,000 oscillation events have been observed since that time. The vast majority of these disturbances are not independently confirmed by utilities or other sources; however, no generator trip or load shedding misclassifications were found during visual inspection of hundreds of collected disturbance plots. It should also be noted that the FNET trigger ignores disturbances of less than 400 MW, 320 MW, and 300 MW in the EI, WECC, and ERCOT regions, respectively.

Event metadata are stored in a database, which facilitates the process of locating suitable training cases. Line trips presented a different challenge since there is currently no trigger for them, and thus no corresponding FNET data. This section describes the process of selecting event cases, extracting (or simulating) their data, performing preprocessing of the data, and training the neural network.

Event Search Interface

Event metadata such as size, location, and time are stored in a MySQL database that resides on the FNET application server. Until the start of this work, there was no way to quickly conduct refined searches on the event database and view the corresponding event plots, much less export the metadata for later use. For this research, a PHP-based web search interface was developed for the generation trip/load shedding database (Fig. 4.4). This interface allows users to narrow their searches by date range, interconnection, and event type. Plots can be viewed by clicking the 'View' button for each event. The table of results can be downloaded in comma-separated value (CSV) format, which is then used by the data extraction tool.

The oscillation event metadata is currently stored in a separate table. Due to technical limitations, it could not be integrated into the load shedding/generation trip table. This did not present significant difficulty, since the same information could still be extracted using SQL queries and saved in CSV format.

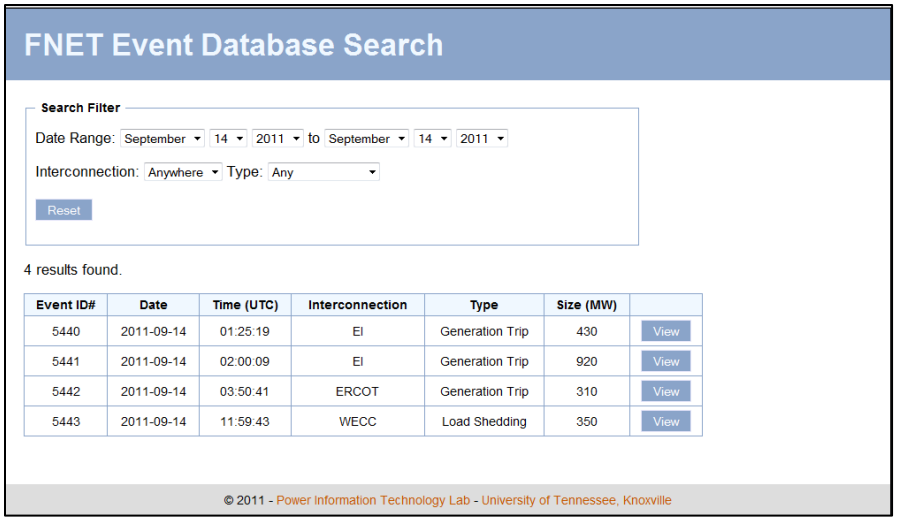


Fig. 4.4: FNET Event Database Search tool interface.

Simulation of Line Trip Data

Because FNET does not currently detect line trips, there was no database of events or corresponding data for this event type. PSS/E, a transmission planning tool, was used to perform time-domain simulations of line trips.

Dynamic simulation in PSS/E is a multi-step process that is quite time-consuming if done manually. Fortunately, a Python API exists for PSS/E that allows this process to be automated. The process can be briefly summarized as follows [55]:

1. Load the network data
2. Convert generators and loads to their Norton equivalents and constant-admittance equivalents, respectively
3. Load the dynamic models
4. Select output channels (measurements) to be recorded
5. Perform simulation for a brief steady-state period before the event
6. Apply the disturbance
7. Perform simulation for some time period after the event

For this project, a 16,000-bus model of the Eastern Interconnection was used for the simulations. Approximately 75 buses corresponding to actual FDRs were selected as the measurement points, and lines adjacent to these buses were tripped one at a time. A 20-second simulation was performed in each case, with measurement points being saved at 0.1-second intervals (to match the reporting rate used by FDRs). The frequency data were saved in PSS/E's proprietary binary output format. Since this format cannot be read by MATLAB, a Python script was written that converted each

file into a MAT file. (It should be pointed out that the extraction process in the script was handled by a library produced by Siemens, the developers of PSS/E.) An example of a simulated line trip is shown in Fig. 4.5.

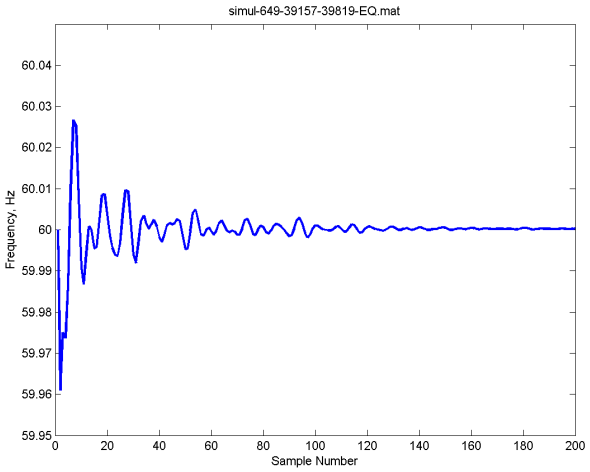


Fig. 4.5: Simulated line trip event

Automated Extraction Tool

FNET data are presently stored in Microsoft Access MDB files. Originally, an entire day's data could fit into a single MDB file, however the addition of more FDRs has made this impossible. Thus, each day's data is now spread across three different files. Locating and extracting the data given the timestamp is nontrivial because the database files are spread across multiple filesystems and use a variety of naming conventions.

A MATLAB-based extraction tool was developed that accepts the CSV file produced by the search interface and extracts the corresponding FDR data for each event from the MDB files into MAT files. The extracted data are kept at full resolution (10 samples per second) so that the preprocessing tool can downsample as needed. Each file is saved in a directory structure that indicates what type of event it represents and the interconnection where it took place. This structure makes it easier for the preprocessor to segregate the training data for each class.

Preprocessor

MATLAB's ANN toolbox requires that the training data be stored in an $m \times n$ matrix where m is the number of input nodes, and n is the number of training cases. The classifications for the training data are stored in a $k \times n$ matrix, where k is the number of classes and n is again the number of training cases. MATLAB refers to this matrix as the target matrix. In the target matrix, each column has a '1' in the row corresponding to the class for that case, while the other rows have zeroes.

A MATLAB-based preprocessor was written to traverse the directory structure produced by the extraction tool and create the training and target matrixes, which are then saved. The preprocessor can also downsample the training data as needed to reduce the number of input nodes. Because the preprocessor relies upon the naming conventions used in the directory structure to delineate the classes, a particular class can be easily excluded or included in the training matrix by simply renaming its corresponding folder. This feature is quite useful for testing different configurations of networks.

Training Case Selection

Significant time was devoted to the selection of appropriate training cases. Because the grid behaves differently from season to season and from year to year, cases were selected to cover these time frames wherever possible [45, 48]. The simulated line trip data did not reflect any seasonal variation. Since line trips are usually a local event whose response does not depend on the system as a whole, this should not introduce significant error. It should be noted that there were too few oscillation cases in the ERCOT power system that could be used for training data, so this type of event was excluded from the classifier. The breakdown of training cases is given in Table 4.1.

Table 4.1: Breakdown of training case event types.

	EI	WECC	ERCOT
Generation Trip	547	415	189
Load Shedding	160	346	95
Oscillation	333	392	0
Line Trip	257	0	0

Initial testing of the ANN using simulated line trip data resulted in poor performance. Upon further examination, it was found that several of the cases did not show much, if any, variation in frequency during the simulation period. These cases were subsequently removed from the training set. Similarly, some FDR data recorded during oscillation events did not appear to show any oscillatory behavior – these cases were also removed. Based upon knowledge of previous oscillations, data were then extracted from FDRs in regions in the Eastern Interconnection that typically oscillate against one another [5], which seemed to yield better results. WECC oscillation and ERCOT load shedding cases were initially excluded from the classifier because of a lack of sufficient data, but were later added once more data became available.

Network Configuration

The next task was to determine what level of downsampling, if any, should be used on the input data. At first glance, it might be tempting to use all 200 points, however this approach is problematic. A large number of input nodes requires more training time and also introduces additional noise into the network. Both of these issues can be solved by downsampling, which essentially low-pass filters the input signal and reduces the number of points. Repeated training/testing runs with downsampling factors ranging from one (every point) to ten (every tenth point) showed that the best performance (lowest overall confusion) resulted when every fourth data point was used. As a result, this downsampling factor was applied when forming the matrixes for testing and training.

The choice of the number of hidden nodes was somewhat more difficult. There is no hard and fast rule for how many hidden nodes are needed for a particular classification problem, although one source states that the number of hidden nodes should never be more than twice the number of input nodes [53]. In order to determine the size of the hidden layer for this problem, a MATLAB program was written that created neural networks with increasing numbers of hidden nodes. Each network with *n* hidden nodes was trained and tested ten times, and from this the average confusion was calculated. The program was run with *n* ranging from two to 70, which yielded the confusion plot shown in Fig. 4.6. Since the performance of the network appeared to roughly flatten (and then increase) after 20 nodes, this number of hidden nodes was used in subsequent tests.

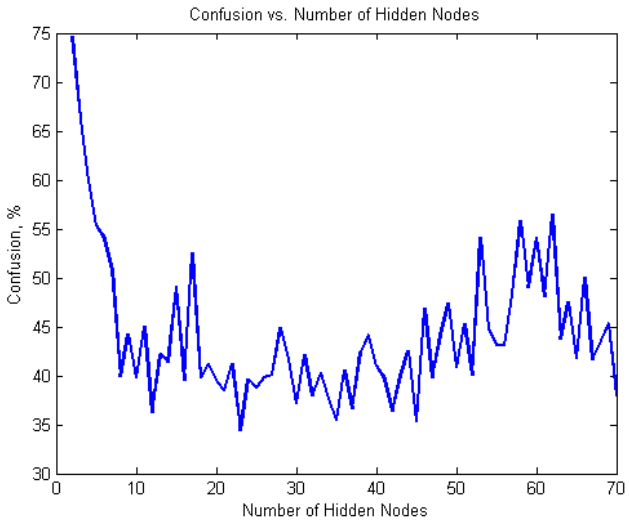


Fig. 4.6: Plot of ANN confusion vs. number of hidden nodes.

Another choice that needed to be made was the overall form of the network. That is, should it try to encompass all possible classes, or should several small networks be used instead? To answer this question, several different training networks were created, each representing a different type of

configuration. For example, one network was designed to identify all possible classes, while another contained only generator trips and load shedding events. Additionally, several simple is/isn't networks were created for each distinct disturbance type; these were later broken into interconnection-specific classifiers.

Training and Testing

Once the training and target matrixes were created, building and training the ANN was fairly straightforward. The toolbox only requires the number of hidden nodes and the desired breakdown of the cases into training, validation, and testing. For this project, the breakdown used was 70% training, 15% validation, and 15% testing. In this context, validation refers to the measurement of network generalization during the training process. If generalization stops improving, the toolbox ends the training process [56].

A MATLAB program was written to create, train, and test networks for the different configurations. Each configuration was trained and then tested against a corresponding test set of previously unseen cases randomly selected from events that occurred during a two-year period beginning in September 2009. Approximately 45 cases of each type were used in the testing set. Because performance can vary significantly from one training instance to another, the training and testing was repeated 10 times for each configuration, and the network having the lowest overall confusion was saved for future use. This process was repeated once the data for WECC oscillations and ERCOT load shedding events became available.

Results

Multi-Type/Location Classifiers

Confusion matrixes for each multi-type/location network configuration are presented in Fig. 4.7-Fig. 4.16. For each matrix, the best performing network for that particular configuration was used. The class types corresponding to each numerical identifier can be found in the tables preceding each figure. Table 4.2 summarizes the overall accuracy of each network, which is defined as the percent of disturbances correctly classified according to their major type (generator trip, load shedding, etc.) without regard for location (EI, ERCOT, WECC).

Table 4.2: Summary results for multi-type classifiers

Classifier Type	Overall Accuracy (Original)	Overall Accuracy (with new categories)
Monolithic (all types)	88.3 %	65.3 %
Generator trip/load shedding	99.1 %	99.2 %
Generator trip/load shedding/line trip	95.1 %	80.8 %
Generator trip/load shedding/oscillation	96.2 %	94.6 %
Oscillation/line trip	93.3 %	80.7 %

At first glance, the monolithic network (Fig. 4.7) containing all possible classes appears to have the worst performance (79.9% accuracy). However, examination of the confusion matrix reveals that the classifier usually determined the correct event type, even if it picked the wrong interconnection. Since the interconnection of each signal is known *a priori*, this confusion does not really matter. If the location confusion is disregarded, the real accuracy is 88.3%. This network experienced some difficulty distinguishing between EI generator trips and oscillations, as well as between line trips and oscillations, a result that is not terribly surprising since nearly all line trips exhibit some form of oscillatory behavior. However, since line trips tend to be local phenomena observed by only a few FDRs, and oscillations can be observed system-wide, this may not present a major problem if additional logic can be used to distinguish between the two. After ERCOT load shedding and WECC oscillations were added to the monolithic classifier, the overall accuracy declined to 65.3%. Inspection of the confusion matrix (Fig. 4.8) reveals that a large part of this error is caused by both EI and ERCOT load shedding events being classified incorrectly as WECC oscillation events. Additionally, ERCOT load sheddings and EI/WECC oscillations were in many cases misclassified as line trips.

By removing everything but generator trip and load shedding cases (Fig. 4.9), significantly better performance (94%) was achieved. Again, this is somewhat expected since these two events have dramatically different frequency signatures. Also, most of the misclassifications were in location, rather than event type. Ignoring the location error, the real accuracy is 99.1%. Addition of ERCOT load shedding events to this classifier (Fig. 4.10) increased the accuracy slightly to 99.2%.

Using only generator trip, load shedding, and line trip cases (Fig. 4.11), the overall performance degraded slightly (86.3%). Here, we see that the network had some trouble distinguishing between generator trips in the EI and WECC regions, and also between generator trips in the WECC and ERCOT regions. If the results for the basic event types are lumped together, the true accuracy is 95.1%. Some confusion between EI generator trips and oscillations can also be observed, which is to be expected given their similar frequency signatures. Once introduced, the ERCOT load sheddings (Fig. 4.12) were frequently misclassified as line trips, dropping the overall accuracy to 80.8%.

The next classifier was designed to distinguish between generator trips, load sheddings, and oscillations. Originally (Fig. 4.13), the network's overall performance was still quite good (92.4% accuracy). Considering that generator trips often contain oscillatory components that would presumably confuse the classifier, this result is very encouraging. As before, the classifier is nearly always correct in choosing the event type (96.2% accuracy), even if the location is incorrect. Even after WECC oscillations are added (Fig. 4.14), the accuracy is still quite high (94.6%).

Because line trips and oscillations appeared to cause confusion for the monolithic classifier, a network was created that could only distinguish between these two types of events (Fig. 4.15). Originally, this network performed quite well, correctly classifying 93.3% of the testing cases. But, the accuracy declined to 80.7% with the addition of WECC oscillations. In many cases, both types of oscillations were misclassified as line trips.

Table 4.3: Composition of training cases: all types

	Generator Trip	Line Trip	Load Shedding	Oscillation
EI	1 (547)	7 (257)	4 (160)	6 (333)
ERCOT	3 (189)			
WECC	2 (415)		5 (346)	

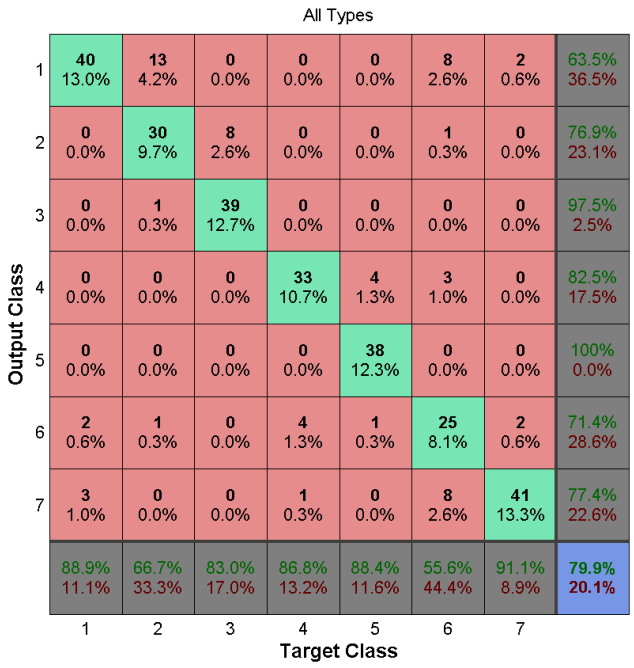


Fig. 4.7: Confusion matrix for network containing all types of events.

Table 4.4: Composition of training cases, all types (new categories)

	Generator Trip	Line Trip	Load Shedding	Oscillation
EI	1 (547)	9 (257)	4 (160)	7 (333)
ERCOT	3 (189)		6 (95)	
WECC	2 (415)		5 (346)	8 (392)

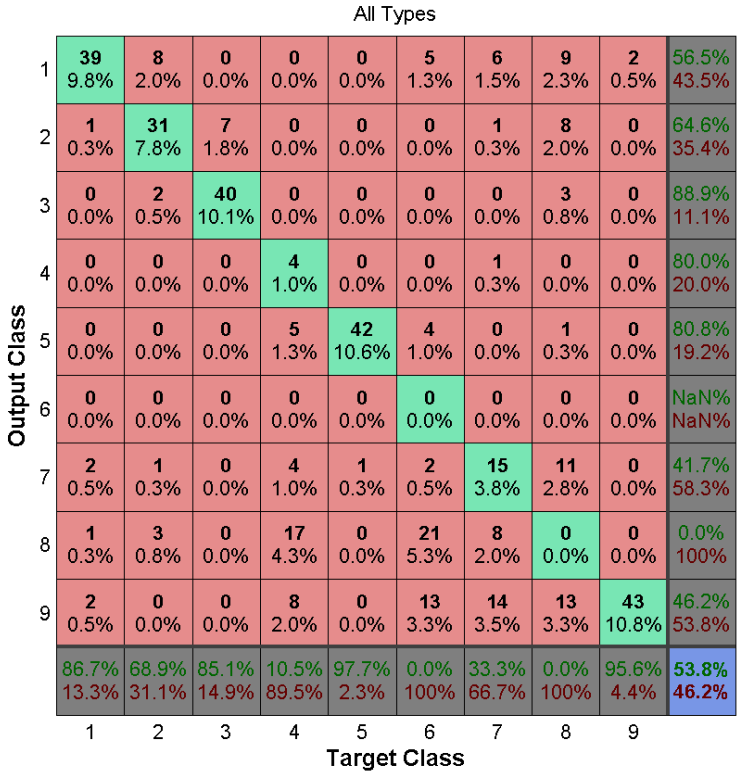


Fig. 4.8: Confusion matrix for network containing all types, including new categories

Table 4.5: Composition of training cases: generator trip/load shedding

	Generator Trip	Line Trip	Load Shedding	Oscillation
EI	1 (547)		4 (160)	
ERCOT	3 (189)			
WECC	2 (415)		5 (346)	

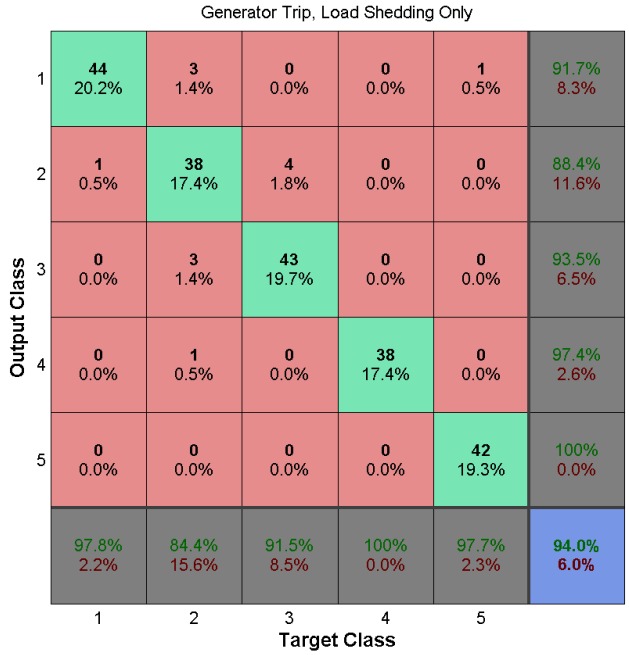


Fig. 4.9: Confusion matrix for network containing generator trip and load shedding cases only.

Table 4.6: Composition of training cases: generator trip/load shedding (new categories)

	Generator Trip	Line Trip	Load Shedding	Oscillation
EI	1 (547)		4 (160)	
ERCOT	3 (189)		6 (95)	
WECC	2 (415)		5 (346)	

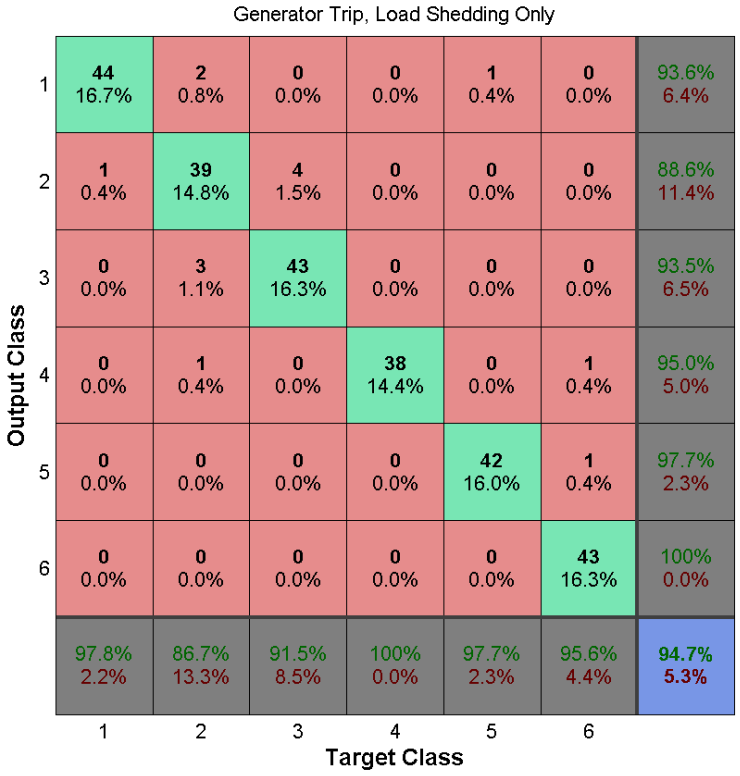


Fig. 4.10: Confusion matrix for network containing generator trips and load shedding only (new categories)

Table 4.7: Composition of training cases: generator trip/load shedding/line trip

	Generator Trip	Line Trip	Load Shedding	Oscillation
EI	1 (547)	6 (257)	4 (160)	
ERCOT	3 (189)			
WECC	2 (415)		5 (346)	

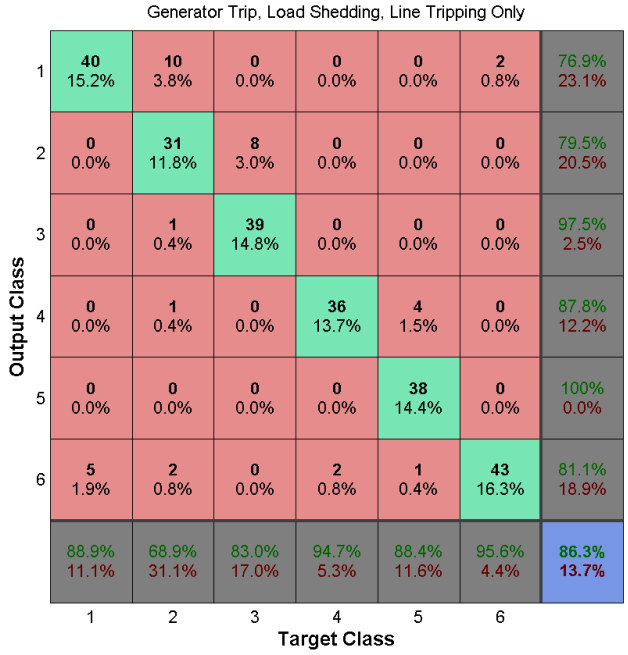


Fig. 4.11: Confusion matrix for network containing generator trip, load shedding, and line trip cases only.

Table 4.8: Composition of training cases: generator trip/load shedding/line trip (new categories)

	Generator Trip	Line Trip	Load Shedding	Oscillation
EI	1 (547)	7 (257)	4 (160)	
ERCOT	3 (189)		6 (95)	
WECC	2 (415)		5 (346)	

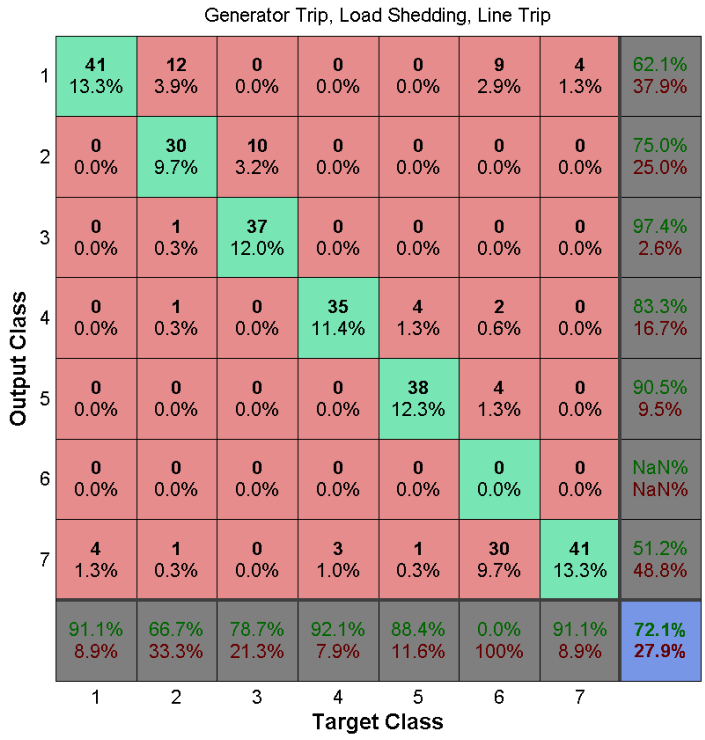


Fig. 4.12: Confusion matrix for network containing generator trip, load shedding, and line trip cases only (new categories)

Table 4.9: Composition of training cases: generator trip/load shedding/oscillation

	Generator Trip	Line Trip	Load Shedding	Oscillation
EI	1 (547)		4 (160)	6 (333)
ERCOT	3 (189)			
WECC	2 (415)		5 (346)	

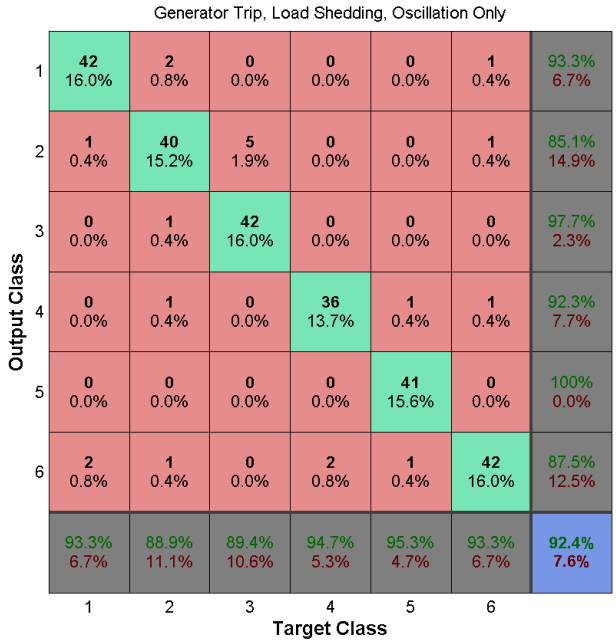


Fig. 4.13: Confusion matrix for network containing generator trip, load shedding, and oscillation cases only

Table 4.10: Composition of training cases: generator trip/load shedding/oscillation (new categories)

	Generator Trip	Line Trip	Load Shedding	Oscillation
EI	1 (547)		4 (160)	7 (333)
ERCOT	3 (189)		6 (95)	
WECC	2 (415)		5 (346)	8 (392)

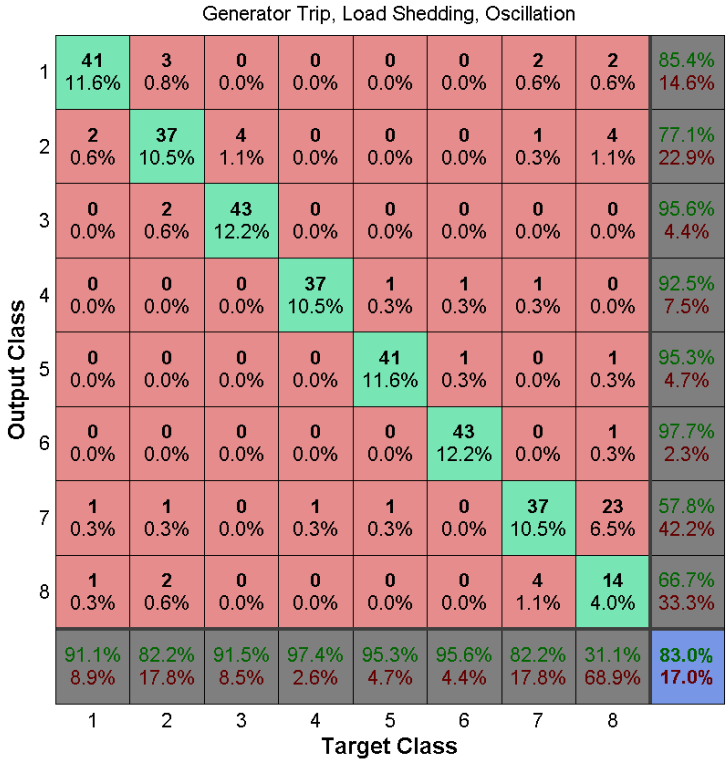


Fig. 4.14: Confusion matrix for network containing generator trip, load shedding, and oscillation cases only (new categories)

Table 4.11: Composition of training cases: line trip/oscillation

	Generator Trip	Line Trip	Load Shedding	Oscillation
EI		2 (257)		1 (333)
ERCOT				
WECC				

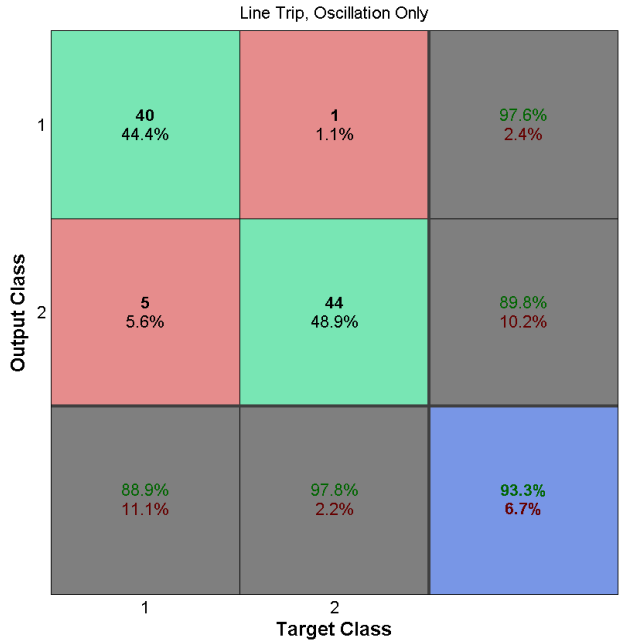


Fig. 4.15: Confusion matrix for network containing line trip and oscillation cases only

Table 4.12: Composition of training cases: line trip/oscillation (new categories)

	Generator Trip	Line Trip	Load Shedding	Oscillation
EI		1 (257)		2 (333)
ERCOT				
WECC				3 (392)

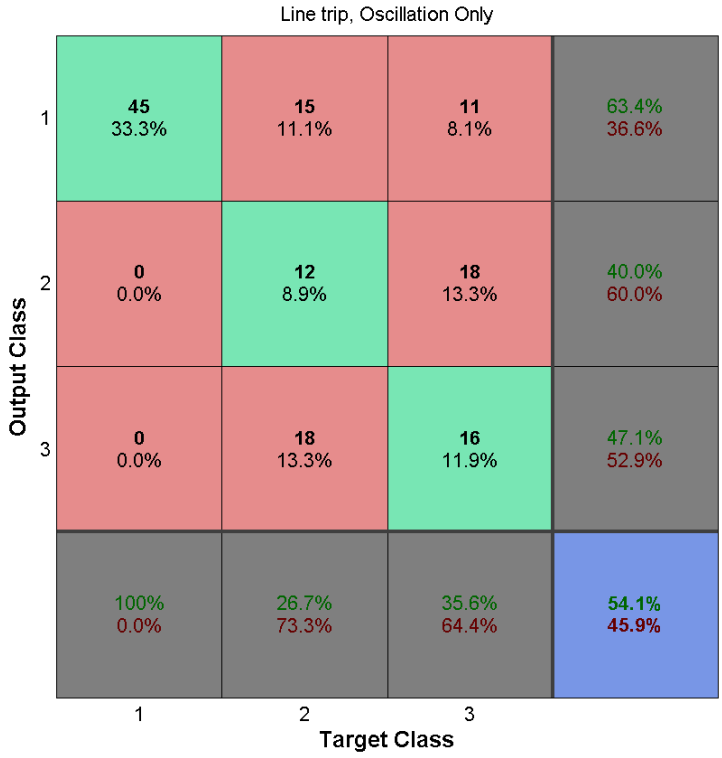


Fig. 4.16: Confusion matrix for network containing line trip and oscillation cases only (new categories)

Simple Classifiers

Confusion matrixes for the simple (yes/no) networks are shown in Fig. 4.17–Fig. 4.20. In the confusion matrixes, class '1' indicates that the case is of the event type, while class '2' indicates that it is not. For these classifiers, the training cases for different interconnections were lumped together to form the training and testing sets. The overall and average accuracies for each classifier are given in Table 4.13. Since the testing sets contained more 'is not' examples than 'is' examples, the overall accuracy is weighted accordingly; here, the unweighted average accuracy is used to provide a fair comparison of the results.

Table 4.13: Summary results for simple classifiers

Classifier Type	Overall Accuracy	Average Accuracy
Generator trip	92.2 %	92.4 %
Load shedding	87.2 %	80.9 %
Line trip	93.2 %	89.4 %
Oscillation	78.1 %	52.0 %

Although the accuracy of the generator and line trip classifiers were similar to that of the multi-type classifiers discussed earlier, the load shedding and oscillation classifiers performed noticeably worse. In particular, the oscillation classifier provided little better accuracy than random chance. One possible reason for this might be that the 'not' cases used to train the network contained the remaining disturbance types, all of which can sometimes include a noticeable oscillatory component. In contrast, the oscillation case examples contained a wide variety of oscillations in terms of magnitude, frequency, and duration. It seems plausible that the network could not adequately generalize the positive cases, an explanation which is supported by the 99.7% negative case classification accuracy for this network.

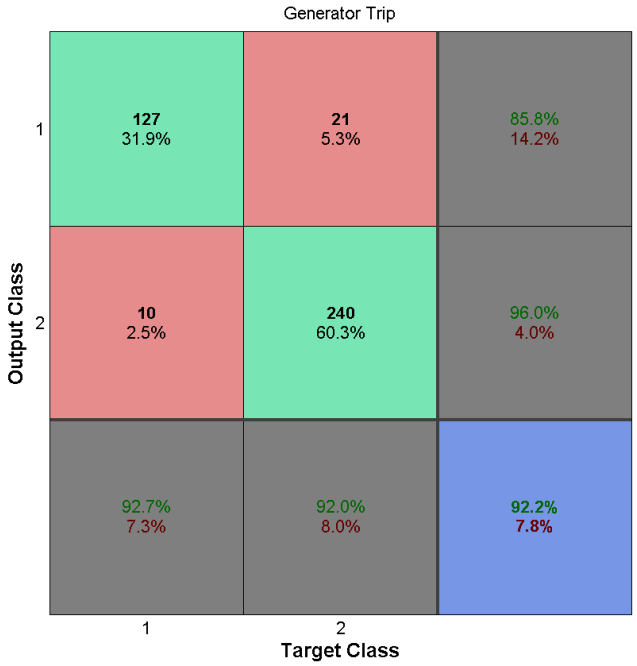


Fig. 4.17: Confusion matrix for generator trip classifier

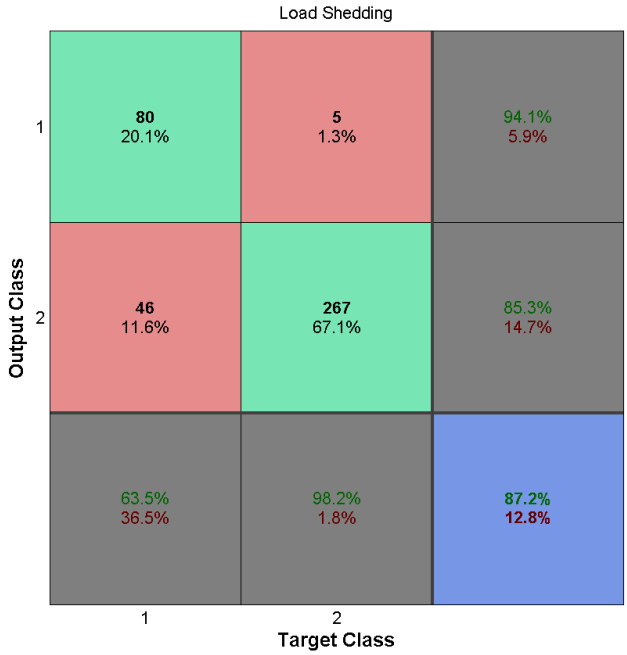


Fig. 4.18: Confusion matrix for load shedding classifier

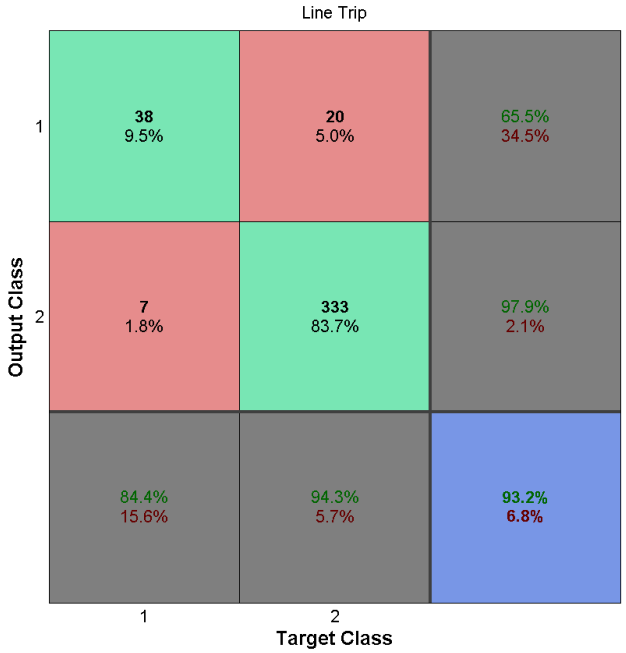


Fig. 4.19: Confusion matrix for line trip classifier

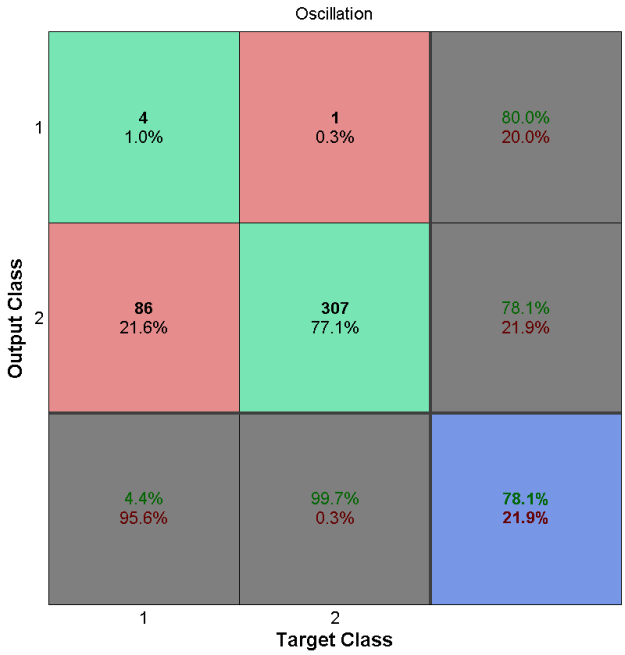


Fig. 4.20: Confusion matrix for oscillation classifier

Interconnection-specific Classifiers

Classification networks presented in the previous sections included training cases from several different interconnections. Although it may be convenient to have a single network that can be applied to examples from different power grids, the results indicate that doing so can introduce confusion and decrease the overall accuracy. This suggests that a better approach might be to develop networks for each individual interconnection.

Table 4.14 and Fig. 4.21 show the numerical class identifiers and confusion matrix, respectively, for an Eastern Interconnection monolithic classifier. This network routinely identifies generator trips, line trips, and load shedding events correctly while frequently misclassifying oscillations. For the first three types of disturbances, the accuracy was greater than or equal to 89.5%, while for oscillations it was correct only 46.7% of the time. Since oscillations were the greatest source of inaccuracy in the classifier, these were removed and a generator trip/load shedding/line trip network was created, which was 97.7% accurate (Fig. 4.22).

To see if better accuracy could be obtained by separating out each event type into its own network, several yes/no classifiers were created using the EI training data; their confusion matrixes are shown in Fig. 4.23 - Fig. 4.26. The results here were somewhat surprising. Although the load sheddings

were categorized correctly nearly 95% of the time with the simple network (Fig. 4.24) compared to 89.5% for the monolithic one, the simple network was slightly less accurate (88.9% vs. 91.1%) in identifying generator trips (Fig. 4.23). Both networks performed equally well in identifying line trips (Fig. 4.25). Most interestingly, the simple oscillation network (Fig. 4.26) misclassified all of the positive examples despite several attempts at retraining. There are two likely explanations for this. First, there is great diversity in the characteristics of EI oscillations, which would make them difficult for the ANN to generalize. Second, it is possible that some of the oscillation cases were coincident with other types of disturbances.

Table 4.14: Composition of training cases: monolithic EI classifier

Generator Trip	Line Trip	Load Shedding	Oscillation
1 (547)	3 (257)	2 (160)	4 (333)

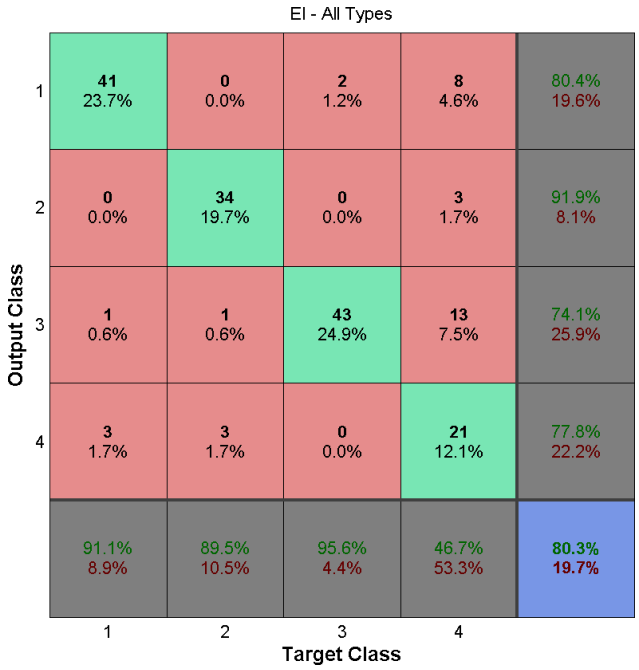


Fig. 4.21: Confusion matrix for EI monolithic classifier

Table 4.15: Composition of training cases for EI generator trip/load shedding/line trip classifier

Generator Trip	Line Trip	Load Shedding
1 (547)	3 (257)	2 (160)

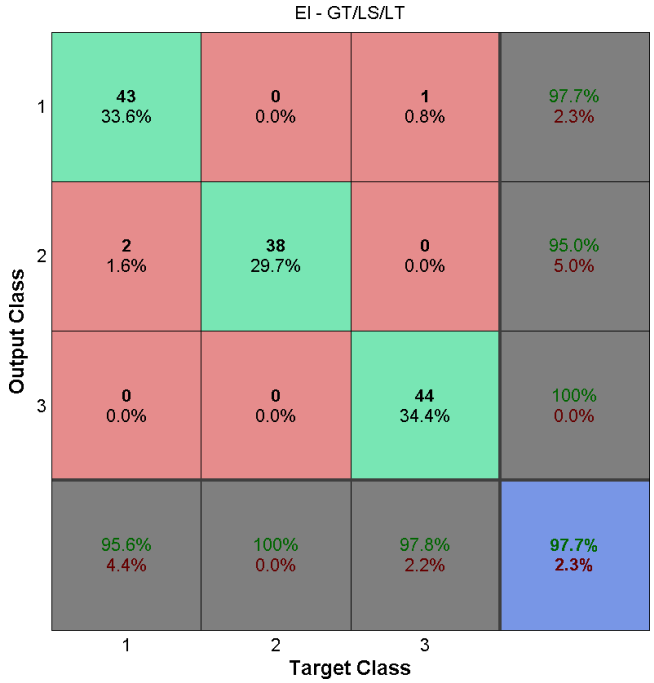


Fig. 4.22: Confusion matrix for EI generator trip/load shedding/line trip classifier

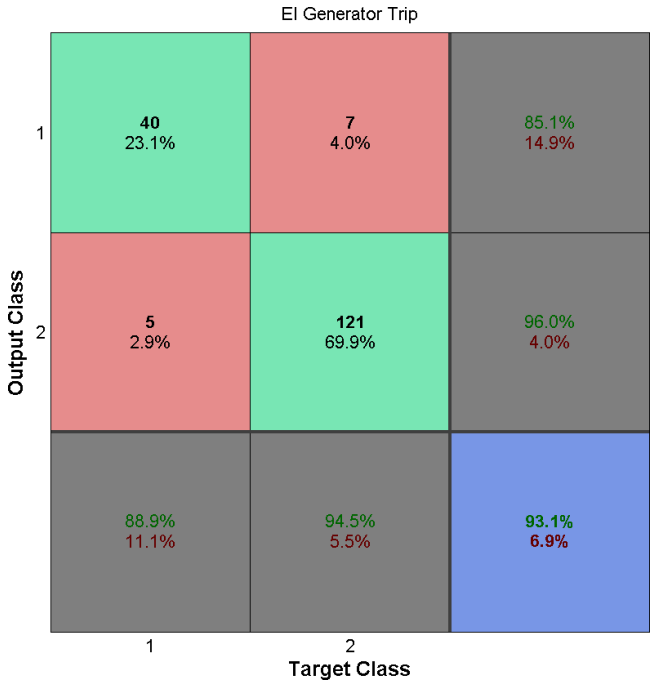


Fig. 4.23: Confusion matrix for EI generator trip classifier

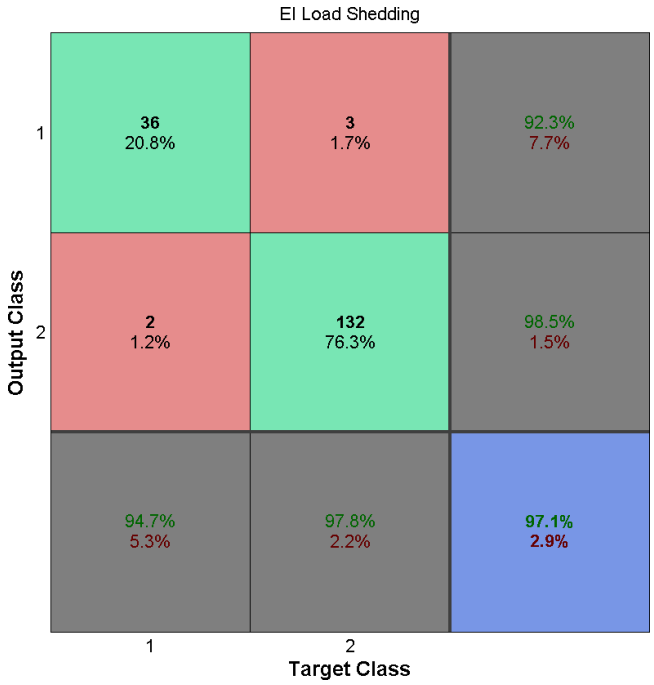


Fig. 4.24: Confusion matrix for EI load shedding classifier

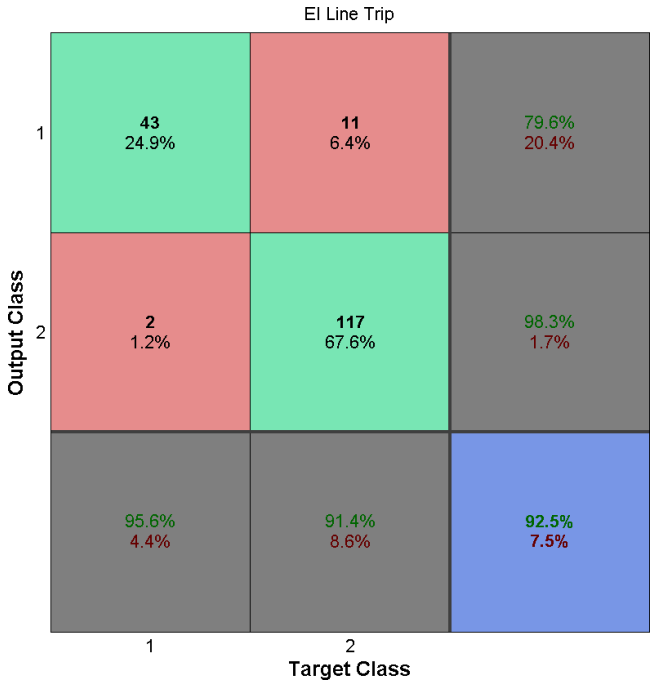


Fig. 4.25: Confusion matrix for EI line trip classifier

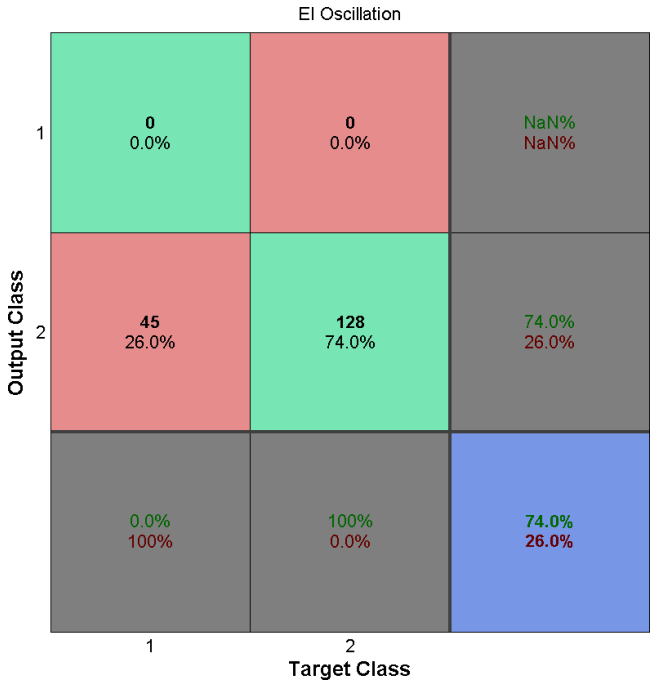


Fig. 4.26: Confusion matrix for EI oscillation classifier

Minimum Required Signal Length

The artificial neural networks developed in this research were designed to operate on 20-second segments of unfiltered FNET frequency data downsampled to every fourth point. This time interval is relatively long compared to the duration of an actual event. Since disturbances should ideally be

classified as quickly as possible, this raises the question: is there some minimum signal length required for the classifier to reliably categorize each disturbance? To answer this, a program was written to create and train networks using input signals with lengths ranging from 0.4 seconds to 10 seconds and downsampling factors of 1, 2, and 4. Here, the EI generator trip/load shedding/line trip classifier described in the previous section was used. Because the MATLAB ANN toolbox randomly selects the example and testing cases from the training set, each network configuration was created, trained, and tested against an independent testing set 10 times. The lowest confusion rate for each configuration was then recorded and plotted in Fig. 4.27. Assuming a desired confusion rate of less than 5%, the minimum signal length appears to be about three seconds, regardless of the downsampling factor. But, since each example case contains approximately one second of pre-event data, the true answer is in fact closer to two seconds, which is much less than the 20 seconds currently required by the FNET server application.

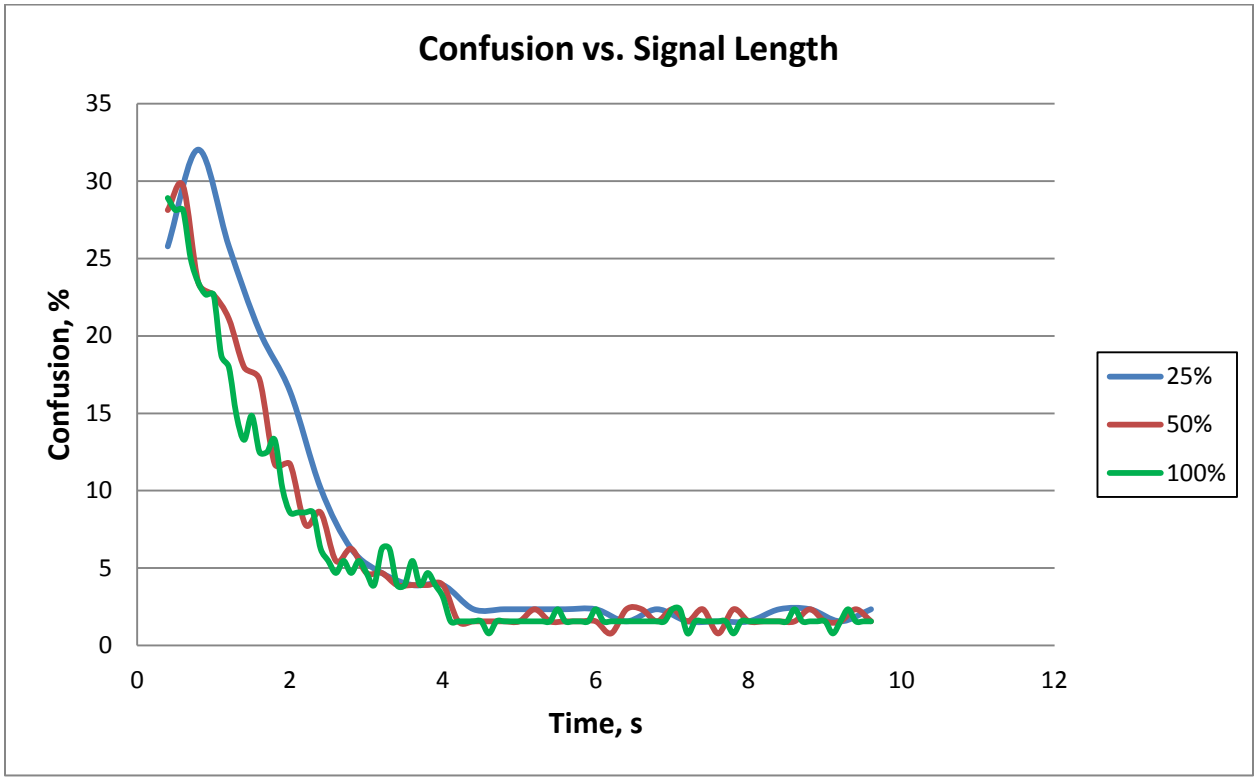


Fig. 4.27: Plot of Confusion vs. signal length for ANN-based classifier

Conclusions

Development of automated classification techniques for various power system disturbances based upon their frequency signatures has traditionally been a tedious, time-consuming task. Conditional logic-based classifiers are difficult to create and maintain due to the slight differences

between individual events of the same type. Thus, a classifier that could be trained using examples of previous events would save both time and effort, and possibly provide better accuracy.

In this chapter, a variety of different neural network-based classifiers were created and tested using data from both PSS/E simulations and FDR measurements associated with disturbances identified by the existing FNET logic-based classifier. First, event data from different interconnections were used to train multi-disturbance networks of varying complexity. Results showed that the networks were generally quite accurate (>90%) in terms of choosing the correct event type. Still, oscillation cases from the Eastern Interconnection tended to be misclassified. Although the networks sometimes chose the wrong interconnection for a given category, this was inconsequential since the interconnection for a particular FDR is already known. As with all the classifiers developed in this work, the accuracy is measured with respect to that of the FNET classifier used to produce the training and testing sets. The results presented here assume that the existing classifier is highly accurate, which seems reasonable given the dearth of misclassifications observed in the FNET event database. Ideally, a list of independently confirmed events would be used for creating the training and testing sets; such a list is currently being developed.

Simple interconnection-agnostic networks for each major disturbance type were then created using data from the three U.S. power grids. Although generator and line trips were accurately categorized with this method, the remaining event types were often misclassified. To further explore the source of the observed classification errors, ANNs using data from only the Eastern Interconnection were created and tested. A monolithic network designed to distinguish between the four event types performed reasonably well, but experienced difficulty with the oscillation cases. Once these were removed, 97.7% accuracy was obtained. Subsequent attempts to create simple yes/no networks for each event type did not result in significantly better accuracy. Thus, location-specific multi-event networks appear to provide the best overall classification performance for this type of problem. Traditional modal analysis techniques like the Prony or matrix pencil methods could then be used to determine if oscillations are present in the disturbance signature.

In the final part of this work, different input signal lengths and downsampling factors were studied to determine how much data is actually necessary for a particular classifier to reliably obtain accurate results. It was found that the minimum signal length was approximately two seconds, regardless of the downsampling rate. This represents a major improvement over classification techniques currently employed by the FNET system.

5. Electromechanical Speed Map Development using FNET Measurements

Introduction

Differences in electromechanical wave propagation speeds have long been observed in different regions of the power grid [39]. Various attempts have been made over the years to develop "speed maps" of the grid, either through analytical means or from measurements. Analytical techniques are computationally difficult and require that an accurate model of the system be known *a priori*, which is generally impossible for a large system such as the Eastern Interconnection. Most, if not all, of measurement-based techniques introduced to-date assume constant propagation speed between the disturbance source and the sensor, which does not reflect reality. This chapter proposes a new method for developing speed maps using FNET data from confirmed disturbances that seeks to eliminate many of the issues faced by previous attempts.

Algorithm

Consider a power system disturbance resulting in an electromechanical transient (either frequency or angle) that propagates throughout the system in all directions (Fig. 5.1). Assuming the grid in question has a time-synchronized wide-area measurement system, this disturbance will be detected at different times by each sensor (Fig. 5.2). In general, though certainly not always, those sensors (which could be PMUs or FDRs) that are closest to the disturbance will detect it first, while those farthest away will observe the resulting transient slightly later.

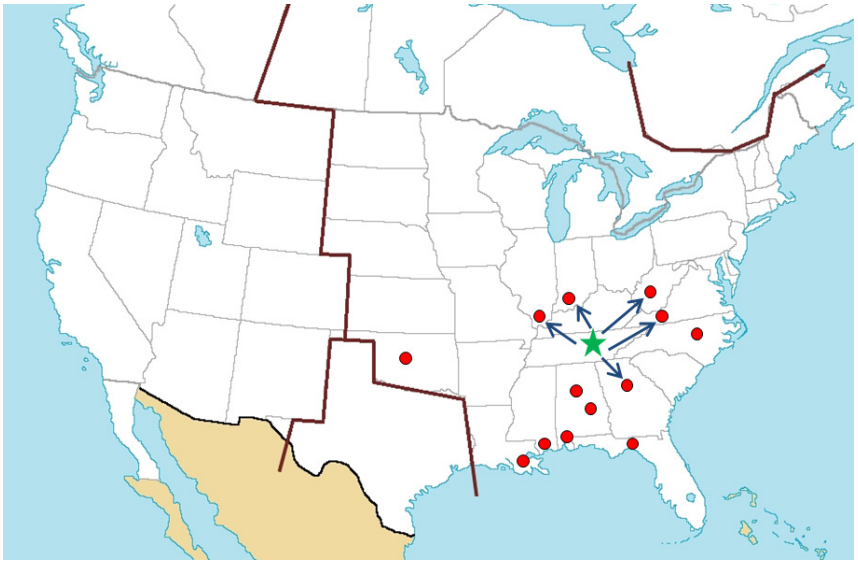


Fig. 5.1: Illustration of disturbance propagation.

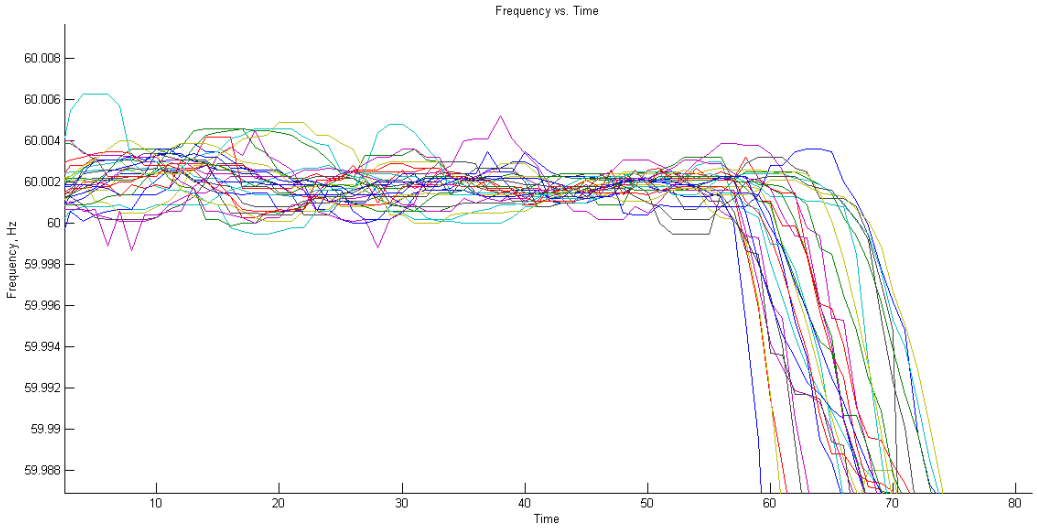


Fig. 5.2: Plot of frequency vs. time for a generator trip.

Since the locations of the sensors are known, it may be tempting to simply divide the distance by the time delay between the sensor and the event, and use this number as the wave speed. However, this approach tends to include regions of the system in between the disturbance and the sensor, which could be problematic if the distance is large and contains areas with different amounts of inertia and impedance. Additionally, the exact time of the event may not be known with great certainty, which further skews the result. Any method of estimating reasonably accurate phase velocities must address both of these concerns. One way of doing this would be to only consider the propagations speeds between sensors that are fairly close together. This would give less weight to inertially heterogeneous regions and eliminate the timing issue since the sensors themselves provide highly accurate timestamps that are independent of the event itself. Previous investigations [35, 36, 40] have revealed that different regions in a power system can exhibit significantly varying wave velocities depending on their relative position to the originating disturbance and the direction of propagation. However, no attempt was made to normalize the results, which could possibly reveal relative differences between regions that remain consistent over time. A new algorithm for calculating regional phase speeds could be summarized as follows:

1. Compute a distance matrix D, where entry $d_{i,j}$ contains the distance between sensors i and j.
2. Determine the time delay of arrival for each sensor S by observing when its frequency (or angle) passes a particular threshold.

3. For each sensor *S*, search the distance matrix to generate a list of neighboring sensors {N_1, N_2, N_n…} that are within some pre-determined radius. Here, a search radius of 300 miles was used.
4. For each neighboring sensor, compute the propagation speed from *S* to N_i using the time delays of arrival found in Step 2.
5. Estimate the propagation speed for *S* by averaging the velocities found in Step 4.
6. If desired, normalize the speeds for each measurement location by the largest observed speed.

Testing Cases

Fifteen Eastern Interconnection generator trip cases were selected from the FNET event database. These particular cases were chosen because a) they occurred within the past two years; and b) their source could be confirmed independently by either a utility or Nuclear Regulatory Commission event notification reports. Limiting the test cases to more recent events helps to provide a current view of the system, while knowing the actual generator allows the effects of location to be examined. Additionally, the Eastern Interconnection contains the largest number of FDRs, which should improve the estimation accuracy. The first nine cases were selected to be large events between 900 and 1300 MW (as estimated by FNET), while the last six were smaller events estimated between 580 and 740 MW. Details for each case are given in Table 5.1.

Table 5.1: Testing Case Descriptions

Case	Generator	Date	Detected Time (UTC)	Estimated Size	Generator Size
1	Watts Bar Unit 1	11/14/2010	11:52:14	1287	1121
2	John Amos Unit 3	11/26/2010	14:20:13	1294	1300
3	Mountaineer Unit 1	11/27/2010	11:34:02	1199	1300
4	Sequoyah Unit 1	12/20/2010	5:50:05	929	1148
5	Watts Bar Unit 1	5/29/2011	5:54:50	1102	1121
6	Sequoyah Unit 1	6/26/2011	20:16:09	1067	1148
7	Gen. James M. Gavin	8/18/2011	6:43:17	1100	1300
8	Sequoyah Unit 1	8/19/2011	2:50:35	1100	1148
9	Browns Ferry Unit 3	9/28/2011	9:15:33	1000	1113
10	Sooner Generating Station Unit 1	11/6/2010	4:07:14	580	569
11	Iatan Unit 2	11/8/2010	11:14:18	610	850
12	Mount Storm Unit 3	1/5/2011	12:56:03	593	567
13	Dolet Hills Unit 1	5/26/2011	3:56:01	622	721
14	Limerick Unit 2	5/29/2011	9:02:42	673	1200*
15	Limerick Unit 1	6/3/2011	14:20:44	740	1200
* Generator was operating at 75% output, according to NRC					

Results

The proposed algorithm was implemented in MATLAB, and the resulting speed calculations were stored in a MySQL database. Although the speeds were initially normalized for each case, this did not provide any real benefit; thus the actual speeds were used for subsequent analyses. The data were then imported into ArcMap, which was used to create shaded gradient maps for each case (Fig. 5.3 through Fig. 5.17). In addition, the median and average speeds found for each measurement location were used to create composite maps, which are shown in Fig. 5.18 and Fig. 5.19. Due to how ArcMap renders the raster image for the gradient, some areas such as Texas and Quebec are shaded, even though the speeds were not computed for these regions.

Although the speeds calculated using this method still demonstrate significant variability, some general observations can be made. Transients appear to travel quickly through the northwestern portion of the EI, and more slowly in the upper northeast. These observed differences can be explained by the fact that the major population centers of the East Coast contain significant amounts of generation (Fig. 5.20) and load (Fig. 5.21)[1], which both contribute to system inertia and act to slow down the propagating wave. In comparison, the northwestern part of the EI is sparsely populated (Fig. 5.22) and has a much smaller amount of inertia. The central portion of the EI near eastern Tennessee and southeastern Kentucky also appears to be slower than the surrounding regions, perhaps due to the large number of generating facilities in that area.

Aggregated speed estimates for each case and FDR location are given in

[1] Fig. 5.20 and Fig. 5.21 were constructed from a 29,000-bus EI model that does not include parts of Florida and New England.

Table 5.2 and Table 5.3, respectively. It is apparent that the mean and median speeds for each case fall within a fairly narrow range, which demonstrates the improved consistency of this technique. It can also be observed that the overall speeds are essentially independent of the event size and location. Calculated speeds were higher than the 500 miles/s figure typically assumed for the EI, and lower than those estimated by Kook using time-domain simulations of the system [35]. One reason for this could be that earlier techniques used the time elapsed since the disturbance occurred in the denominator of the speed formula, which would result in lower speeds. This approach is problematic since in practice the exact time of a disturbance is usually not known to sub-second resolution. The method presented here should provide a more realistic estimation of the local propagation speed in each region.

Table 5.2: Aggregated speed estimates for each case

Case	Mean (mi/s)	Median (mi/s)	Minimum (mi/s)	Maximum (mi/s)
1	1260	1202	531	2922
2	1228	1176	428	2539
3	1311	1252	687	2441
4	1098	968	533	2710
5	1232	1180	615	2441
6	1196	1203	211	2246
7	1101	940	466	2307
8	1035	1044	418	1844
9	1216	1120	452	2077
10	1178	1242	593	1878
11	1291	1293	367	2539
12	886	742	375	1991
13	1144	1038	459	2758
14	793	693	40	2406
15	1048	975	448	1930

Table 5.3: Aggregated speed estimates for each FDR

FDR	City	State	Mean (mi/s)	Median (mi/s)	Minimum (mi/s)	Maximum (mi/s)
673	Montgomery	AL	1169	1210	720	1796
666	Danbury	CT	778	762	533	1072
663	Gainesville	FL	1389	1395	894	1930
796	North Palm Beach	FL	1344	1203	780	2406
674	Pensacola	FL	1181	1190	373	1971
623	Plant City	FL	767	775	466	1001
786	Tallahassee	FL	1197	1390	263	1591
675	Atlanta	GA	1052	789	606	1891
754	Cedar Falls	IA	1016	1008	524	1913
714	Urbandale	IA	1408	1293	317	2922
620	Chicago	IL	935	852	537	1469
767	Marion	IL	809	746	459	1521
778	Matton	IL	978	982	422	1342
755	Urbana	IL	1025	992	863	1215
710	Carmel	IN	961	952	87	1578
713	Hammond	IN	959	904	611	1383
797	Hammond	IN	1039	1039	905	1173
682	Holyoke	MA	769	739	40	1203
684	Waltham	MA	701	719	45	1110
715	Grand Rapids	MB	1395	1539	471	2166
718	Kelsey	MB	960	847	424	2246
719	Thompson	MB	977	1057	469	1829
665	Winnipeg	MB	1780	1962	728	2539
733	Bangor	ME	1047	948	75	2018
621	Detroit	MI	1150	1194	675	1591
679	Grand Rapids	MI	1090	1022	235	1819
712	Houghton	MI	1173	1006	245	2758
750	Kingsford	MI	1179	1132	368	2233
760	Elk River	MN	1171	1299	238	1794
790	Elk River	MN	1816	1816	1816	1816
720	Fergus Falls	MN	1640	1749	813	2383
619	St. Paul	MN	1288	1303	121	2336
832	St. Paul	MN	1469	1405	1100	1951
616	Kansas City	MO	1085	1120	705	1612
781	Kirksville	MO	1085	1054	636	2102
756	Union	MO	1006	996	614	1620
672	Gulfport	MS	1158	1093	46	2134
730	Glendive	MT	895	895	895	895

753	Hampstead	NC	1185	1416	330	1627
664	Bismarck	ND	2080	2226	1164	2710
740	Devil's Lake	ND	1691	1731	1321	2077
729	Dickinson	ND	1295	1247	888	2039
726	Williston	ND	1519	1382	1209	2213
804	Atlantic City	NJ	633	633	613	652
678	Princeton	NJ	751	695	375	1067
667	Leroy	NY	1136	1165	960	1417
835	Potsdam	NY	1844	1844	1844	1844
707	Troy	NY	836	867	279	1265
830	Cincinnati	OH	1164	1131	1117	1269
696	Cleveland	OH	1184	1333	428	1795
670	Gahanna	OH	1199	1222	674	1617
803	Gahanna	OH	1044	1044	1044	1044
759	Norman	OK	1292	1388	925	1388
669	Simpsonville	SC	1141	1002	611	2191
739	Big Stone	SD	1449	1446	905	1948
868	Gallatin	TN	683	683	683	683
692	Knoxville	TN	602	728	254	730
770	Knoxville	TN	558	484	211	1174
722	Oak Ridge	TN	480	459	418	580
683	Lubbock	TX	1292	1388	925	1388
661	Alexandria	VA	962	918	500	1394
639	Blacksburg	VA	1000	1035	560	1309
785	Blacksburg	VA	980	911	771	1308
857	Fredericksburg	VA	1209	1209	1209	1209
668	Newport News	VA	1303	1303	1303	1303
601	Richmond	VA	857	875	582	1067
686	Roanoke	VA	1086	1093	518	1448
688	Charleston	WV	995	1002	574	1350
744	Morgantown	WV	1052	1062	781	1435

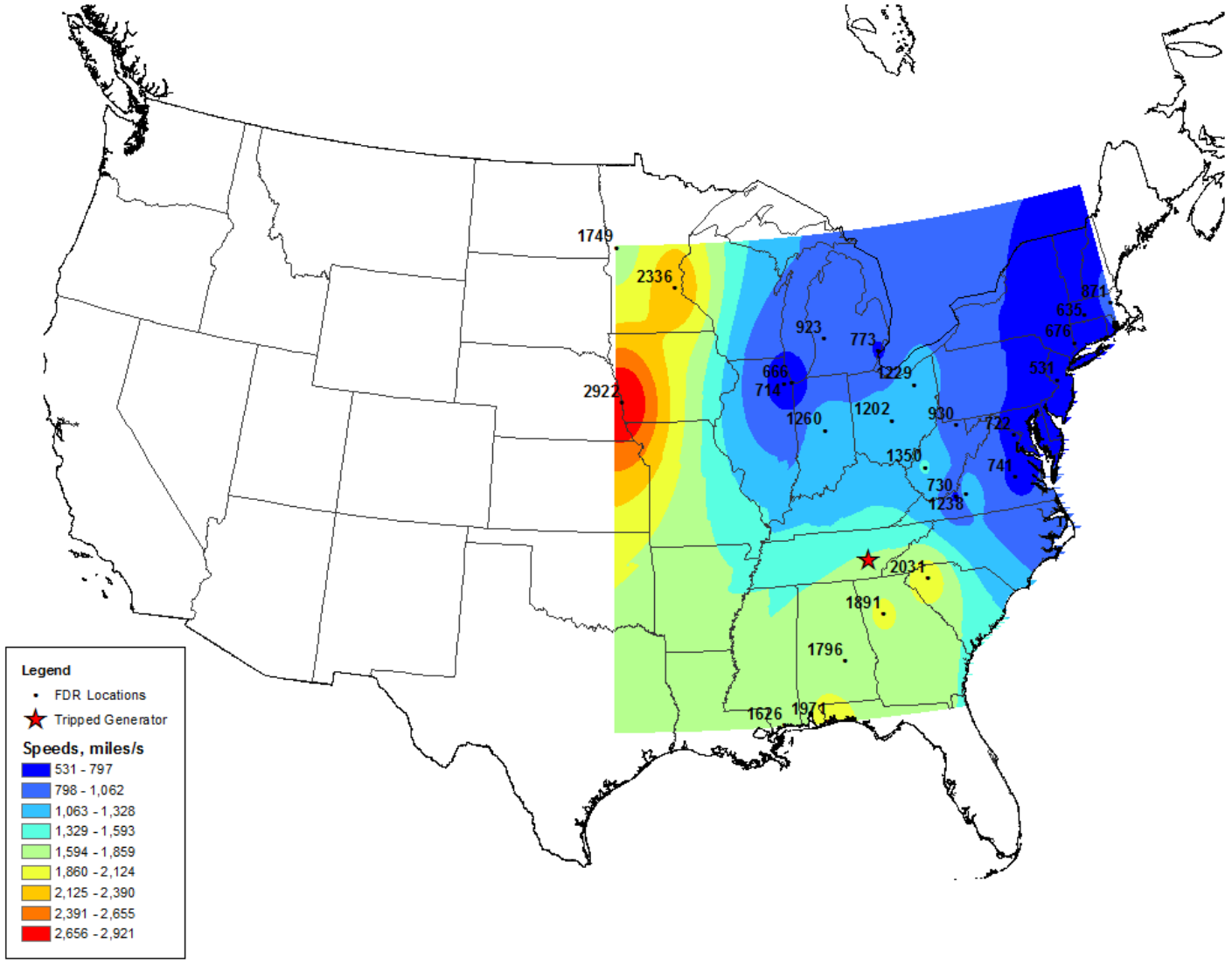


Fig. 5.3: Case 1 - Watts Bar Unit 1

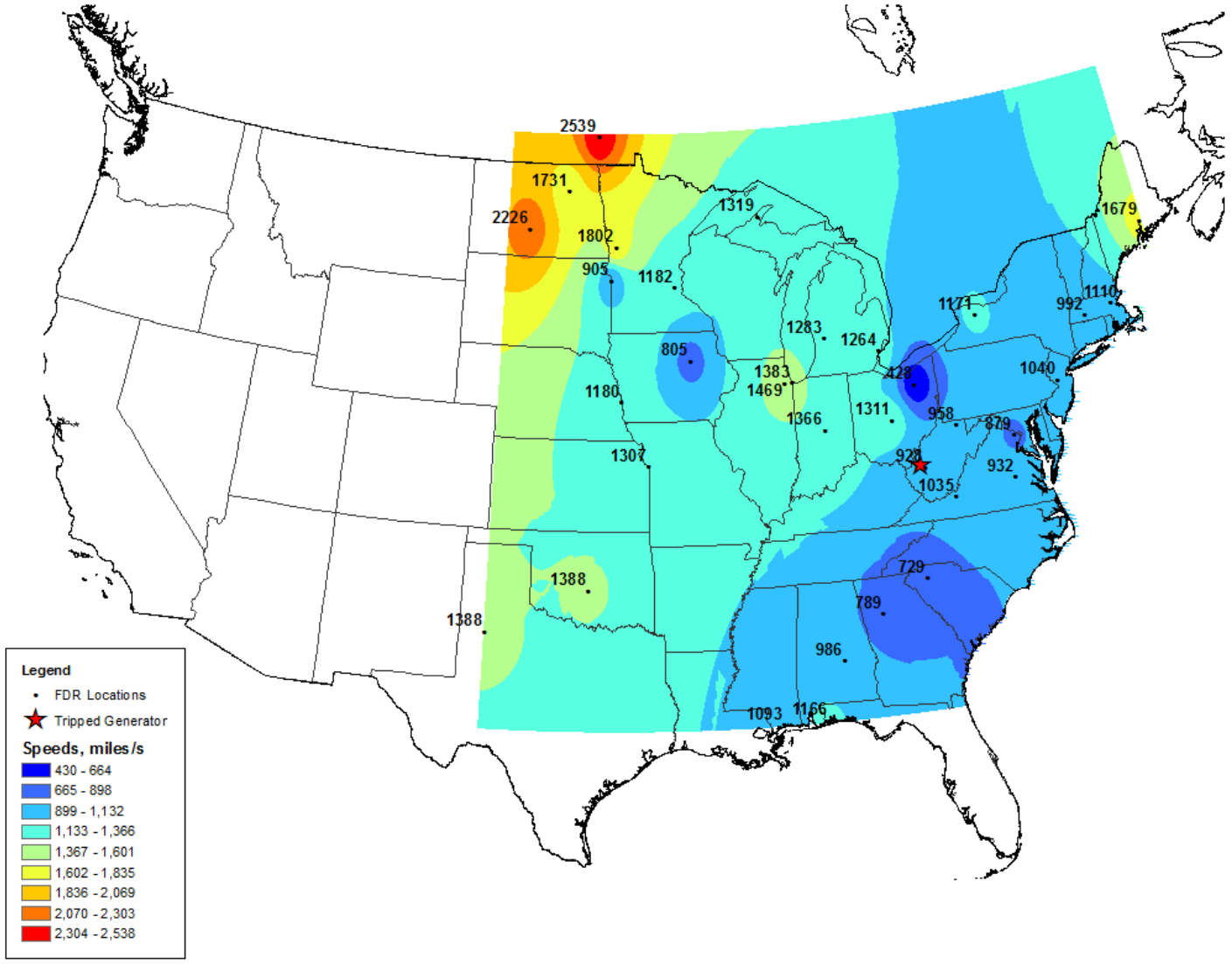


Fig. 5.4: Case 2 - John Amos Unit 3

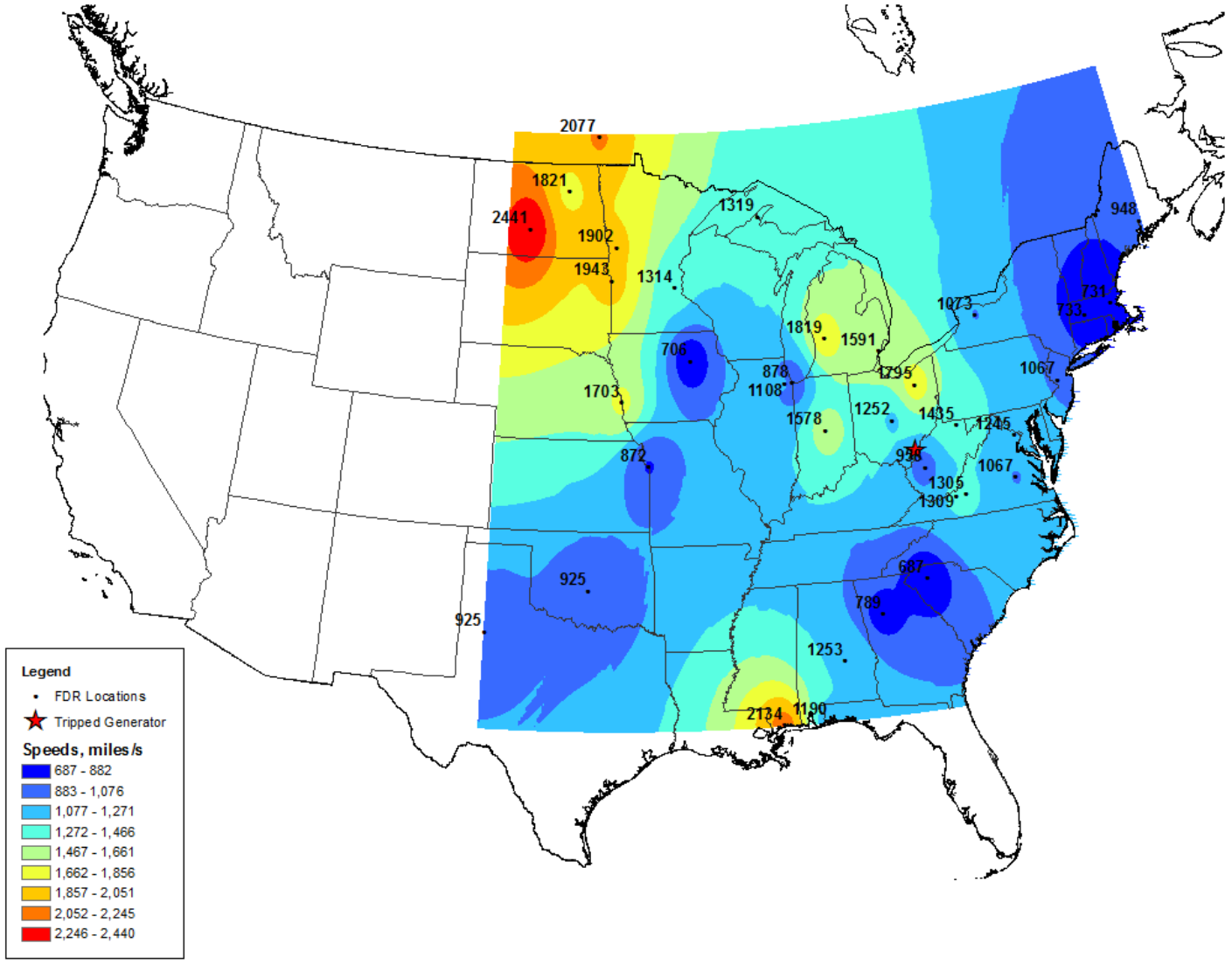


Fig. 5.5: Case 3 - Mountaineer Unit 1

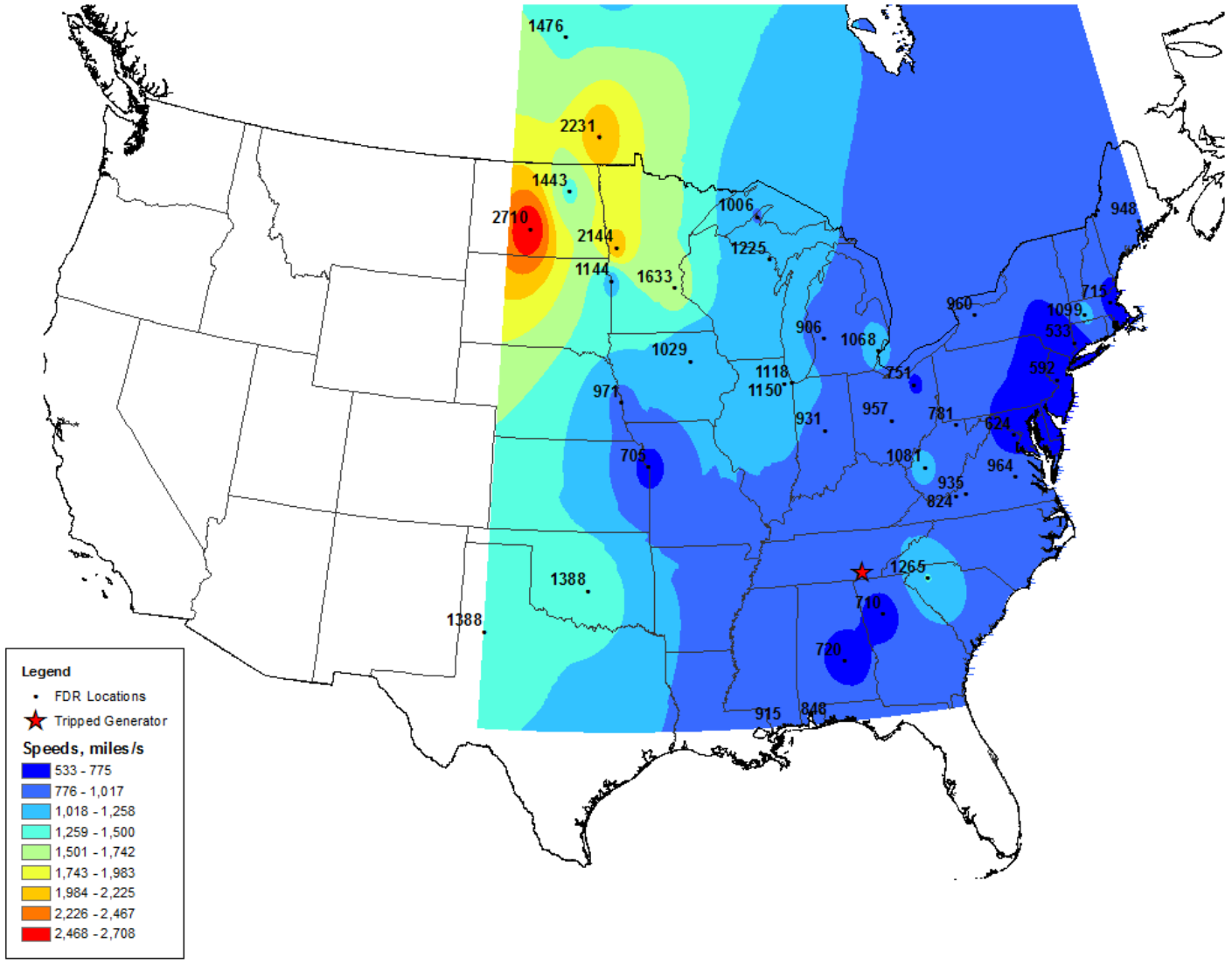


Fig. 5.6: Case 4 - Sequoyah Unit 1

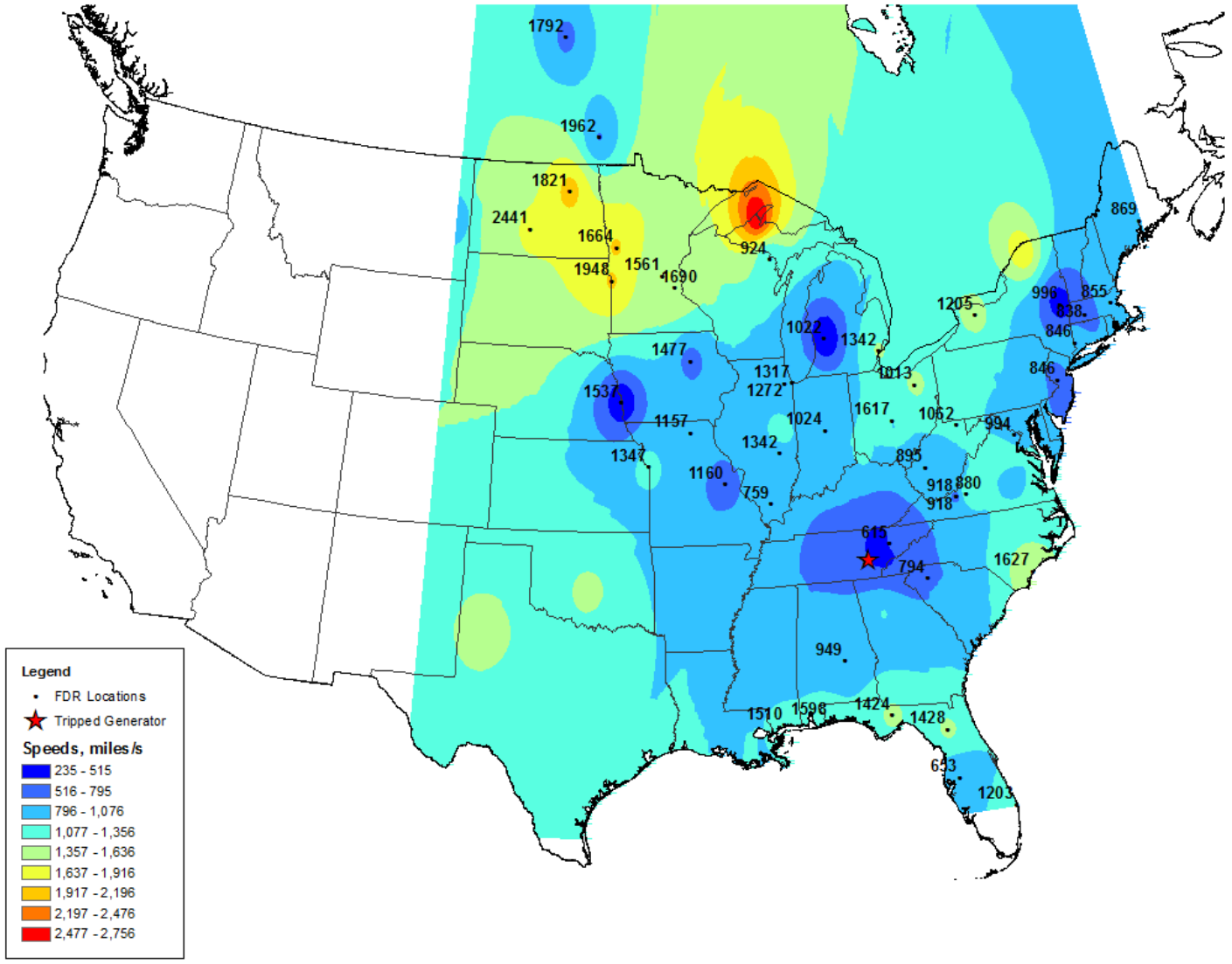


Fig. 5.7: Case 5 - Watts Bar Unit 1

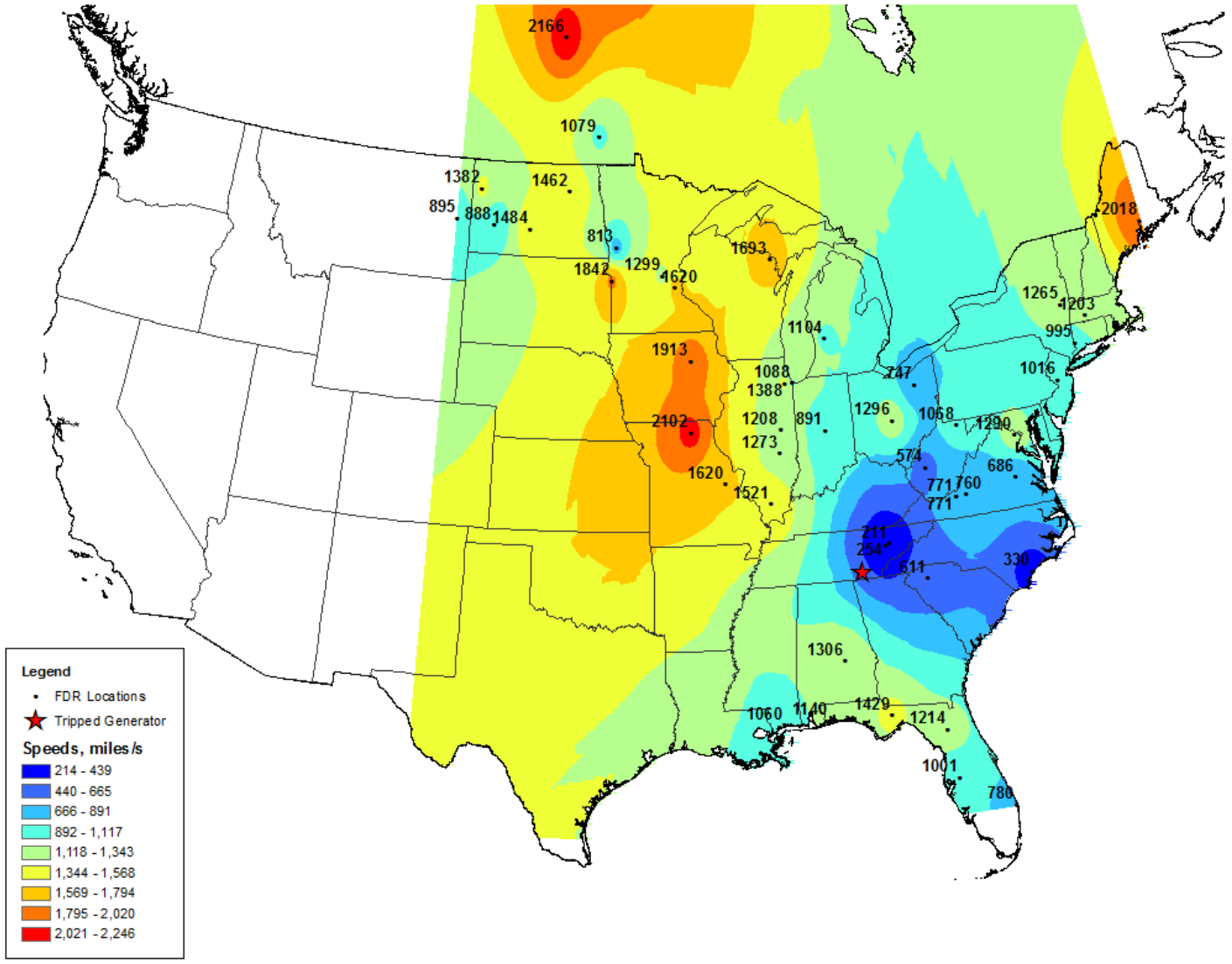


Fig. 5.8: Case 6 - Sequoyah Unit 1

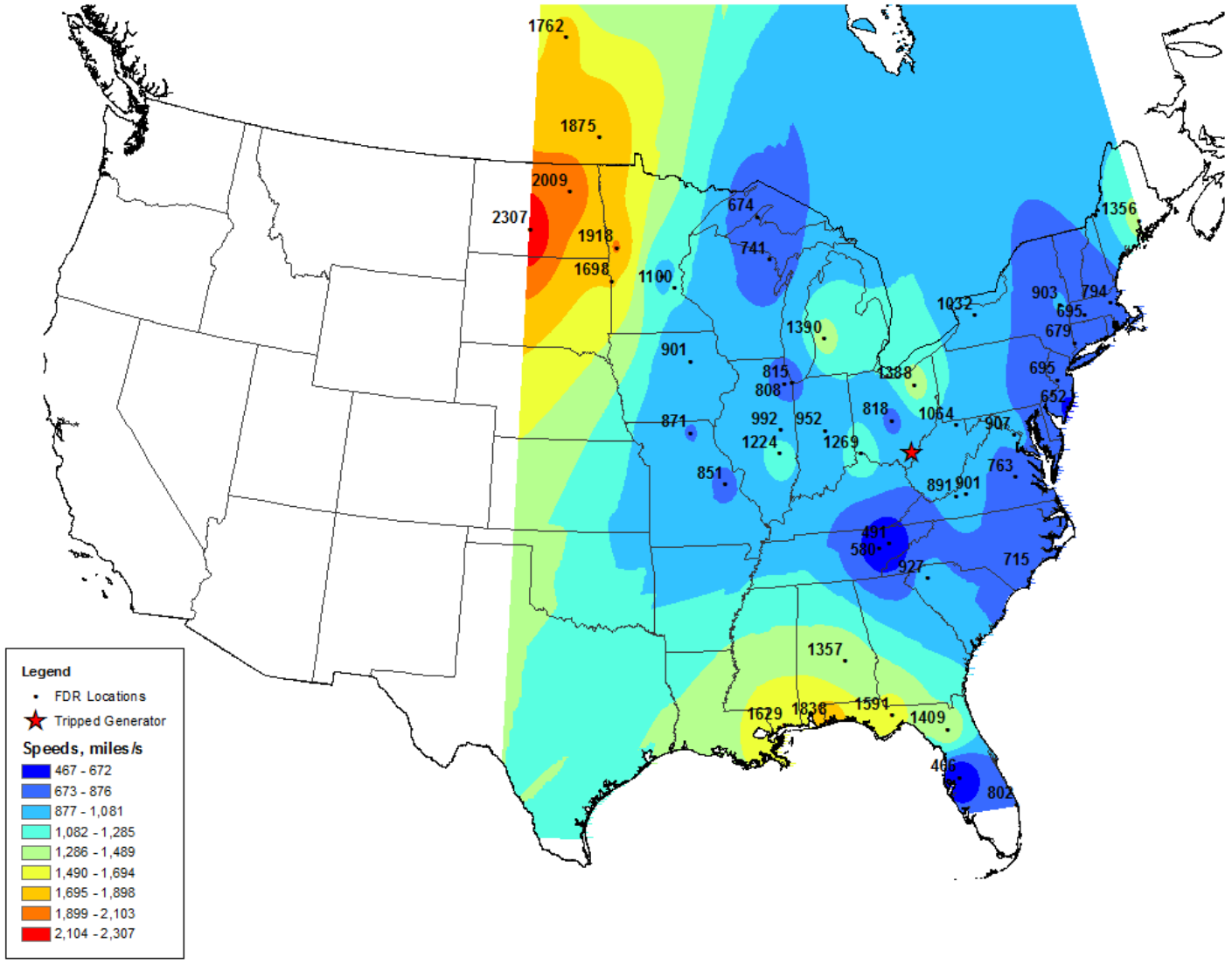


Fig. 5.9: Case 7 - Gen. James M. Gavin

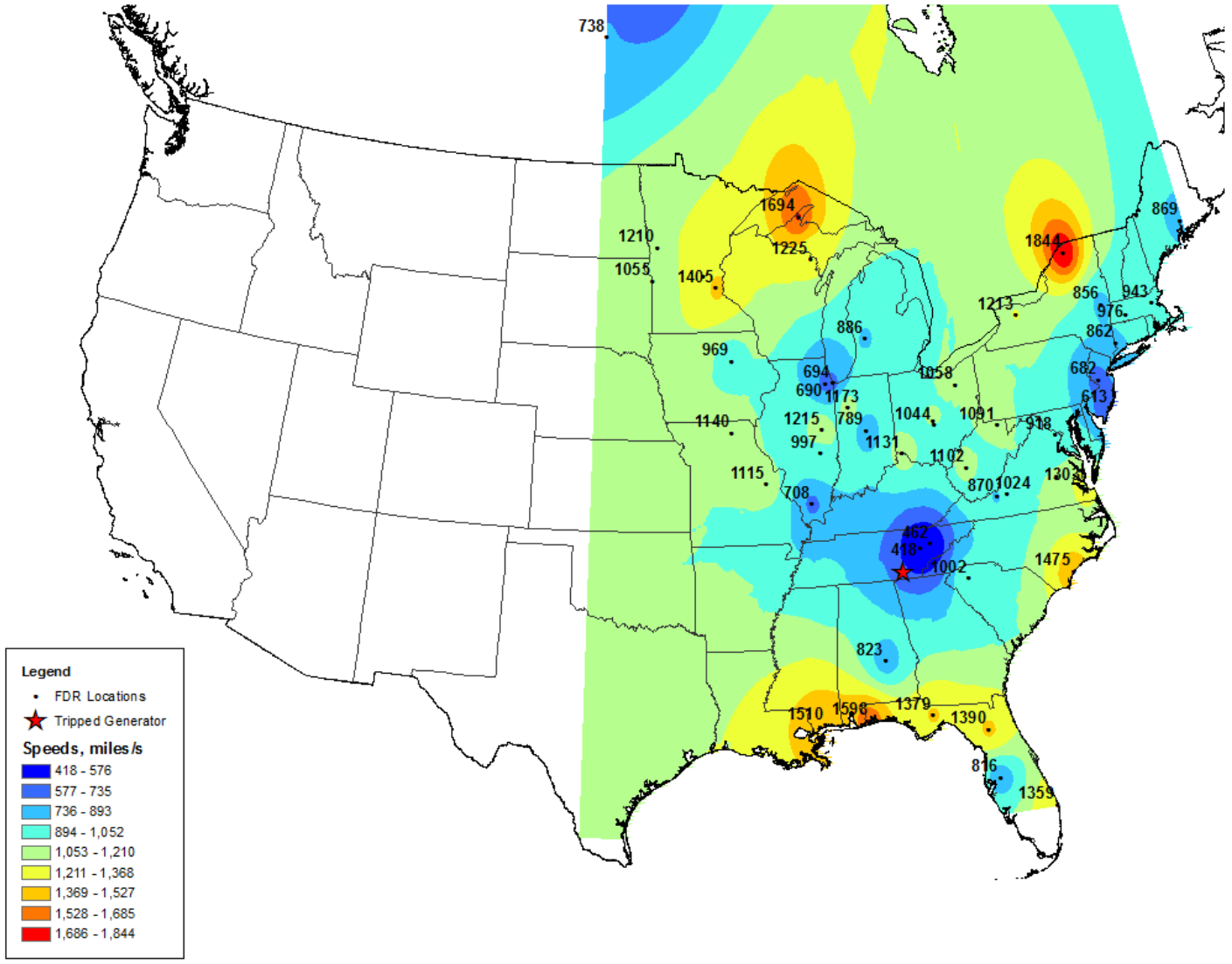


Fig. 5.10: Case 8 - Sequoyah Unit 1

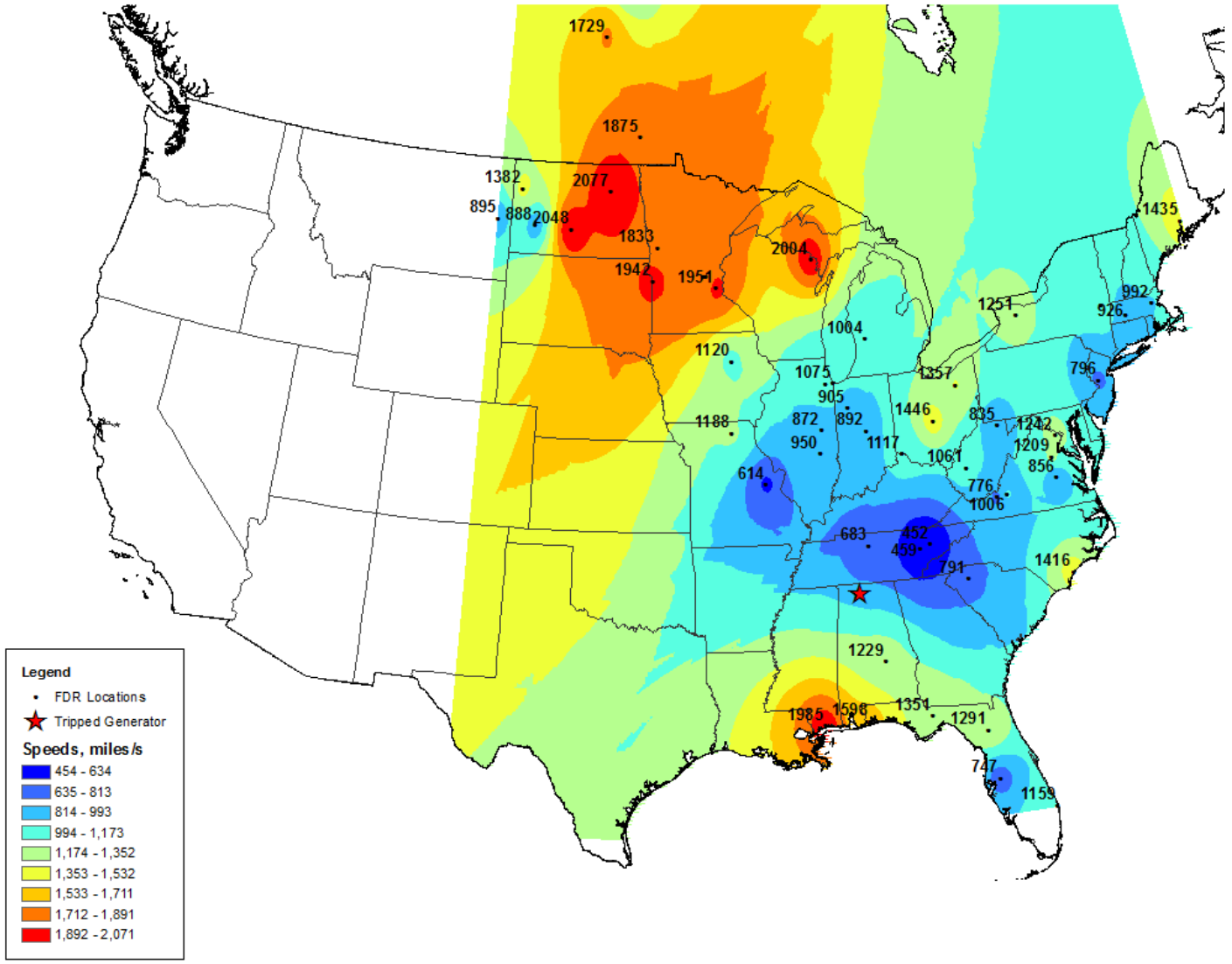


Fig. 5.11: Case 9 - Browns Ferry Unit 3

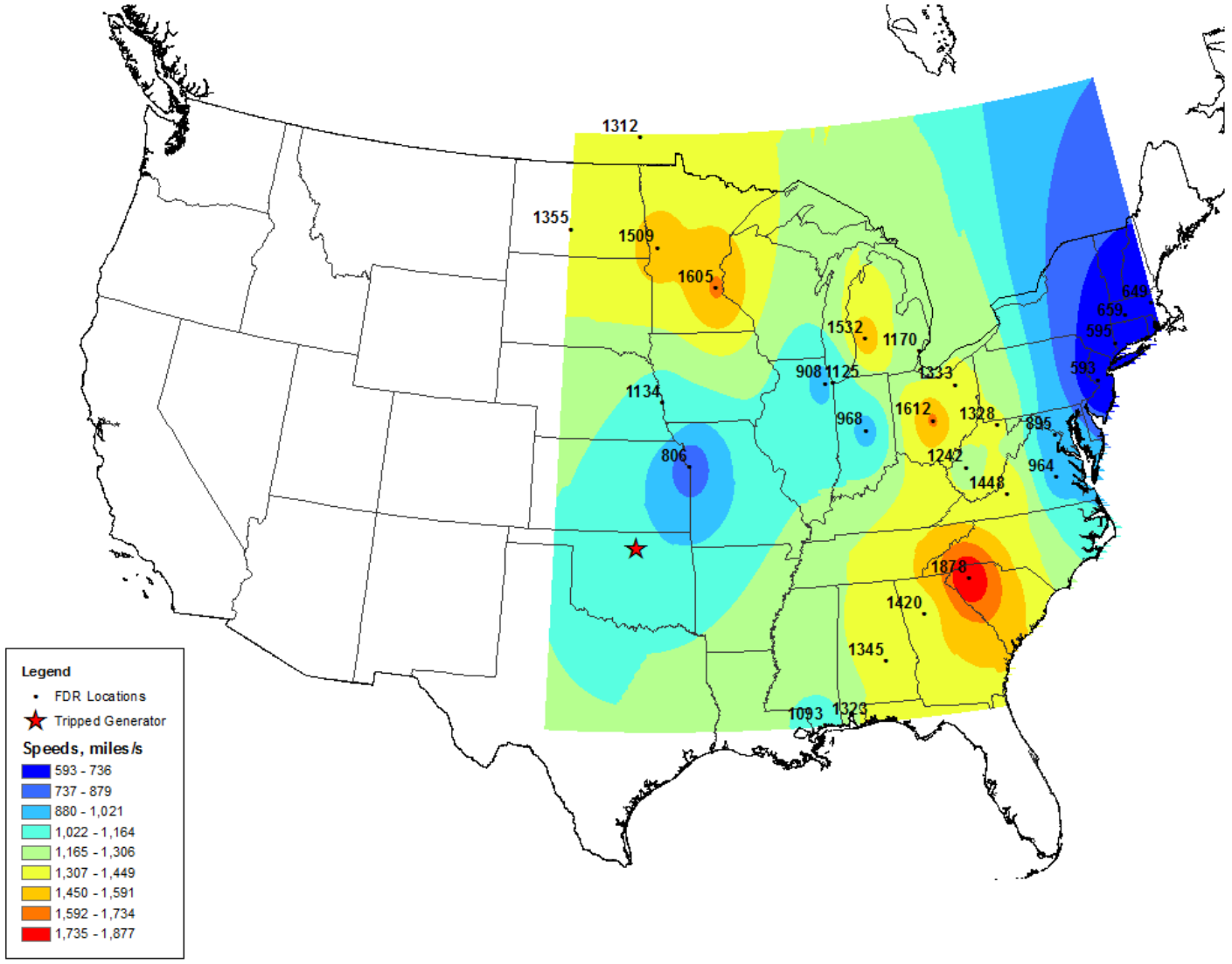


Fig. 5.12: Case 10 - Sooner Unit 1

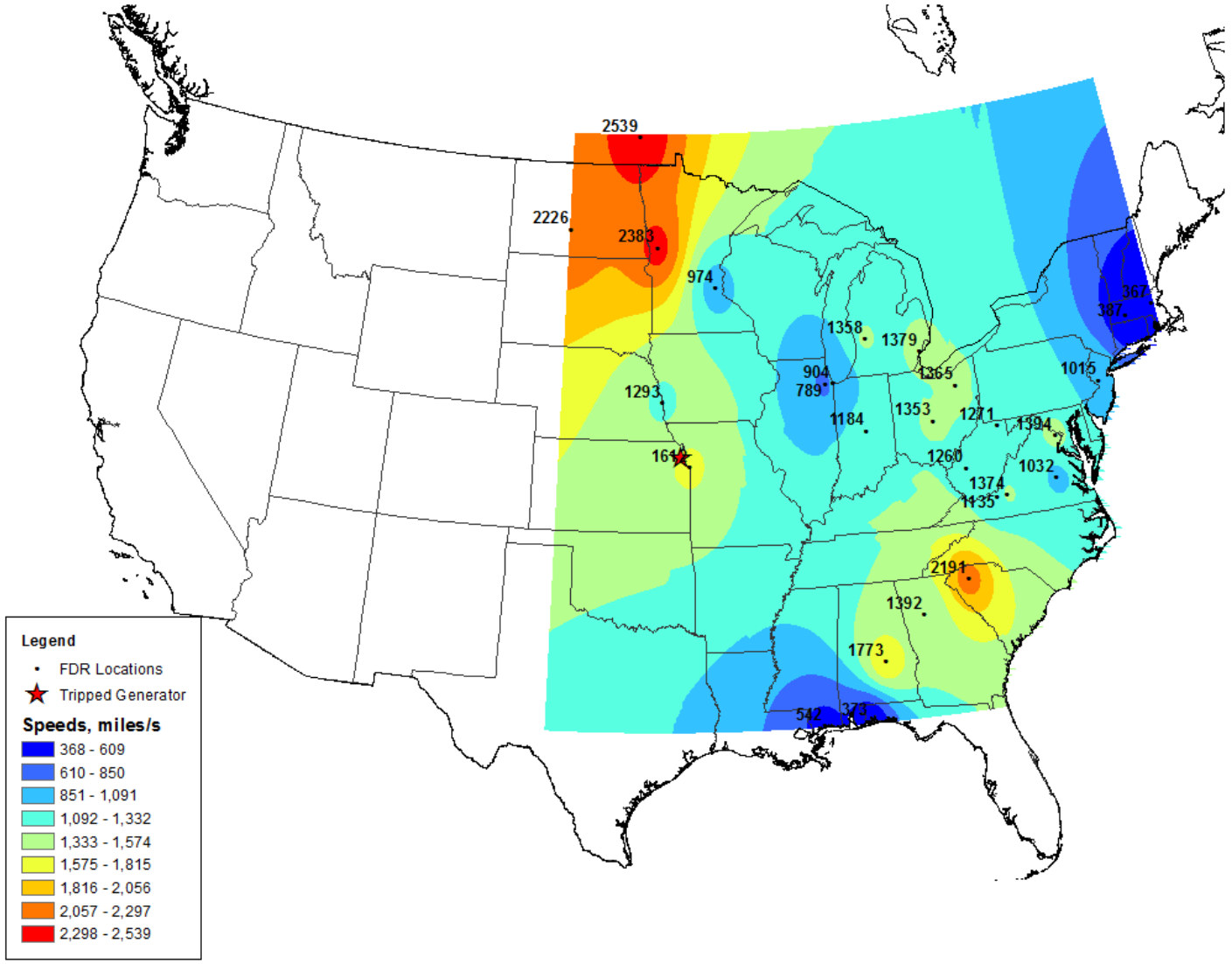


Fig. 5.13: Case 11 - Iatan Unit 2

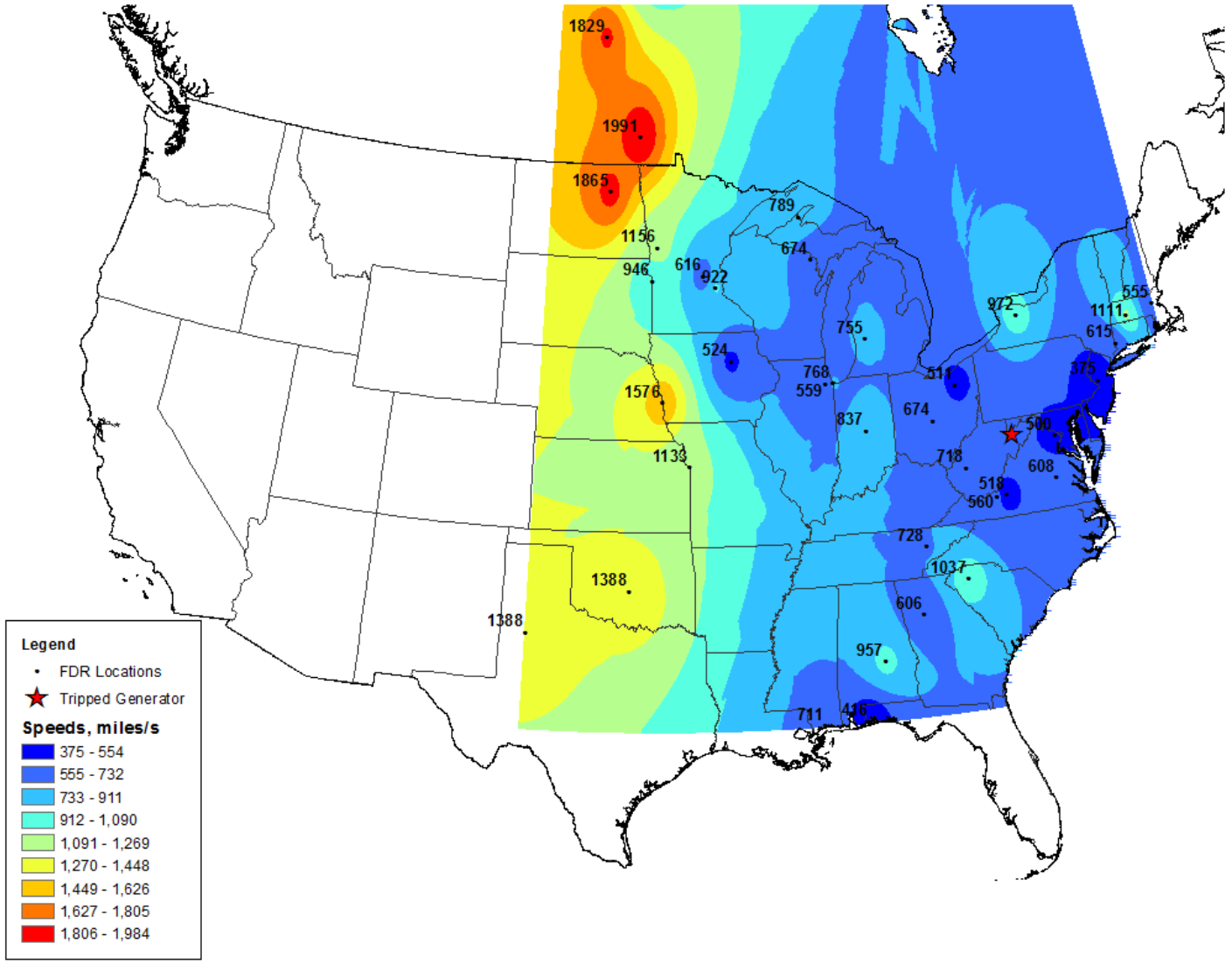


Fig. 5.14: Case 12 - Mt. Storm Unit 3

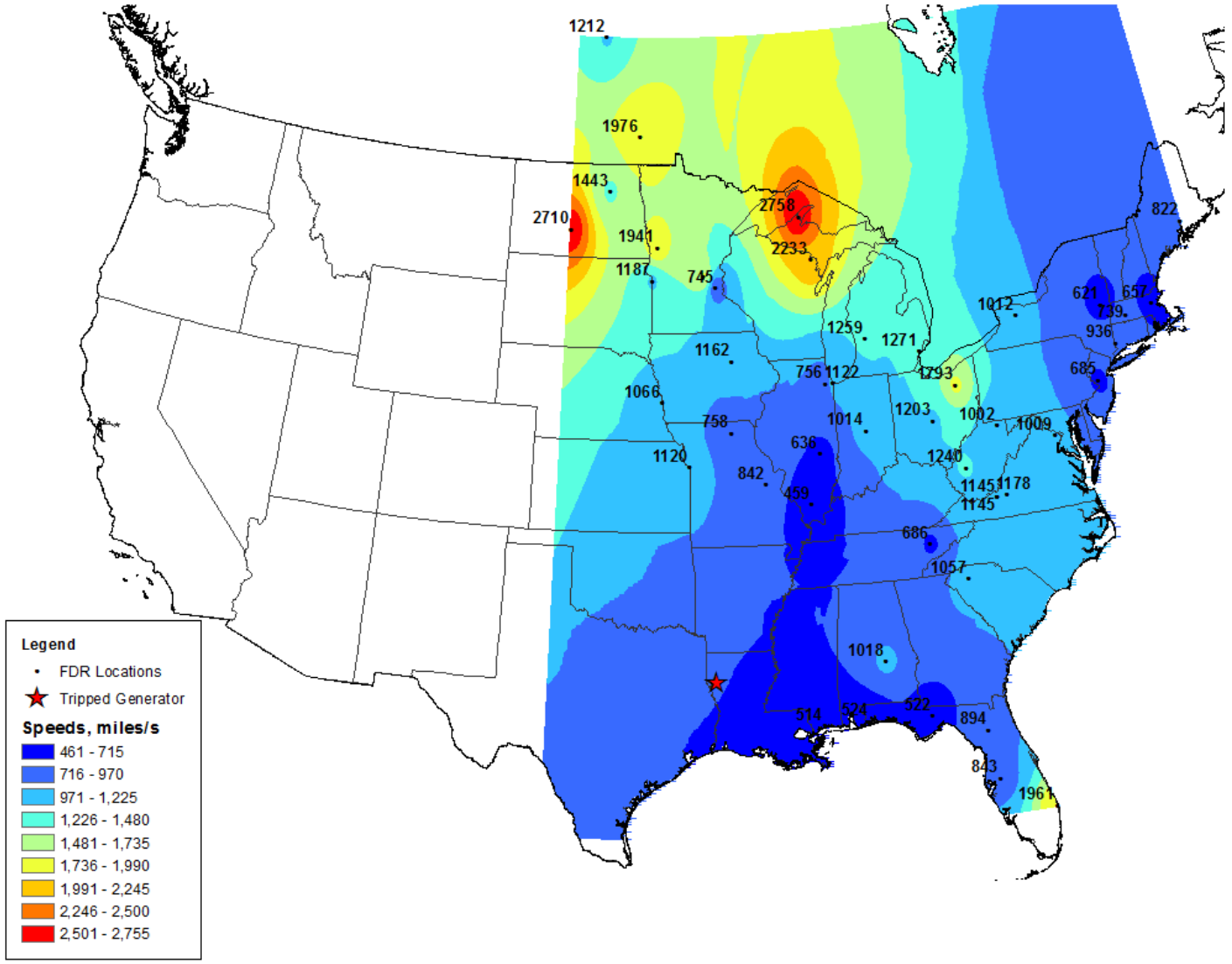


Fig. 5.15: Case 13 - Dolet Hills Unit 1

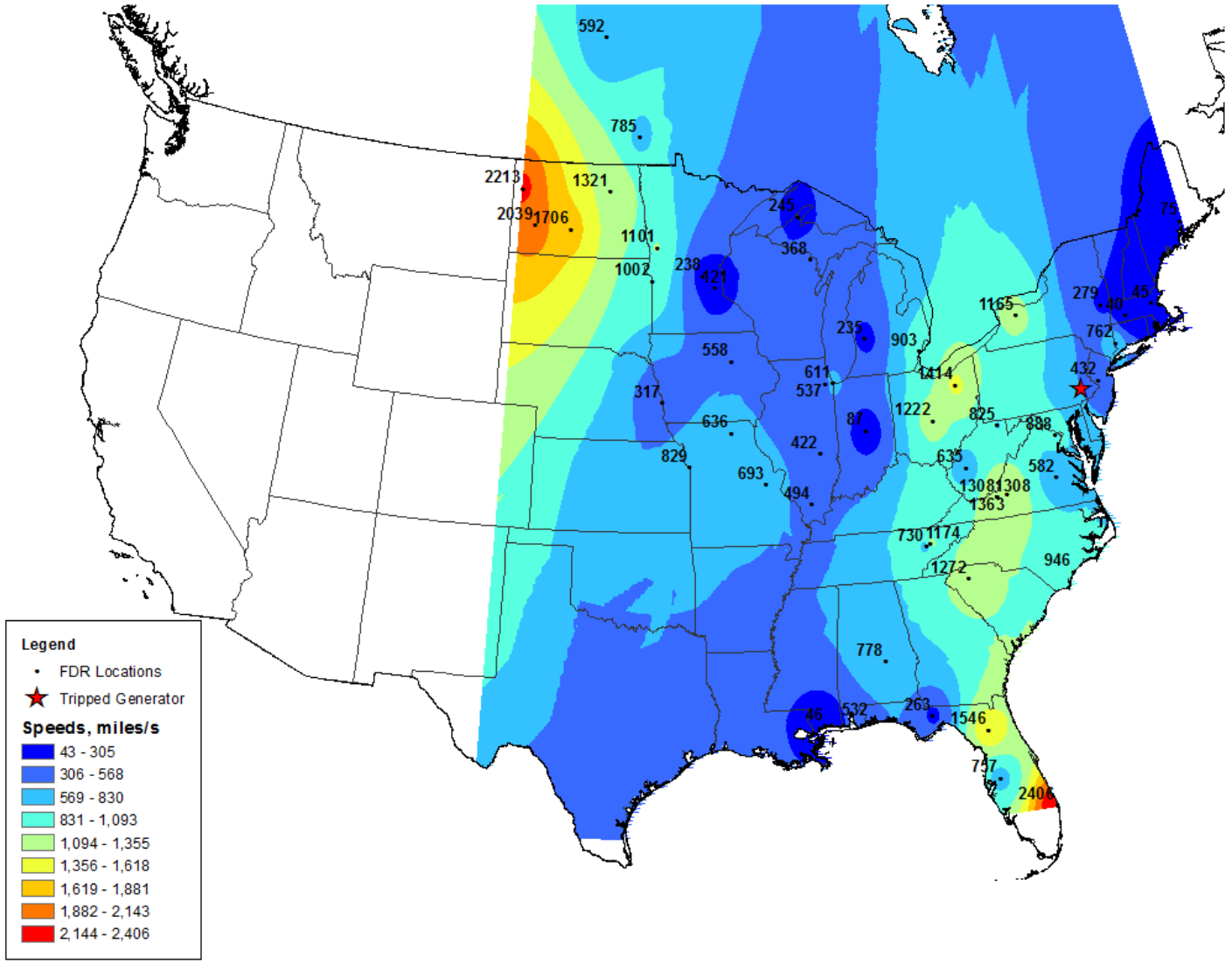


Fig. 5.16: Case 14 - Limerick Unit 2

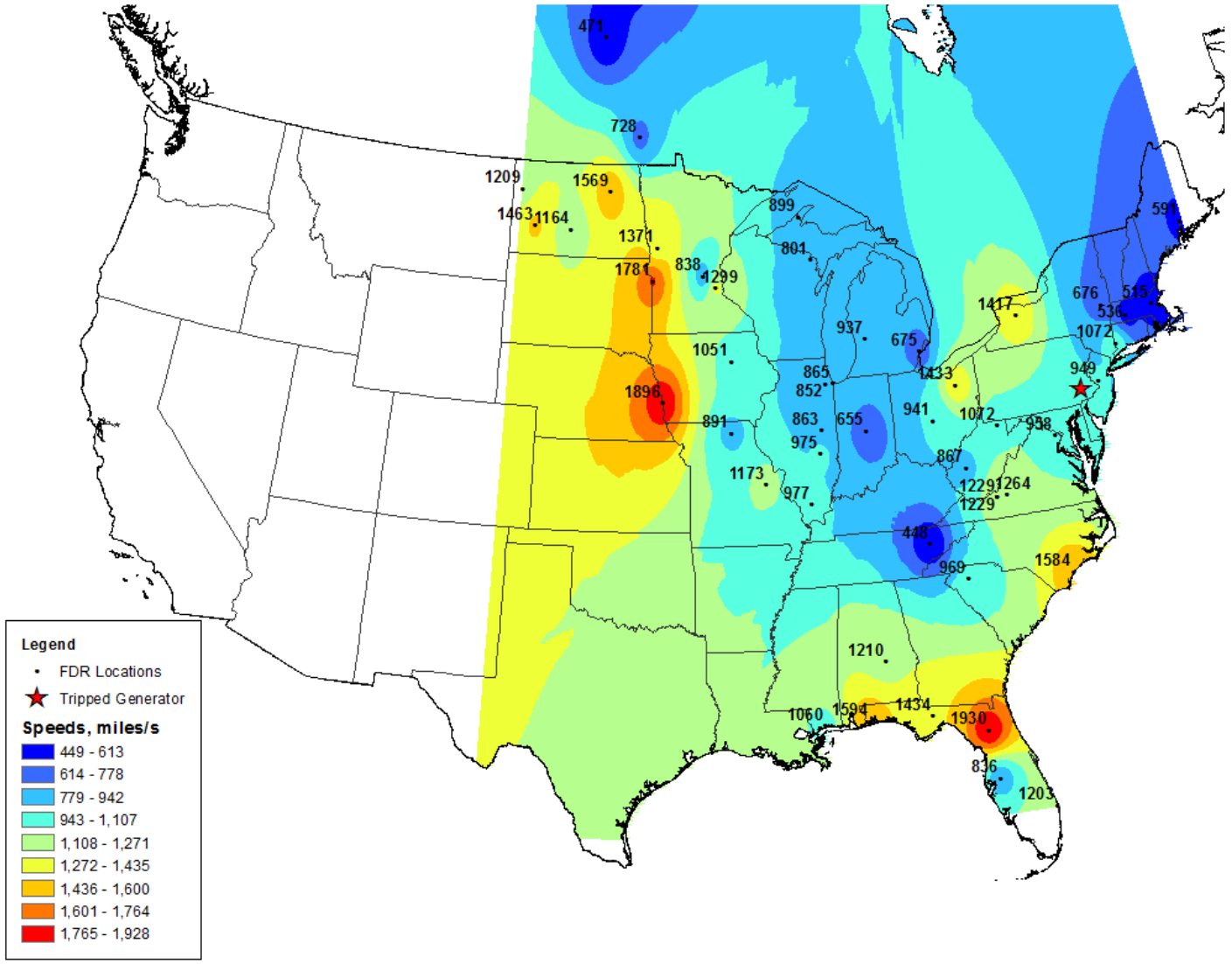


Fig. 5.17: Case 15 - Limerick Unit 1

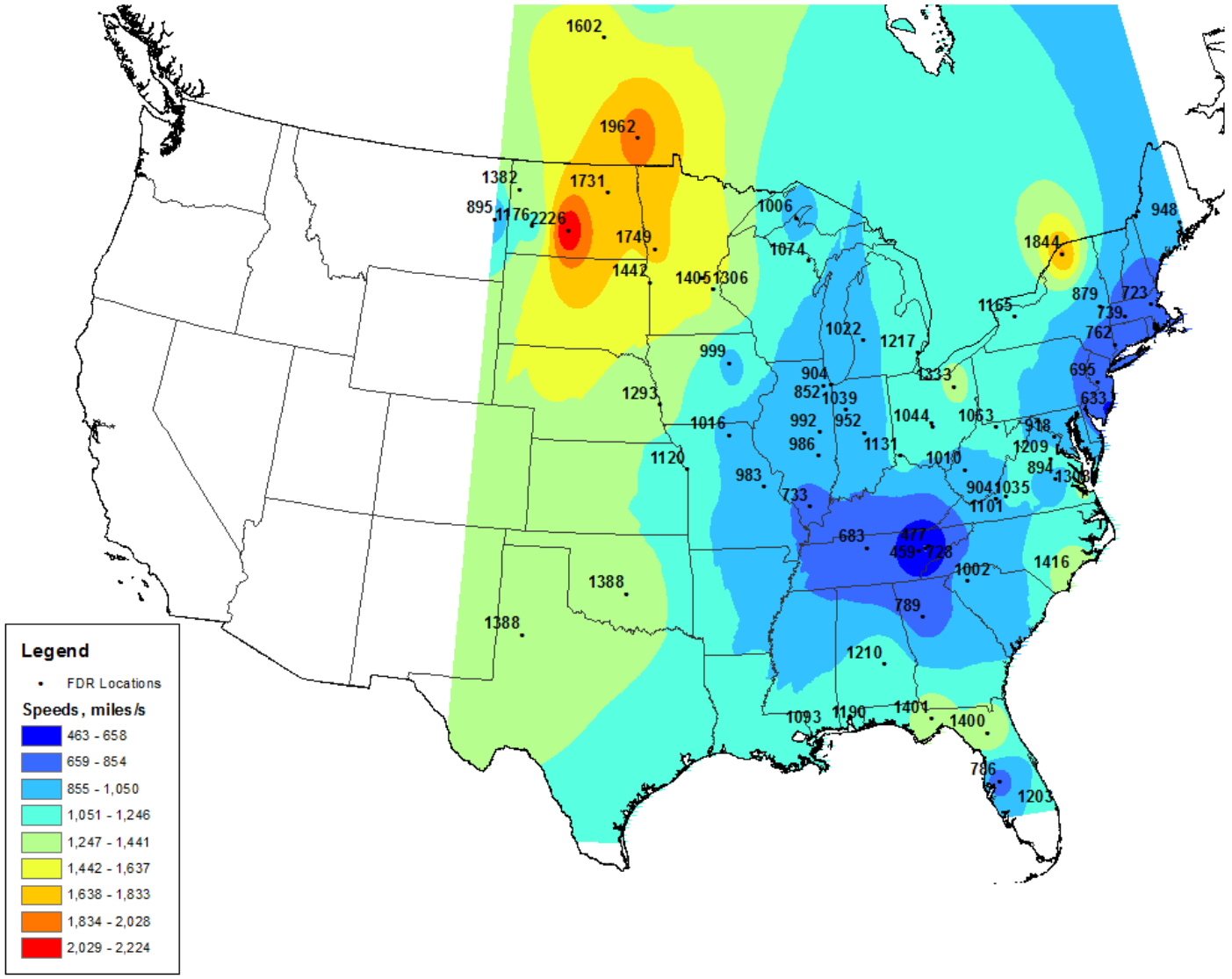


Fig. 5.18: Median speeds - all cases

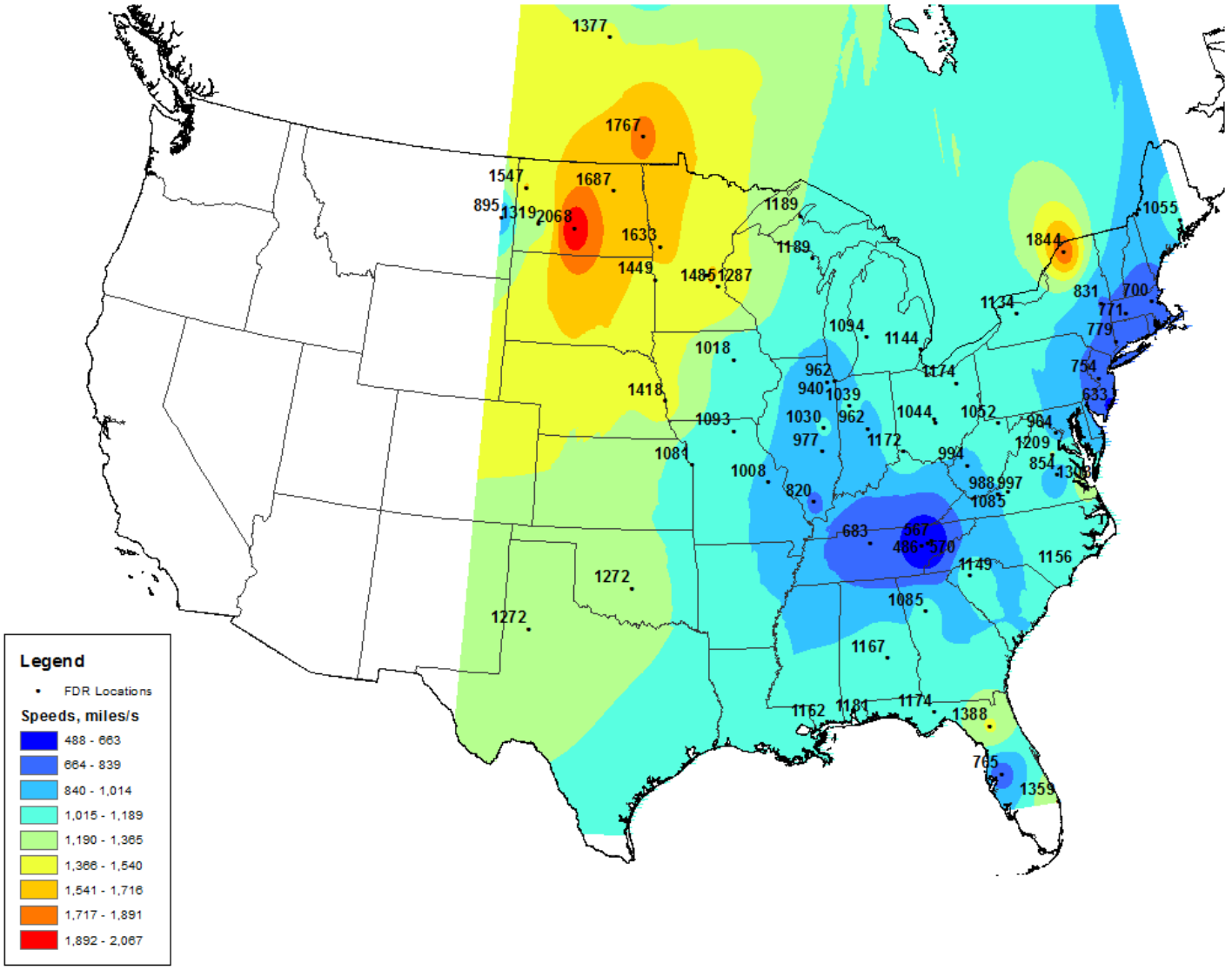


Fig. 5.19: Average speeds - all cases

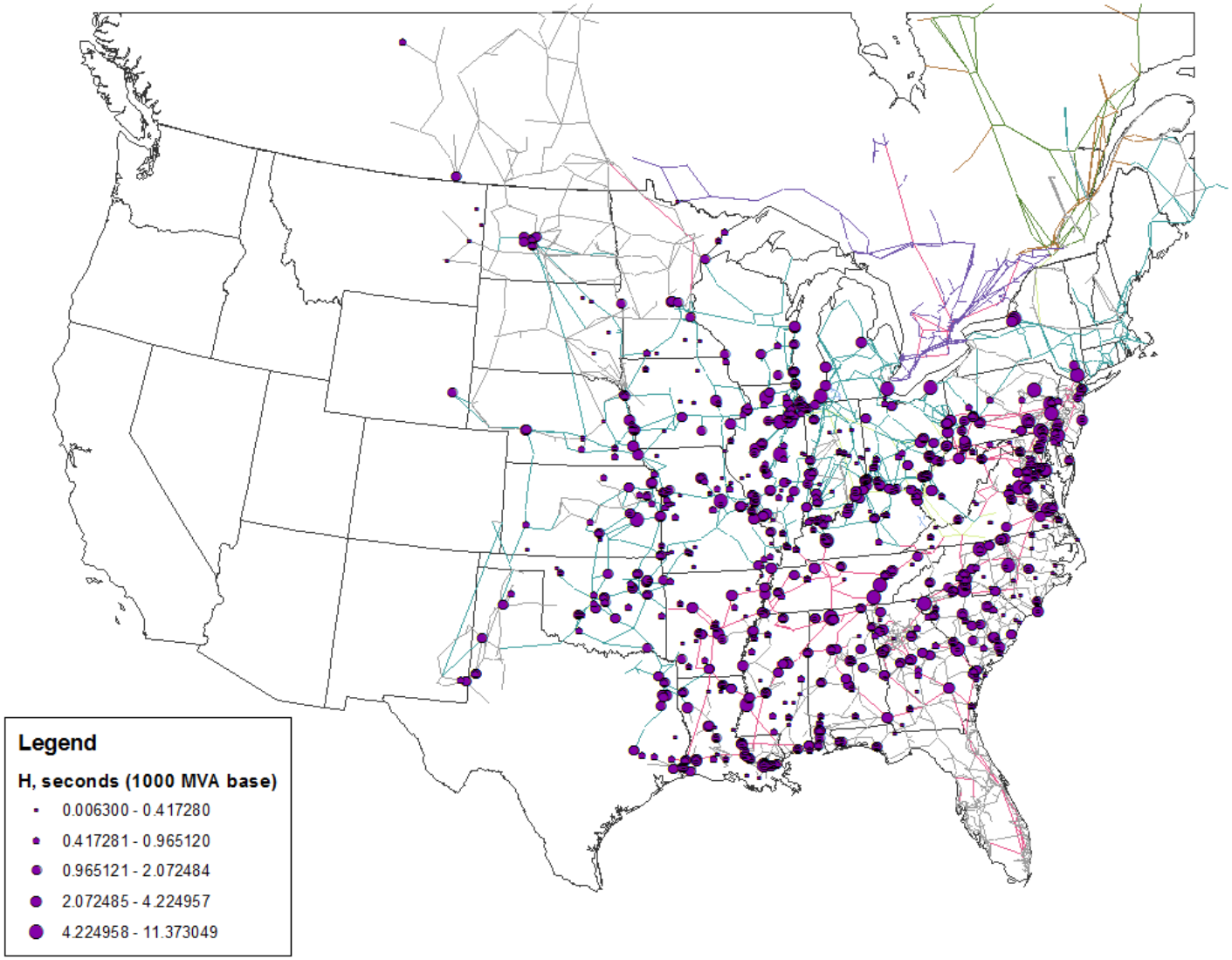


Fig. 5.20: Inertia constant (H) of generators in the Eastern Interconnection

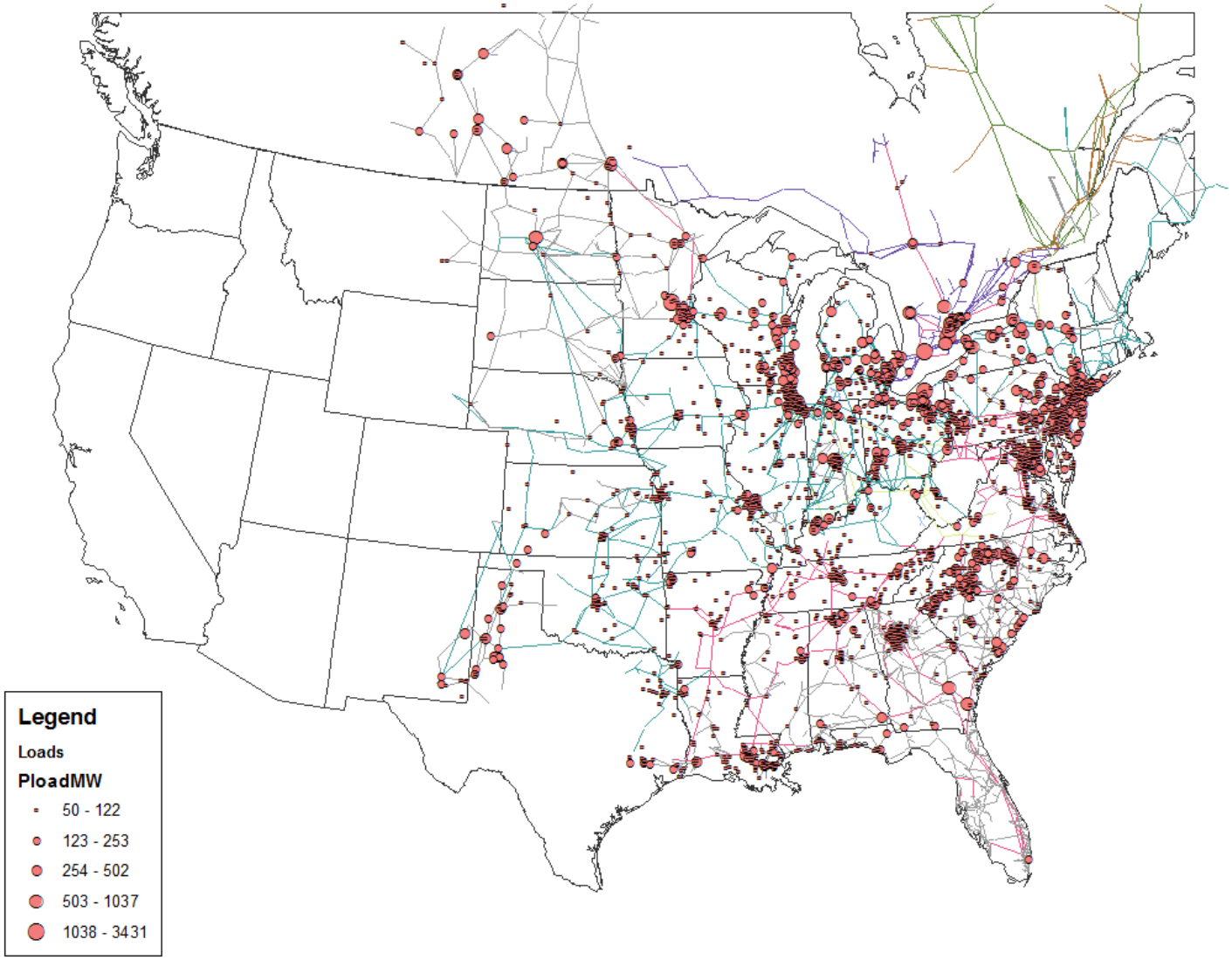


Fig. 5.21: Buses with loads greater than 50 MW

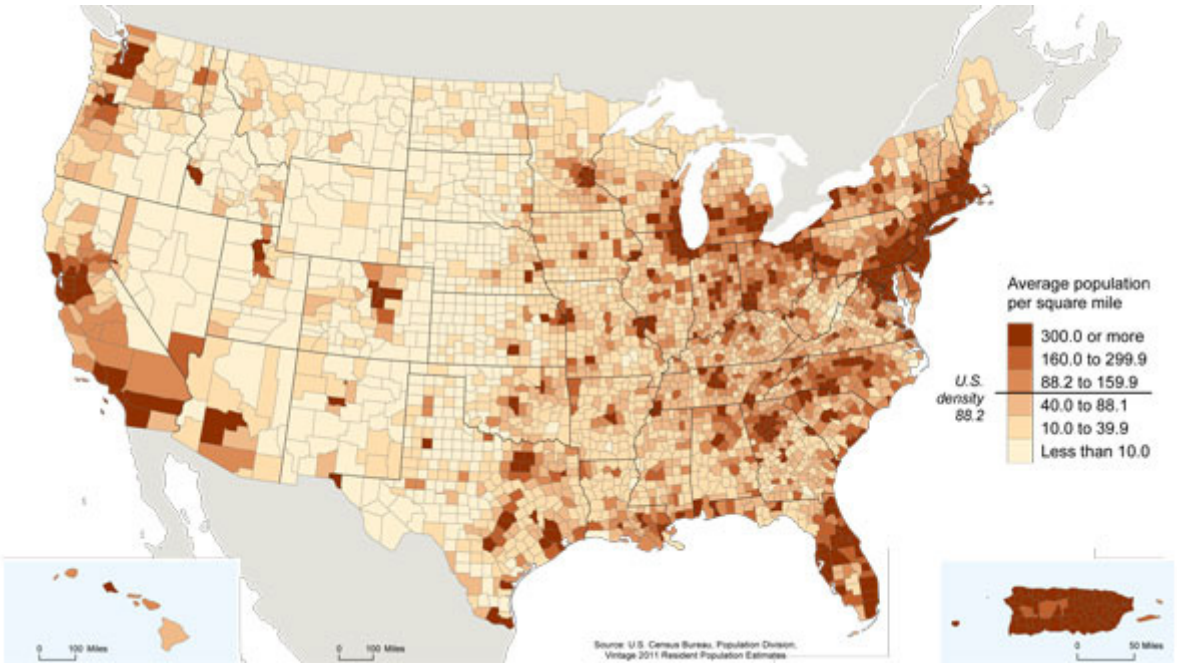


Fig. 5.22: Population density map of the United States (2011)

Conclusions

This chapter has presented a new, model-agnostic method of determining electromechanical wave propagation speeds in power systems using synchronized phasor measurement data. Unlike previous techniques that averaged the behavior of large portions of the system together, this method only considers the propagation speeds over small areas. By removing the dependence on knowing the exact time of the disturbance, a more consistent estimation of the propagation speeds can be obtained. Several conclusions can be drawn from this work:

- It is probably infeasible to produce a single speed map that is always true for the system being considered. The topology of a power system is constantly changing, and this has a significant effect on its transient behavior. Still, the system should not vary considerably over small periods of time (e.g., several hours). This suggests that perhaps the speed map could be continually updated based upon the most recent disturbances, which would certainly be possible given that most power systems experience several events each day. Such a map should improve the accuracy of FNET's triangulation algorithm and could play an important role in future wide-area control strategies.
- Additional measurement points would provide better granularity and lessen the effects of random errors in the data. Thus, the accuracy of this method should improve as more FDRs are added to the system.

- Frequency Disturbance Recorder measurements are timestamped at 100-ms intervals. This introduces error in the speed calculation since the time delay must always be rounded up to the nearest tenth of a second. Ideally, a faster reporting rate should be used.
- The frequency cutoff value used to determine when a given location observes the disturbance should not dramatically affect the estimated speed provided that it is chosen in the linear region immediately following the event. While the cutoff is usually easy to obtain visually, it may be difficult to determine automatically, particularly if significant noise or oscillation is present in the data.

6. FNET Software Framework

Introduction

Since its inception in 2004, FNET has experienced steady growth in the number of installed Frequency Disturbance Recorders. The 131 FDRs (as of the time of this writing) that have been deployed around the world each transmit about 55 bytes of data 10 times per second. This means that the FNET server must process and store roughly 6 GB of data per day, which is currently archived in Microsoft Access databases. Accessing and preparing the data for offline analysis has traditionally been a tedious task for a number of reasons. This chapter examines some of the challenges faced by those working with the FDR data, and describes an object-oriented software framework designed to simplify a number of common tasks. It is hoped that this framework will allow future researchers to spend more time analyzing the data and less time retrieving it.

Structure of FDR Data

The FNET server program continuously saves the incoming FDR data to a Microsoft Access MDB file, which is rotated three times per day to avoid exceeding the 2-GB size limit of the Access file format. Thus, a single day's data can span multiple files. Each database file contains a table for each FDR within FNET. Table names are of the form "FRURawData*XXX*," where *XXX* is the unique identification number of the FDR in question. Each row in the table stores a single measurement point, which includes the frequency, phase angle, voltage magnitude, and timestamp. Additional FDR-specific information such as the number of locked-on GPS satellites and the latitude/longitude are stored at regular intervals in the *FirstFreq* field, which normally contains the results of the frequency estimation algorithm before resampling takes place. The *FirstFreq* field is not typically used in offline analysis.

Challenges Associated with FDR Data

FNET has used a variety of data storage conventions throughout its history. When the number of FDRs was very small, several days' data could be contained within one MDB file. Once more FDRs were deployed, it became necessary to create a new database file each day. Eventually, two files were needed, then three. These files are named deterministically based on the local (EST/EDT) date and time the file is created, which occurs at fixed times throughout the day according to UTC time. This schedule has changed over the years as more files were needed. As a result, the filenames for an event's data depend on what period the data were recorded, the rotation scheme in use at that time, and whether or not Daylight Saving Time (DST) was in effect. Thus, determination of the MDB file corresponding to a

particular time span is non-trivial, and is one of the primary difficulties encountered when working with the FNET data.

Once the correct MDB file has been found, the next step is to extract the necessary data into a usable form. A program was eventually developed to perform the location and extraction procedure. However, this program is somewhat difficult to use and can only output the data to ASCII text files, which are an inefficient means to store what is largely numerical content. Additionally, the current extraction program encodes timestamps in a non-standard fashion that requires further calculation to decode. Since most analysis work takes place in MATLAB, additional steps are needed to import the text files into the workspace. Thus, it is desirable to have a way of easily locating and extracting the data to MATLAB's native MAT format, or perhaps some other text-based format.

After extracting and importing the data, it must still be checked for the errors and inconsistencies that sometimes appear as a result of hardware, software, or network problems. For example, missing segments of data need to be flagged, and if desired, replaced with values interpolated from surrounding points. Duplicate timestamps must also be detected and corrected. Only after all of these steps have been completed are the data in a form that is suitable for plotting and analysis.

An Object-Oriented Perspective on FNET Data

Most modern computer programming languages are object-oriented; that is, they allow programmers to represent items as "objects" having particular attributes and actions. Attributes are simply properties of an object, such as color, weight, or temperature. Actions, or "methods," are functions that an object can perform [57]. For example, a balloon object might have properties like size, material, pressure, and color, and methods for inflating and popping. When writing software that models real-world phenomena, the object-oriented paradigm can greatly simplify the development process and encourage good programming practices.

Offline FNET data analysis programs reflect procedural design patterns more common in older software. For example, the data for an FDR might be loaded into several different vectors (frequency, angle, voltage, etc.) that are processed independently and have no connection with each other. While this approach is sufficient for dealing with a single FDR, the code quickly becomes unwieldy once multiple FDRs are introduced. The framework described here views the FDR as an object with measurement data "attached" to it. Multiple FDRs are stored in an array, making it possible to work with multiple measurement sets at once. Aside from their data, the FDRs each have unique properties such as their name, identification number, and map coordinates; the framework accounts for these as well.

Framework Overview

The FNET software framework was developed in MATLAB since that is the analysis tool most commonly used in the Power Information Technology Laboratory. It consists of a number of different classes and helper functions implemented as m-files that can be readily incorporated into a given project.

The basic unit of the framework is the *fdr* object, which contains both the FDR's measurement data as well as its unique properties. It provides a number of helpful methods used to prepare the data for analysis. Users can call the methods for each *fdr* object individually, or they can use the *fdrlist* object to perform operations *en masse*. A user includes the framework by first creating an *fdrlist* object to hold the *fdr* objects. The *fdr* objects are populated using functions that can read data from either text files or Access MDB files. The metadata for each FDR can be loaded directly from the FNET application server's MySQL database, which ensures that the information is up-to-date. Alternatively, the metadata can be imported from a configuration file should the application server be unavailable. Next, several functions can be called to check for bad data, interpolate missing data, and perform median filtering. Other functions allow the data to be easily exported to tab-delimited or comma-separated value (CSV) files. Finally, the framework allows the FDR data to be "replayed" over the network in the proprietary FNET message format. This feature enables the creation of "software" FDRs, which could be very useful for testing and developing real-time analysis algorithms for the FNET server.

Class Structure

The class structure for the FNET application framework is described below. Example programs are given in Appendix B.

fdr Class

Attributes:

- name – The FDR's given name, which usually reflects its location or host.
- id – The FDR's unique numerical identifier.
- city
- state
- interc – The power system interconnection where the FDR resides (e.g., EI, WECC, ERCOT).
- latitude – The latitude of the FDR's geographical location, in decimal degrees.
- longitude – The longitude of the FDR's geographical location, in decimal degrees.
- numSats – The number of currently locked-on GPS satellites.

- arrivalTime – A scalar to hold the time when the FDR observes an event.
- isRef – A Boolean value to signify if this unit is the reference unit.
- time – A vector containing UTC timestamps for measurements.
- frequency – A vector containing frequency measurements.
- angle – A vector containing angle measurements.
- voltage – A vector containing voltage measurements.
- port – The remote TCP port this FDR connects to on the server.
- unwrapped – A Boolean value to indicate if the angle vector has been unwrapped.
- normalized – A Boolean value to indicate if the angle vector has been normalized.

Methods:

- fdr(*name,id,latitude,longitude,city,state*) – A simple constructor for an FDR object.
- loadDataFile(*folder*) – Load data from an appropriately named text file in *folder* into the FDR object. This function looks for a text file matching the naming convention used by the existing data extraction utility, which is *YYYYMMDD-hhmmss-FDRNAMEnnn.txt* where *nnn* is theFDR's unique numerical identifier. The data are expected to be in format '2' produced by the extraction utility. One the data have been imported, the timestamps are converted into MATLAB serial date numbers and the 11-second firmware correction is applied.
- loadDataFromAccess(*start_date,start_time,duration,unit*) – Load data for this FDR from the appropriate Access database file given the starting date/time and duration. Different duration units can be specified, such as second, minute, or hour. This function assumes that FDR data are located in one or more hard drives attached to the computer. The database files should be placed in the root directory of the hard drive organized by the year and month they were created. For example, the files for data from August 2012 should be located in F:\2012\08. The function automatically determines which hard drive contains the data in question. One the data have been imported, the timestamps are converted into MATLAB serial date numbers and the 11-second firmware correction is applied.
- trimData(startTime,endTime) – Trim data vectors so that they only contain data from *startTime* to *endTime*. This function is useful if the data from different FDRs do not completely align.
- unwrapAngle() – Perform angle unwrapping. FDRs report voltage angle measurements in radians in the range [0, 2π]. Thus, the resulting waveform resembles a sawtooth-type signal, which is essentially useless for analysis purposes. Unwrapping the angle creates a smooth signal that is suitable for analysis and comparison with other units.

- normalizeAngle() – Subtract the first angle value from the angle vector as a whole.
- medianFilter(*n*) – Perform *n*-point median filtering on the frequency data.
- missIndex() – Return a vector containing the indices where there are gaps in the time vector.
- interpolateMissingData() – Fill in missing data for all data vectors using linear interpolation.
- getArrivalTime(*variable,threshold,direction*) – Determine the time when *variable* crosses *threshold* in *direction* (above/below), where *variable* could be frequency or angle, and *threshold* is some fixed value.
- checkTimestampDuplication() – Examine time vector for timestamp duplication and correct any duplications found.
- exportToFile(*folder,format*) – Export the FDR's data to a file in *folder* using the specified *format*. Both comma-separated and tab-delimited formats are supported. The resulting output uses conventional representations of dates and times that include the conversion interval and can be easily imported into another program. The FDR object itself can be saved in a .mat file for future use.
- sendData(*host,protocol,port,override,offset*) – Transmit the FDR's data over TCP/IP to *host* in the specified *protocol.* Only the native FNET protocol is presently supported, however IEEE C37.118 support could be added in the future. Specifying the *override* option will replace the loaded timestamps with the local system time. An optional *offset* (in hours) can be added to simulate Greenwich Mean Time, for example.

fdrlist Class

The *fdrlist* class acts as a wrapper for the *fdr* class. Most of its functions simply call the corresponding *fdr* class methods for each FDR in the list. This reduces the number of *for* loops that need to be written when dealing with a large number of FDRs.

Attributes:

- referenceUnit – The numerical identifier of the reference unit for the list.
- newestStartTime – The most recent starting timestamp found in the time vectors for each FDR comprising the list.
- oldestEnd Time – The oldest ending timestamp found in the time vectors for each FDR comprising the list.
- units – A cell array to hold the *fdr* objects.

Methods:

- buildListFromFiles(*dataDir*,*[configFilename]*) – Create list of FDR objects based on data files located in *dataDir*. This function first creates a list of *fdr* objects and obtains their metadata from a database by default. If a filename is specified as the third argument, the FDR metadata will be loaded from there instead. It then loads the data using each FDR's loadDataFile() function.
- buildListFromArray(*idArray,method,[filename]*) – Build a cell array of *fdr* objects from an array containing their ID numbers. The FDR metadata can be loaded from either a database or a configuration file by specifying 'database' or 'configfile' for the *method* argument. If the configuration file option is used, a filename must be specified as the last argument.
- getStartEndTimes() – Determine the newest start time and the oldest ending time across all the FDRs in the list. This information is used for the trimData() function.
- getArrivalTime(*var,threshold,direction*) – A wrapper function for *fdr* objects in the list.
- relativeArrivalTime() – Subtract the earliest arrival time from each unit's arrival time so as to determine the relative arrival times.
- getSize() – Return the number of *fdr* objects in the list.
- checkForMissingData() - A wrapper function for *fdr* objects in the list.
- interpolateMissingData() - A wrapper function for *fdr* objects in the list.
- unwrapAngle() - A wrapper function for *fdr* objects in the list.
- normalizeAngle() - A wrapper function for *fdr* objects in the list.
- checkTimestampDuplication() - A wrapper function for *fdr* objects in the list.
- trimData() - A wrapper function for *fdr* objects in the list.
- medianFilter(*n*) - A wrapper function for *fdr* objects in the list.
- subtractReferenceAngle() – Subtract the reference FDR's angle vector from the other FDR's angle vectors.
- referenceUnit(*id*) – Set the reference unit for the list. This function performs the necessary preprocessing such as interpolation, angle unwrapping and normalization before calling subtractReferenceAngle() .
- exportToFile(*folder,formatString*) - A wrapper function for *fdr* objects in the list.

FNETEvent Class

Attributes:

- dateStamp – The UTC date of the event.

- timeStamp – The UTC time of the event.
- location – The interconnection where the event occurred.
- eventType – The type of event that occurred, e.g., generator trip, load shedding, etc.

Helper Functions

Several helper functions are included in the framework. They are:

- getConfigFromFile(*id,filename*) – Look up an FDR's information from an FNET UnitConfig file given its ID number. This function returns the name, port, latitude, longitude and interconnection.
- getDataTables(*con*) – Return cell array of FDR data tables in database with connection handle *con*.
- getFileByType(*folder,filter*) – Return a cell array containing a list of files matching *filter* in the given *folder*, where *filter* supports normal Windows wildcard expansions like *.txt, etc.
- getFolders(*directory*) – Return a cell array of folders in *directory*.
- getMDBFileName(*date,time,duration,unit*) – Determines the filename of the Access database for an event occurring at *date* and *time* lasting for *duration* expressed in terms of *unit*. This function also returns the starting and ending timestamps necessary to query the data.
- getMDBFiles(*path*) – Return cell array of .MDB files in *path*.
- loadFDRInfoFromDB(idArray) – Return a cell array of *fdr* objects populated with their ID number, latitude, longitude, city, and state for each FDR ID contained in *idArray*. FDR information is obtained from the FNET application server database.
- whichDataDrive(*date*) – Determines the folder containing Access database file for an event occurring on *date*.

Conclusion

This chapter has presented a MATLAB-based object-oriented software framework for the development of FDR data analysis applications. It greatly simplifies the process of locating, importing, preparing, and exporting the FNET data and dramatically reduces the amount of code needed to be written. Although designed specifically for FNET, the framework could easily be extended to support PMU data as well. In addition, virtual FDRs can be created that emulate devices in the field. By removing barriers to working with the data, the framework allows users to focus almost exclusively on research and development.

7. Environmental Regulation Impacts on Eastern Interconnection Performance

Introduction

The U.S. Environmental Protection Agency (EPA) recently finalized the Mercury and Air Toxics Standards (MATS) and the Cross-State Air Pollution Rule (CSAPR), which are regulations designed to reduce power plant emissions such as mercury, NO_x, SO_2, and ozone [29, 30]. Assuming these rules pass judicial review, as much as 30 GW of generation capacity (mainly coal and oil-fired units) will be taken offline within the next few years [31], mainly from the Eastern Interconnection. Most of this lost capacity is being replaced with natural gas-fired generation, such as gas turbines and combined-cycle plants. Since power injections are being removed from some points in the grid and added to others, the flow of power will be altered, which could have important implications for voltage stability and equipment ratings. This chapter presents a study designed to simulate and quantify the effects of MATS/CSAPR-related generator deactivations on bus voltages and transmission line flows in the Eastern Interconnection over the next few years.

Mercury and Air Toxics Standards

According to the Environmental Protection Agency, coal and oil-fired power plants are responsible for approximately half of the airborne mercury emissions in the United States [30]. Upon entering water, biological processes convert the metal to methylmercury, an even more toxic substance that bioaccumulates in fish and other aquatic wildlife. Consumption of methylmercury-contaminated fish by pregnant women and children is particularly dangerous since it can affect nervous system development. In addition, coal and oil-fired power plants release toxic metals such as arsenic, chromium, and nickel, which are believed to be carcinogenic. The Mercury and Air Toxics Standards establish limits on the amounts of these substances that can be released into the environment, while also creating work practices designed to reduce emissions of organic air toxics such as dioxins and furans.

Power plants can limit their emissions by employing any of several different pollution control technologies, such as flue gas desulfurization (FGD), activated carbon injection (AJI), or fabric filtration [58]. However, retrofitting existing plants with these controls is both expensive and time consuming. In many cases, it would be uneconomical for plant owners to bring their generating stations into compliance. Thus, MATS will effectively shutter a large number of coal and oil-fired power plants.

Cross-State Air Pollution Rule (CSAPR)

In July of 2011, the EPA finalized the Cross-State Air Pollution Rule (CSAPR), which applies to 23 states located within the Eastern Interconnection. CSAPR replaces the Clean Air Interstate Rule, and aims to reduce SO_2, NO_x, ozone, and fine particulate emissions produced by power plants. The EPA cites a variety of health and environmental benefits resulting from the rule, including the annual avoidance of:

- 13,000-34,000 premature deaths
- 15,000 nonfatal heart attacks
- 19,000 emergency room visits
- 1.8 million lost work days or school absences
- 400,000 asthma attacks [29]

One of the most striking aspects of CSAPR is the speed at which the EPA expects compliance. The first reductions were scheduled to begin in January of 2012, with total implementation being achieved in 2014. The rule is currently being challenged by utilities in the federal court system, so the actual implementation timeline and number of affected plants is unknown. However, some utilities have already begun the process of retiring older plants, regardless of whether or not CSAPR takes effect [59-63].

Steady-State Analysis

Electrical power is unique compared to other commodities in that it must be consumed at the same time it is produced. Complicating matters, electrical energy is difficult to store in large quantities, and cannot be easily routed along a particular transmission path. Rather, the flow of power is determined by the structure of the system itself [64]. It is this particular property that makes it difficult for significant topology changes to be made to the system without dramatically affecting the line flows and bus voltages.

Power flow (also called “load flow”) studies are the primary means of steady-state analysis used by system planners, and form the basis for dynamics studies as well. Given a model of the system and the loads it supplies, the load flow computes the load bus voltages, line flows, and generator bus angles. These quantities can then be used to determine which system components are operating outside of their limits. Because the power flow problem is non-linear, numerical techniques such as various forms of the Newton-Raphson method are employed to find the solution. Several commercial software packages exist for this purpose, including PSS/E, PowerWorld, and PSLF.

Environmental regulations such as MATS and CSAPR will most likely result in a large number of power plants being shut down. This represents a significant loss of generation capacity, which could be as small as 14.5 GW, or as large as 30 GW [31]. (Estimates of the actual amount vary depending on who is doing the calculation.) Obviously, this capacity must be replaced by some other means. Recent advancements in drilling technology have unlocked previously uneconomical natural gas reserves, most notably the Marcellus Shale. The drop in natural gas prices created by this additional supply has led many utilities to invest in gas-fired generating units, which can be quickly built at a much lower cost than nuclear plants [65]. One important constraint for gas plants, however, is that they must be built near both major gas pipelines and high-voltage transmission lines.

The shutdown of so many large coal and oil-fired units and their replacement with gas-fired generators could present some significant challenges to the grid. New plants will in many cases be built in different locations than the generators being shut down and will thus alter the topology of the system. As a result, the flow of power will be different than it is now. For example, lines that are presently operating below their capacity could become congested, which would have implications for system reliability and the locational marginal price (LMP) of electricity. Other areas could see voltage problems. NERC studied the potential impacts of the draft regulations from the perspective of reserve adequacy [32], but has not examined the possible steady-state consequences. Thus, there is a clear need for this type of study.

Study Design

The Institute for Energy Research (IER) has compiled a list of power plants that it claims will likely be shut down as a result of MATS and CSAPR, which can be found in Appendix C [31]. A map of these plants is given in Fig. 7.1. It should be noted here that IER is tied to the energy industry and should not be considered a completely unbiased source [66]. But, because IER would presumably be very conservative in their estimation of the affected plants, their list can be thought of as a worst-case scenario. Using this list and a 29,000-bus model of the Eastern Interconnection, the proposed regulations' effects on the grid under a variety of different conditions were studied to identify those regions that would be negatively impacted.

Before any simulations could be performed, it was necessary to first develop the information infrastructure needed to create, manage, and analyze a large number of simulation cases. The first portion of this task involved creating a database table of all generators that could conceivably be added or removed from the model. For the affected generators, the IER list mentioned earlier was used to populate the table. In a few cases, this list was augmented or corrected by media reports issued after its

publication. Data for new generation facilities and nuclear power plants were obtained from the 2010 U.S. Energy Information Agency Form 860 data file [67] and the Nuclear Regulatory Agency [68], respectively. The table structure for the generator data is given in Table 7.1.

Table 7.1: Description of generator database table.

Field	Description
generatorId	Primary key; a unique identifier for this generator independent of PSS/E
fuelType	Numerical indicator to represent fuel type
busnum	Corresponding bus number in PSS/E. A foreign key tied to the 'bus' table.
unitId	Corresponding unit ID in PSS/E
capacity	Summer capacity, in MW
plantName	Name of the plant
startYear	Year the plant went into operation, if known
endYear	Year the plant is scheduled to be taken offline
lat	Latitude
lng	Longitude
city	City
state	State
inModel	Boolean variable indicating if the generator exists in the model or not
syncon	Not used

Additional tables were created to store data related to the lines, buses, areas, and zones found in the PSS/E model. Foreign key constraints were applied to maintain data consistency and integrity. These constraints made it impossible to, for example, link a generator to a bus that did not exist in the model.

The Eastern Interconnection model used in this study represents the Summer 2015 peak load case. It was developed by the Multiregional Modeling Working Group (MMWG) in 2010, and contains approximately 29,000 buses and 4,000 generators. Most of the EI system is included in the model, however Florida and some parts of the extreme northeast have been removed. Although the model contains the necessary mathematical attributes describing each system component, it does not include their geographical information such as latitude and longitude. Fortunately, the Energy Visuals models available for Power World include this data for most of the buses in the system (Fig. 7.2). Since the bus numbers in both models are generally the same, the latitudes and longitudes could easily be merged using SPSS and added to the database.

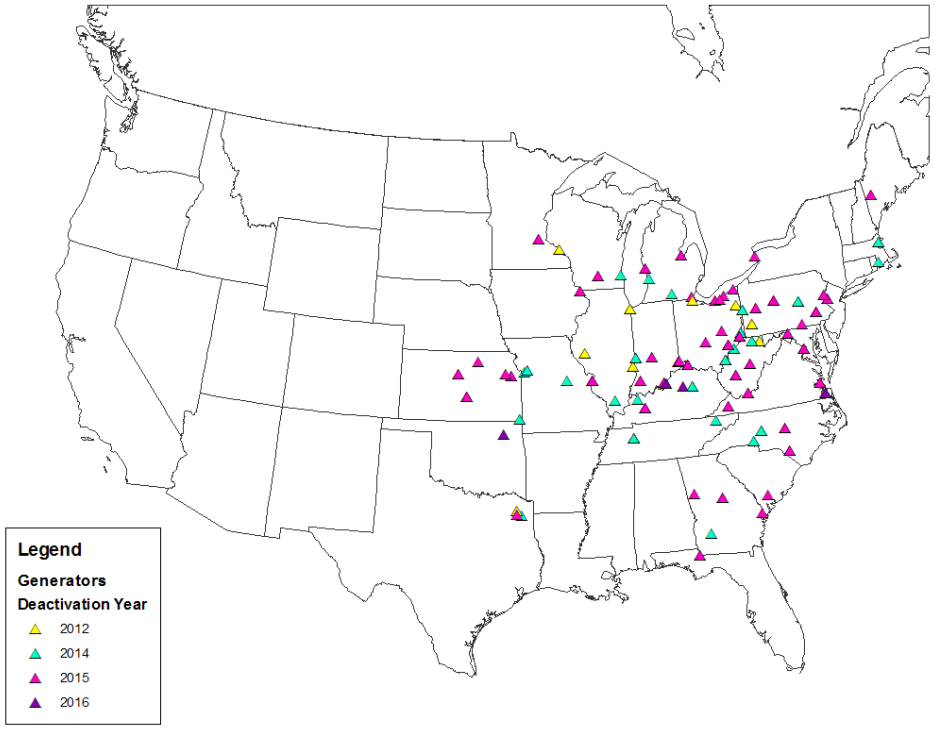


Fig. 7.1: Map of generators by deactivation year

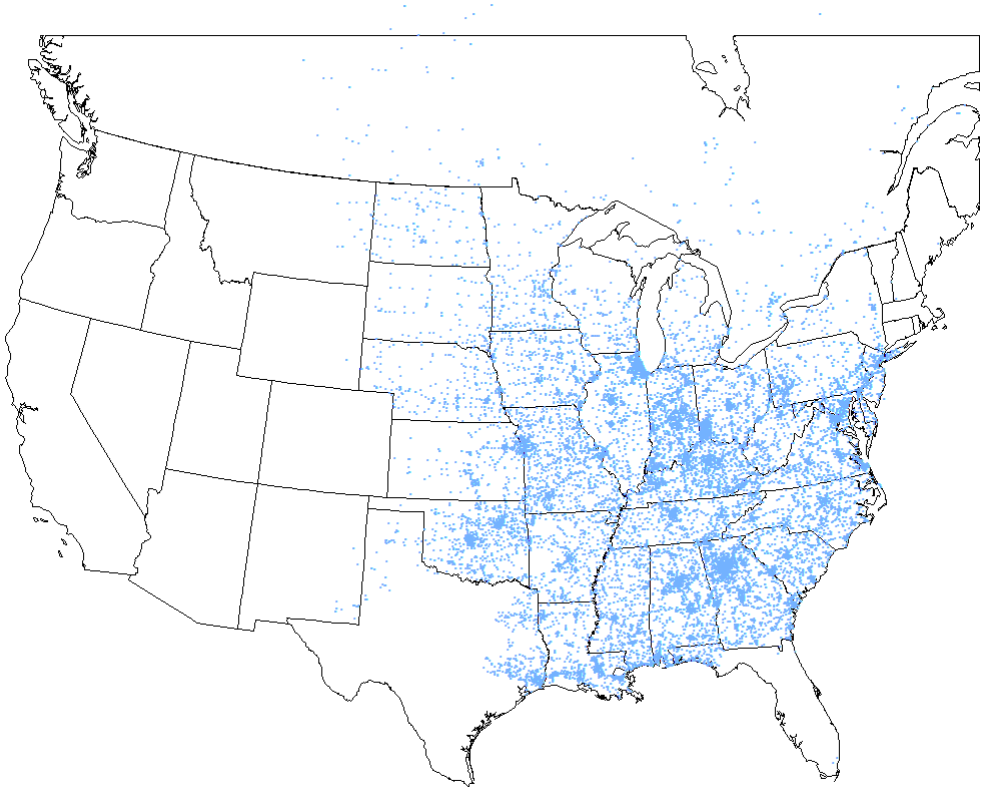


Fig. 7.2: Bus locations in 2015 MMWG EI model

Because the objective of this research was to quantify the effects of adding and removing different generators from the system, the next step was to determine which machines in the model (as identified by a bus number and two-character unit identifier) corresponded to those found in the generator table. Several of the MATS/CSAPR-affected generators had already been located during previous studies, which greatly simplified this task. However, some generators (particularly new ones) still needed to be located. This was done by manually searching for the plant name in the Energy Visuals model to yield the associated bus number, which was then matched against the machines found in the PSS/E model. In all, 151 of the 223 generators identified by IER and media reports were successfully located using this method.

Generators that could not be found, as well as new units that needed to be added to the model, presented a unique challenge. In order to adequately model the power flow, each unit must be connected to a bus in the model that approximates its physical location in the system. This was done by utilizing the Google Maps API in Python to convert each generator's city and state to a latitude and longitude, which was then used to find the nearest high voltage (>100 kV) bus in the model.

To avoid double-counting generators, the PSS/E model needed to be checked to ensure that the new units were not already included. This was done by using the Jaro-Winkler string similarity measure to compare the 12-character PSS/E bus names to the complete plant names found in the database. The Jaro-Winkler distance was chosen because it is well-suited for short strings and provides a similarity between zero and one, with one being a perfect match [69]. To limit the number of comparisons that needed to be made, the plant names were only matched against buses within 0.5 degrees of latitude and longitude. A similarity of at least 0.7 was required for a bus to be considered as a match candidate. The list of possible matches was then reviewed manually and the generator database was updated to reflect any duplications found.

Simulation Scenarios

For this research, sixty different scenarios were developed that reflect the possible evolution of the Eastern Interconnection in the next few years. These scenarios are based on projected generator deactivations, forecasted demand growth, and likely construction of new generating capacity. Table 7.2 describes several base cases that serve as the experimental controls. Each cell contains the unique numerical identifier assigned to a particular scenario. The first row represents the "do-nothing" or "business as usual" case, that is, it assumes that no generators are taken offline, but includes expected load growth. The next row represents the removal of generators as a result of MATS and CSAPR. Row three describes scenarios where some of the affected generators are instead used as synchronous

condensers[2]. Summer demand forecasts for the Eastern Interconnection from 2012 through 2017 are given in Table 7.3 [70]. Since the model used in this study lumps some portions of the system together, its load must be scaled proportionally as given in Table 7.4.

Table 7.2: Base Cases

	2012	2013	2014	2015	2016	2017
Anticipated load increases only, all generators in place	1	23	27	7	17	14
With MATS/CSAPR coal-fired generators removed	2	24	28	10	16	15
With synchronous condensers added	50	51	52	49	53	54

Table 7.3: Forecasted Summer Demand for the Eastern Interconnection

	Summer Demand, MW					
Region	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**
FRCC	51,499	52,645	53,641	54,862	56,100	57,346
MRO (US)	48,009	48,786	49,536	50,288	51,101	51,799
NPCC (US)	66,219	66,952	67,604	68,210	68,758	69,299
RFC	195,700	198,400	201,100	203,600	206,200	208,600
SERC	221,590	225,650	230,208	234,597	238,792	243,056
SPP	47,012	47,715	48,428	49,152	49,876	50,640
MRO (CA)	6,650	6,717	6,780	6,763	6,821	6,869
NPCC (CA)	50,392	50,476	50,546	50,347	50,452	50,655
Total (MW)	**687,071**	**697,341**	**707,843**	**717,819**	**728,100**	**738,264**

Table 7.4: Adjusted Demand for the 29,000-bus EI Model

	2012	2013	2014	2015	2016	2017
Total Load (MW)	554,798	563,120	571,567	580,140	588,842	597,675

Much of the generation capacity lost due to CSAPR and MATS will be replaced with gas turbines or combined-cycle plants. According to the U.S. Energy Information Agency, approximately 21.5 GW of gas-fired generation will be built between now and 2017 [71]. A list of planned gas power plants can be found in Appendix D. The scenarios given in the first row of Table 7.5 represent the projected installations of new gas turbines given anticipated increases in load. The second row includes the 8 GW

[2] A synchronous condenser is a synchronous machine operated in such a way that it can either produce or absorb reactive power without consuming real power. To the system, it can appear as either a capacitor or inductor. Synchronous condensers are used to provide power factor correction, which decreases line losses and improves the voltage profile.

of wind turbine generation planned in the next five years. A list of planned wind power installations can be found in Appendix E.

Table 7.5: Gas Base Cases

	2012	2013	2014	2015	2016	2017
With no wind	3	25	29	12	18	20
With planned wind buildout	22	26	30	13	19	21
Gas Base Case with synchronous condensers added	55	56	57	58	59	60

Although fossil fuels will continue to supply a large portion of the electrical demand in the United States for the foreseeable future, concerns over global warming and future environmental regulations have led to a renewed interest in nuclear power [72]. The U.S. Nuclear Regulatory Commission currently has 12 applications for new reactors under consideration [73]. Of these, six are scheduled for completion in the next few years [71, 74]. A list of these plants can be found in Appendix F. The nuclear base case described in the first row of Table 7.6 includes the new gas-fired generation from Table 7.5 and assumes that these reactors are finished on schedule, and that the license renewals for existing reactors are granted. Cases in the second row include planned wind generation. Finally, the last row describes scenarios where all nuclear plants are removed from the system, similar to what is being planned in Germany and Switzerland [75, 76].

Table 7.6: Nuclear Base Cases

	2012	2013	2014	2015	2016	2017
With no wind (includes gas)	31	32	33	34	35	36
With planned wind buildout	37	38	39	40	41	42
Nuclear Base Case with synchronous condensers added	61	62	63	64	65	66
With all nuclear plants removed	43	44	45	46	47	48

Each case was then described using a number of different parameters, which were then stored in a database table. The table structure for the case descriptions can be found in Table 7.7.

Table 7.7: Case table description

Field	Description
caseId	Primary key, a unique numerical identifier for the case
gensRemoved	Boolean to indicate if MATS/CSAPR-affected generators should be disabled
synCon	Boolean to indicate if synchronous condenser conversion should be performed
wind	Boolean to indicate if wind generation should be added
gas	Boolean to indicate if gas-fired generation should be added

nukesRemoved	Boolean to indicate if existing nuclear reactors should be removed
addNukes	Boolean to indicate if planned nuclear reactors should be added
year	The year this case represents
baseCaseFile	The previously saved PSS/E model file that this case should modify

Simulation Procedure

The large number of generators (487) involved in this study made it virtually impossible for the PSS/E model to be modified by hand. Thus, the PSS/E Python API was used to automate the process of adding and removing generators from the system and running the power flow. A Python script was written to load the case description from the database along with its corresponding base case model. Next, the script removed each generator from the IER list one-at-time, increasing the remaining generation to compensate accordingly, and solved the power flow. If the power flow successfully converged, it then saved a backup file to serve as a restore point for further simulations. If the power flow failed to converge after removing the generator, it was turned into a synchronous condenser by setting its real power output to zero. The power flow solution was then re-attempted. In the event of an unsuccessful simulation, the last restore point was reloaded and the simulation continued from that point. New generators were added by first scaling the existing generation amount down by the capacity of the new unit, followed by the creation of a 22-kV generator bus with a transformer connected to the previously identified high voltage bus. Next, the generator was added to the 22-kV bus using the parameters calculated according to Table 7.8. In order to improve convergence, the per-unit voltage setpoint of the generator was set to be equal to that of the associated high-voltage bus. For this study, it was assumed that each generator was designed to operate at a power factor of 0.8. The power flow was then solved before adding the next unit. This procedure was repeated for each of the 60 simulation cases.

Table 7.8: Generator Parameter Calculations and Descriptions

Parameter	Calculation	Description
P_{GEN}	capacity × 0.95	Scheduled real power output, MW
P_{MIN}	capacity × 0.20	Minimum real power generation capacity, MW
P_{MAX}	capacity	Maximum real power generation capacity, MW
Q_{MAX}	(capacity/0.8) × 0.6	Maximum reactive power generation capacity, MVAr
Q_{MIN}	-(capacity/0.8) × 0.6	Minimum reactive power generation capacity, MVAr
M_{BASE}	capacity/0.8	Machine base apparent power, MVA

Upon completion of the simulations, a Python script was used to extract the high-voltage (>230 kV) buses and lines exceeding their voltage and current ratings, respectively. This information was recorded in the database tables named *busresult* and *branchresult*, which were then imported into ArcGIS to create intuitive visualizations of each scenario's results.

After performing the initial group of simulations where a large number of MATS/CSAPR-affected generators were removed and replaced with gas and wind generators, the bus voltages were used to locate areas in the system where synchronous condensers might be needed. Deactivated generators whose capacities were greater than 100 MW and within 50 miles of an out-of-limit bus were set as synchronous condensers, and the power flow studies were re-run.

Once all simulations had been performed, the converged cases were checked against the case descriptions found in Appendix G and summarized using a Python script. Deviations from the prescribed scenario, swing bus output, and the number of out-of-limit bus voltages and overloaded lines can be found in Attachment A. In general, deviations resulted when a particular machine could not be removed from the model without causing it to diverge during simulation. Since the goal of this study was to model the changes in the grid as accurately as possible, these machines were left in the model so that the effects of removing the remaining generators could be examined.

Results

The following maps show the out-of-limit bus voltages and overloaded transmission lines resulting from each simulation, which are overlaid on the high voltage transmission grid of the Eastern Interconnection. The cases are grouped by the year they represent in the order given by Table 7.2, Table 7.5, and Table 7.6. Due to the fact that no new nuclear plants are scheduled to begin operation between now and 2014, the cases described in the first three rows of Table 7.6 for those years are not shown, since they duplicate the corresponding cases in Table 7.5.

2012

Fig. 7.3 shows simulation results from the 2012 business-as-usual case where no generators are deactivated and no additional generating capacity is added. It is apparent that many bus voltages are above 1.05 p.u., a trend that is visible in all the simulations performed for this study. This is largely due to the fact that the model used in these simulations is configured for the high-voltage transmission system buses to be slightly above their nominal ratings so that a given contingency will not result in low voltages that might violate NERC standards [77]. When the 23 MATS/CSAPR-affected generators are removed (Fig. 7.4), little, if any change is observed, possibly due to the relatively small amount of generation (4,240 MW) being removed from the system that year. However, it was found that the Bay Shore Power Plant in Oregon, Ohio needed to be operated as a synchronous condenser in order for the remaining MATS/CSAPR generators to be removed; this was required for many of the subsequent

simulations. Since very few of the affected generators are located near the out-of-limit buses, only a few were converted into synchronous condensers (Fig. 7.5); their effect appears to be negligible.

The voltage profile appears to be relatively stable when 6,580 MW of new gas-fired generation (Fig. 7.6) and 1,526 MW of wind turbines were added (Fig. 7.7). As before, the addition of synchronous condensers did not make much difference (Fig. 7.8). For the final scenario involving the removal of all nuclear generators, 48 out of the 92 reactors were successfully disabled before the simulation failed to converge. This scenario resulted in a general lowering of bus voltages, with one bus being below 0.95 p.u. as shown in Fig. 7.9. Also, alterations made to the system did not have much effect on the number of overloaded lines, which stayed essentially constant for each of the 2012 scenarios. Because the geographical information for these lines could not be found, they are not shown on the maps below.

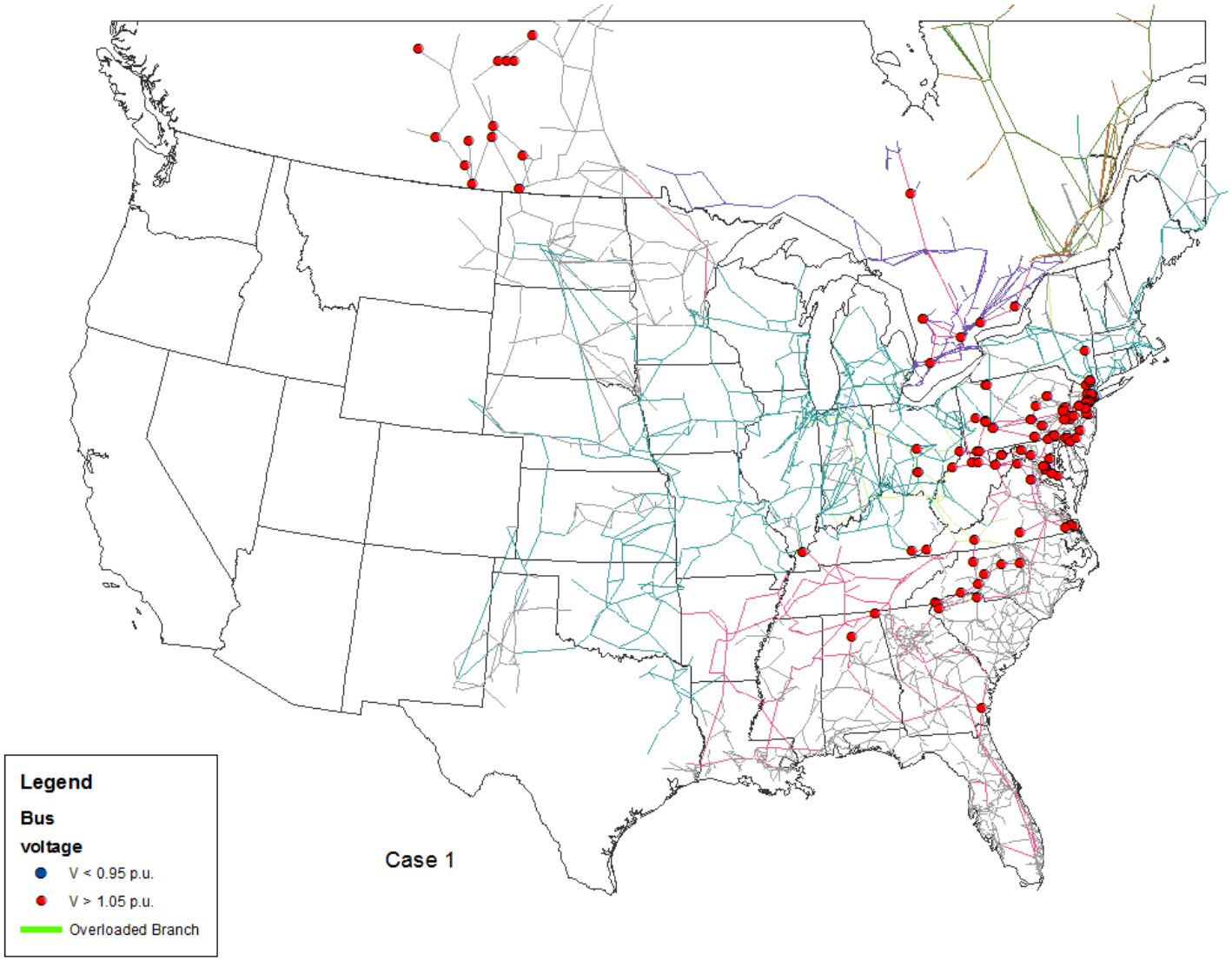


Fig. 7.3: 2012 base case

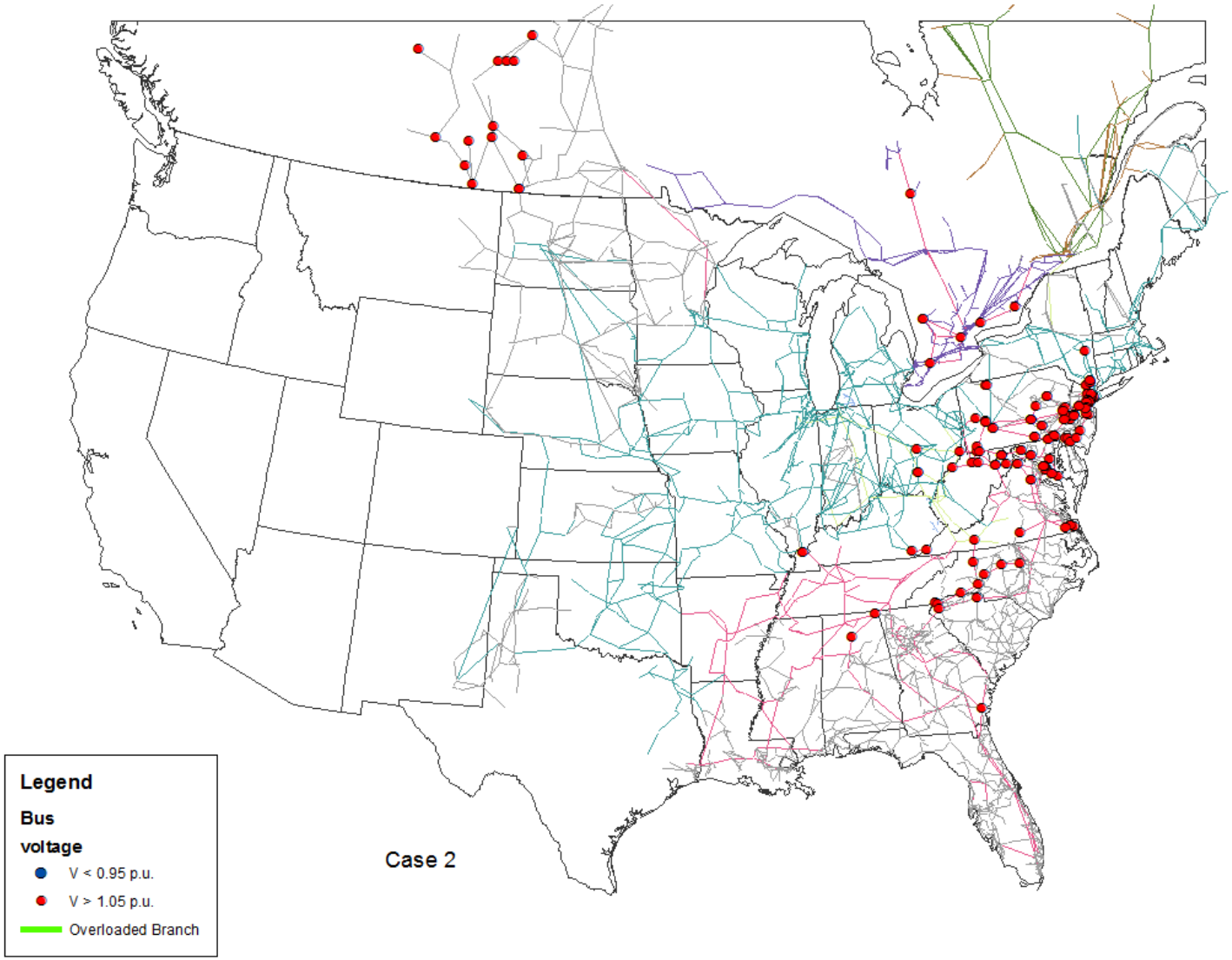


Fig. 7.4: 2012 base case with generators removed

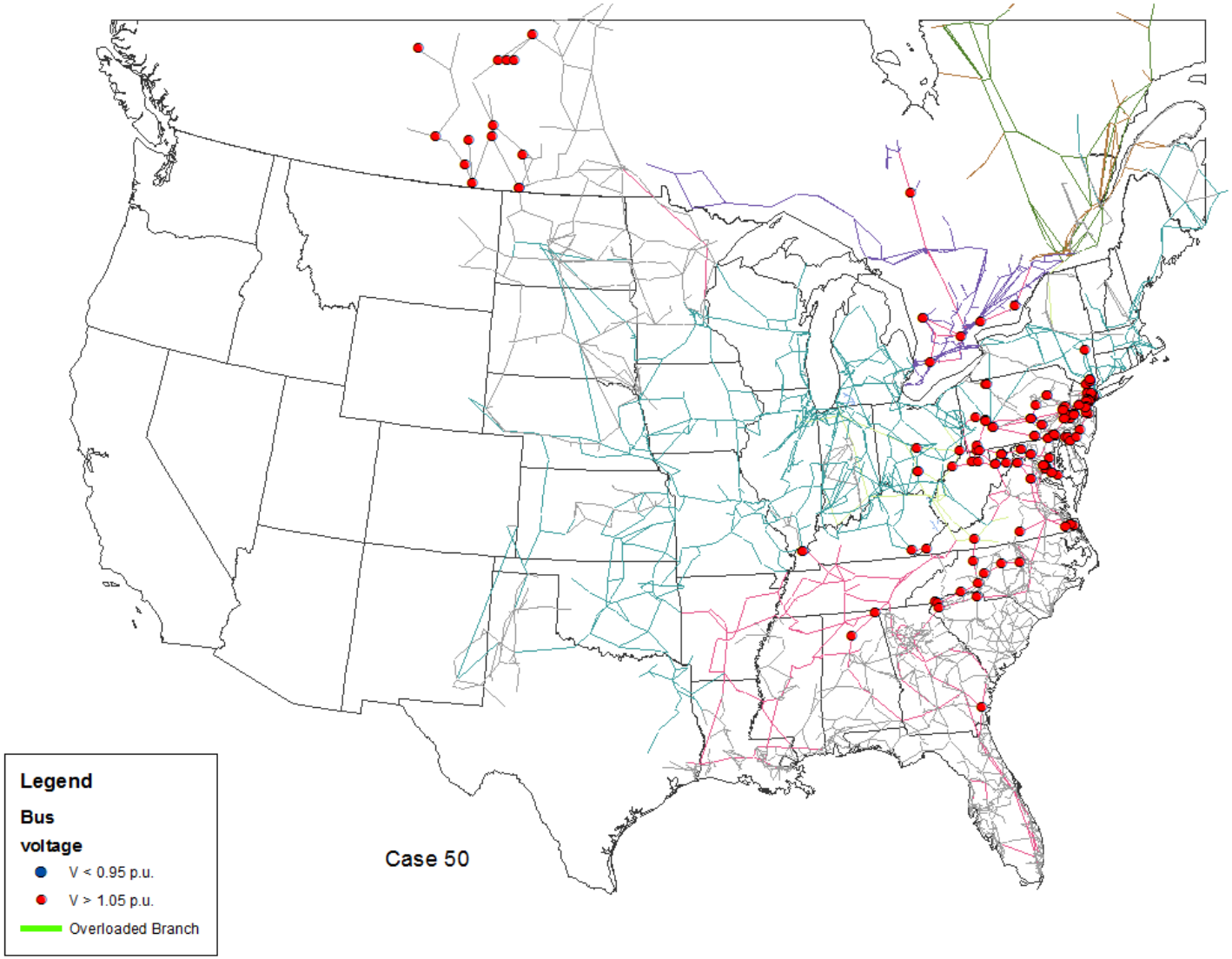


Fig. 7.5: 2012 base case with generators removed, synchronous condensers added

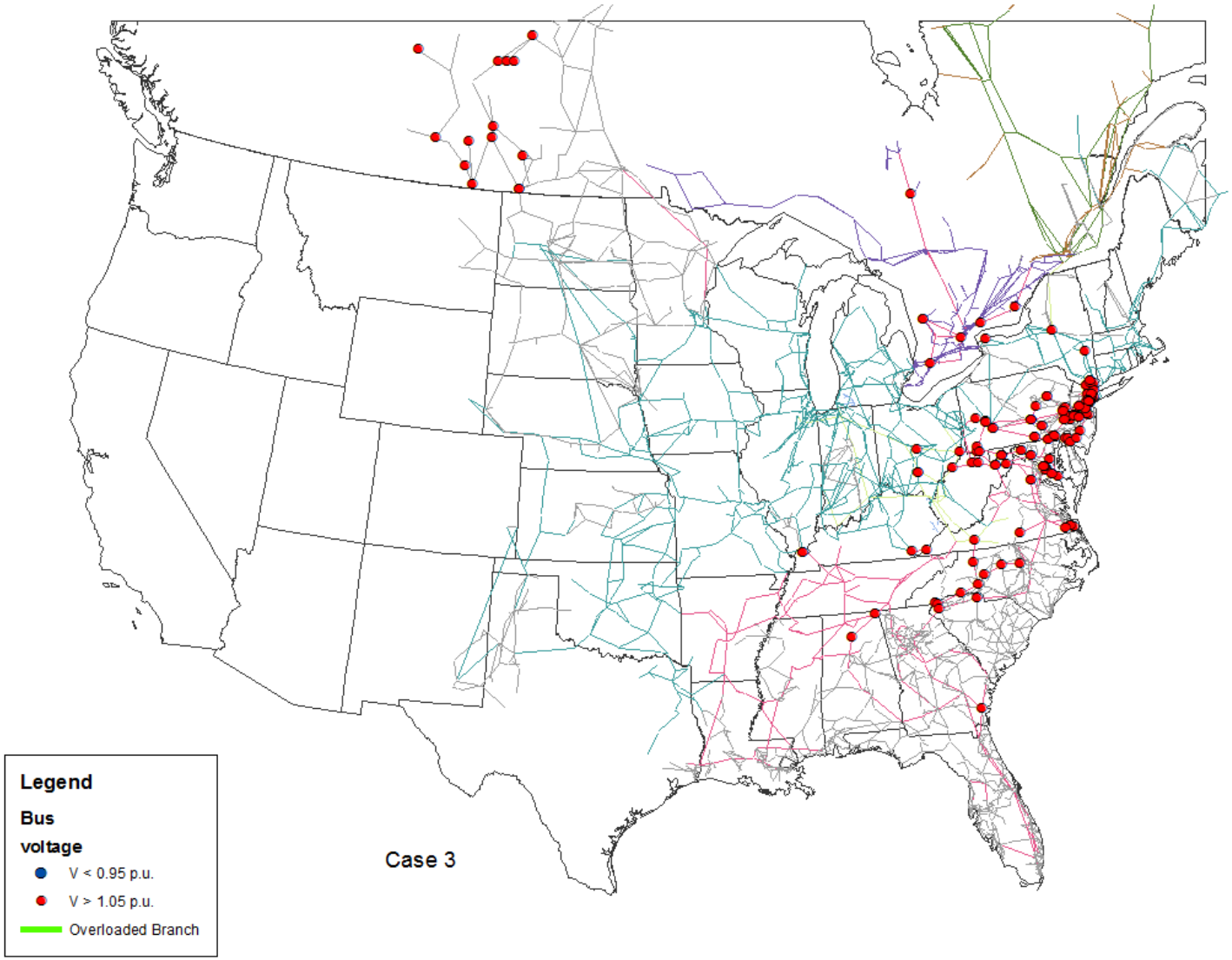


Fig. 7.6: 2012 gas base case

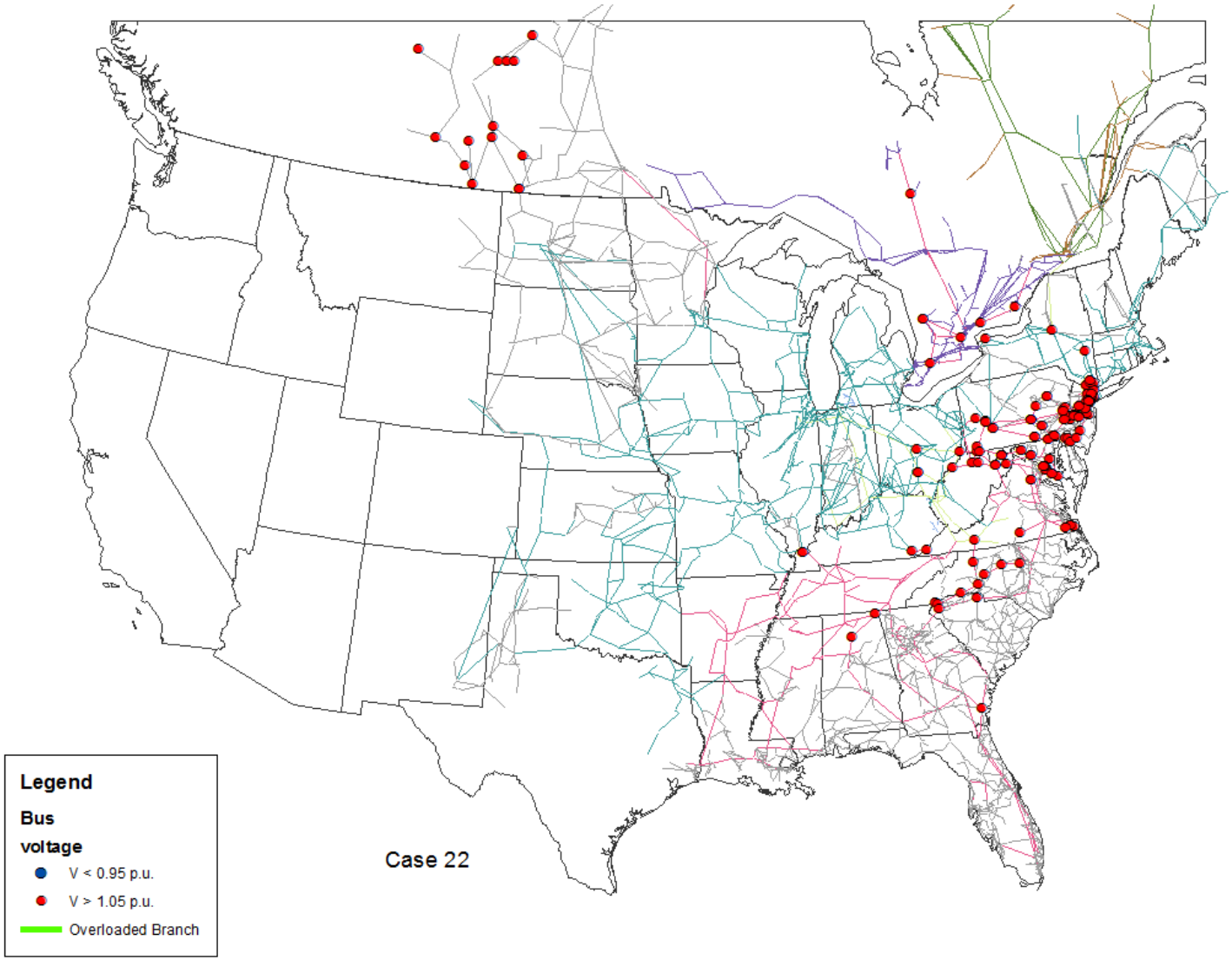


Fig. 7.7: 2012 gas base case with wind

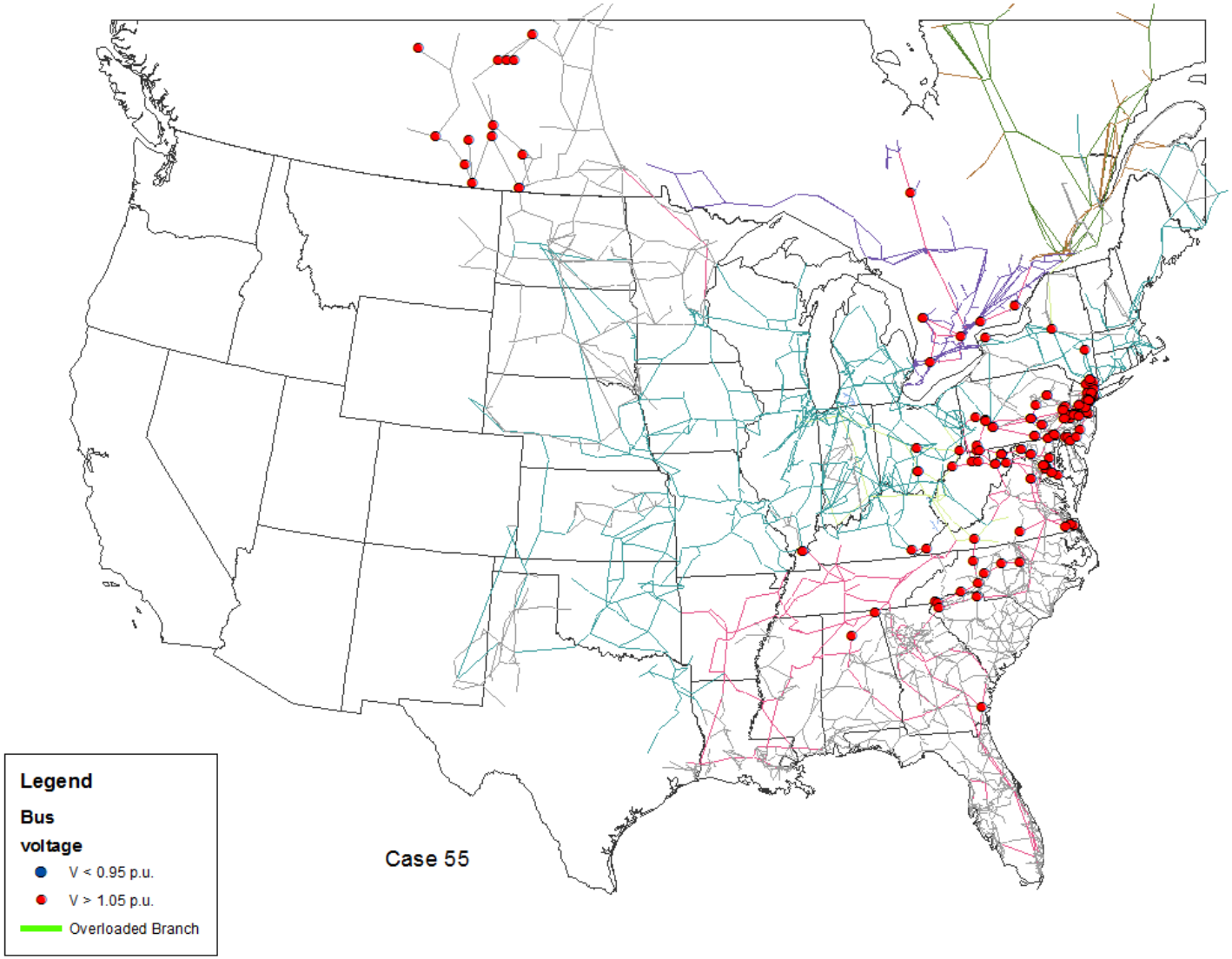


Fig. 7.8: 2012 gas base case with wind generation and synchronous condensers

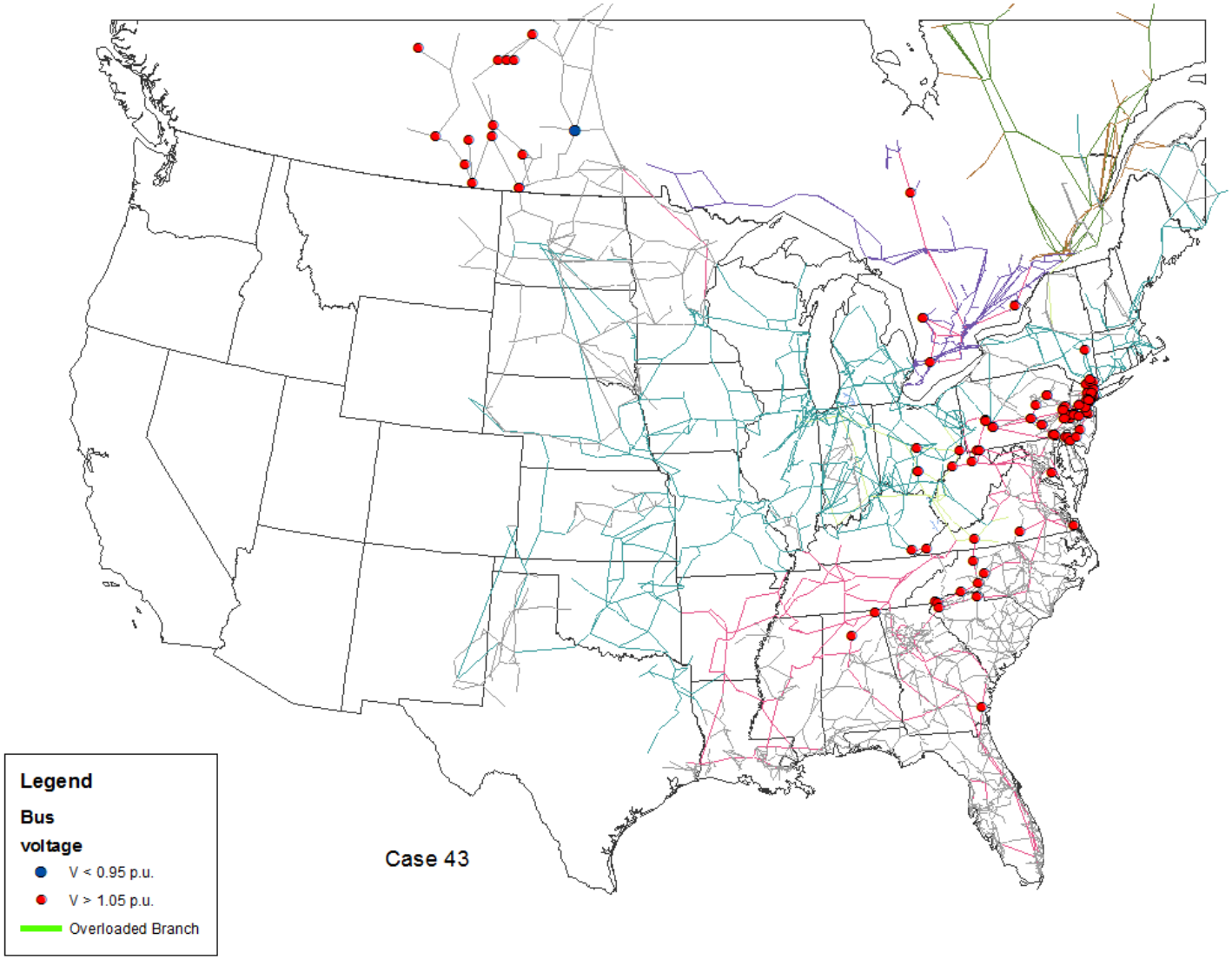


Fig. 7.9: 2012 gas and wind added, nuclear plants removed

2013

Simulation results from the 2013 base case are shown in Fig. 7.10. Like the 2012 base case, it shows a large number of overvoltage buses, though not quite as many. The 1.5% load growth from the previous year included in this case is probably responsible for the slight decline in bus voltages. No generators are due to be taken offline in 2013 as a result of MATS/CSAPR, thus Fig. 7.11 strongly resembles the corresponding figure from the 2012 scenarios. The conversion of some of the previously removed generators to synchronous condensers did not significantly affect the bus voltages (Fig. 7.12).

Only about 2,000 MW of gas-fired generation are to be added to the EI in 2013. This appears to add about 30 overvoltage buses to the system (Fig. 7.13) compared to the base case. The addition of 1,276 MW of wind generation (Fig. 7.14) has a negligible effect on the system, as does the conversion of some of the deactivated generators to synchronous condensers (Fig. 7.15). Only 38 nuclear reactors could be removed before the system failed to converge (Fig. 7.16), which is fewer than in the previous year. In most cases, only three high-voltage transmission lines were operated above their Rate B limits, the only exception being when the nuclear reactors were removed, leading to five overloaded lines.

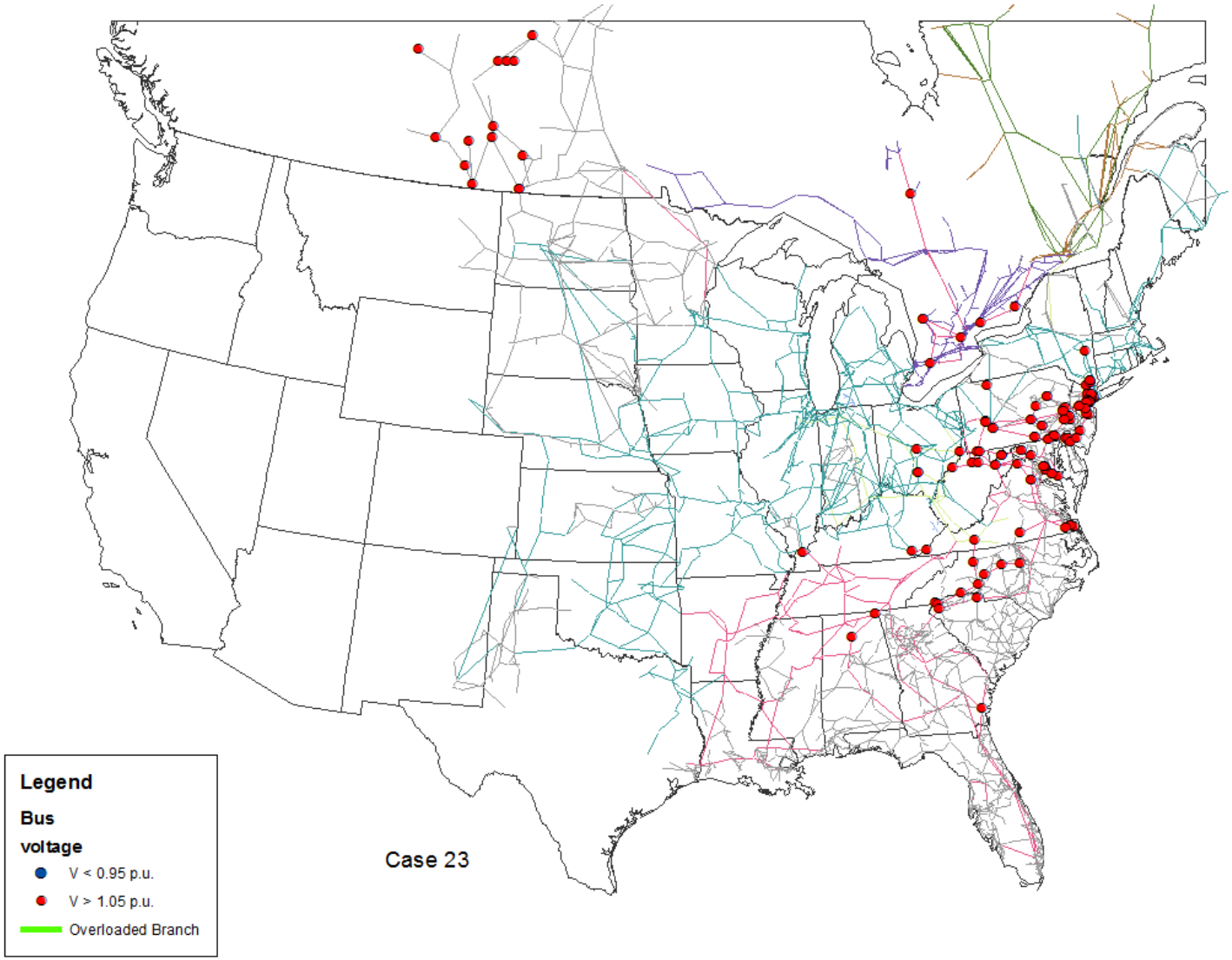


Fig. 7.10: 2013 base case

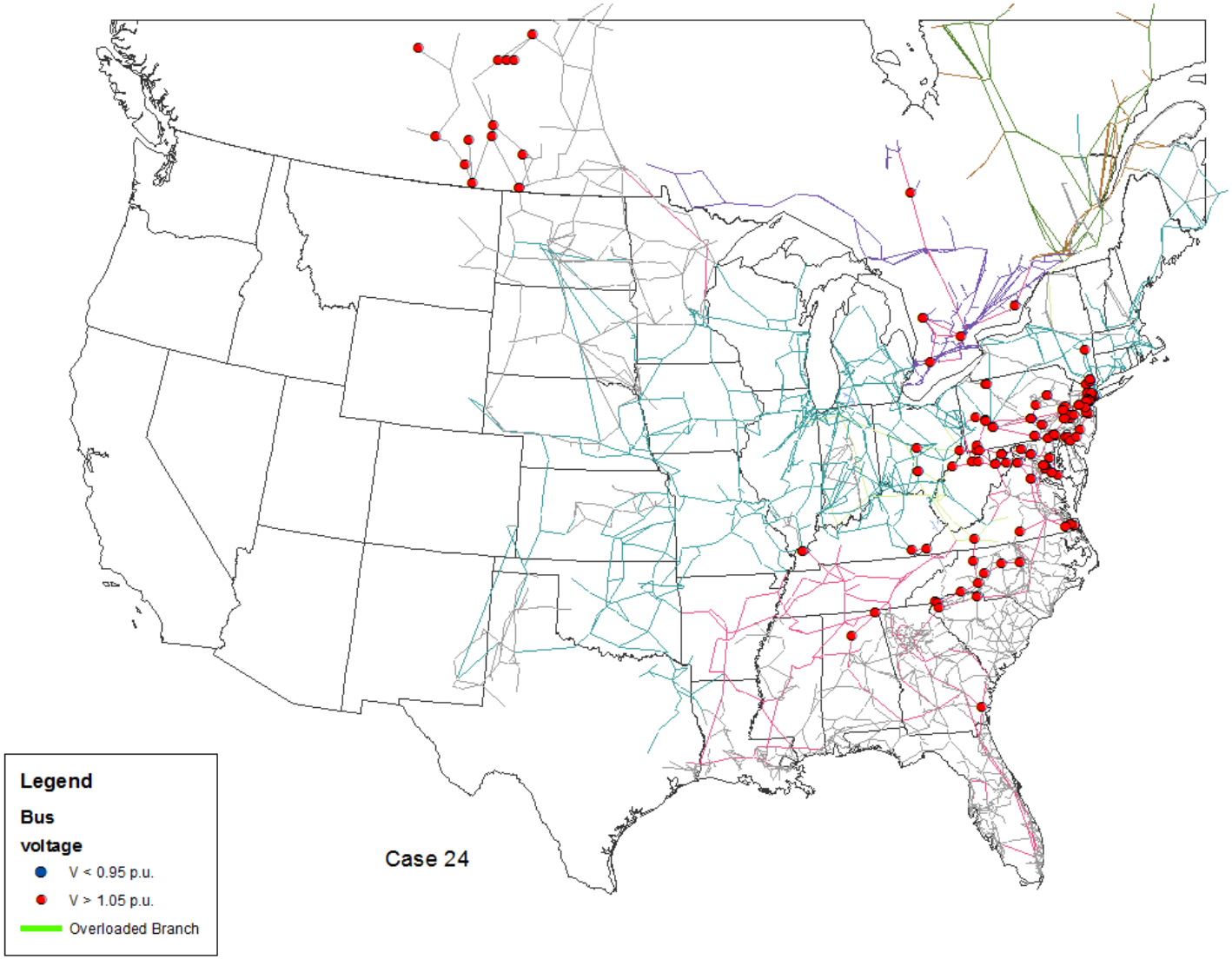


Fig. 7.11: 2013 base case with generators removed

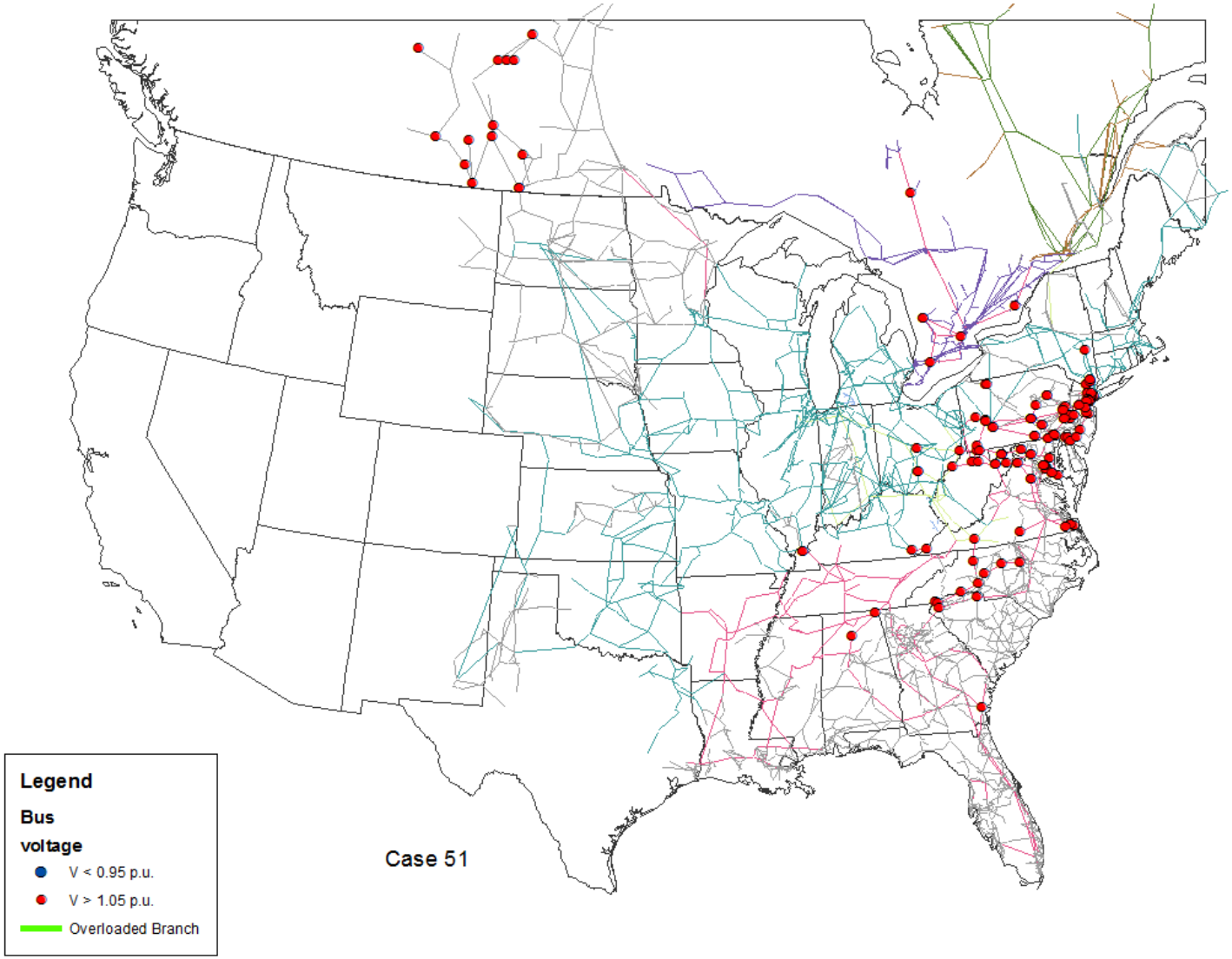


Fig. 7.12: 2013 base case with generators removed, synchronous condensers added

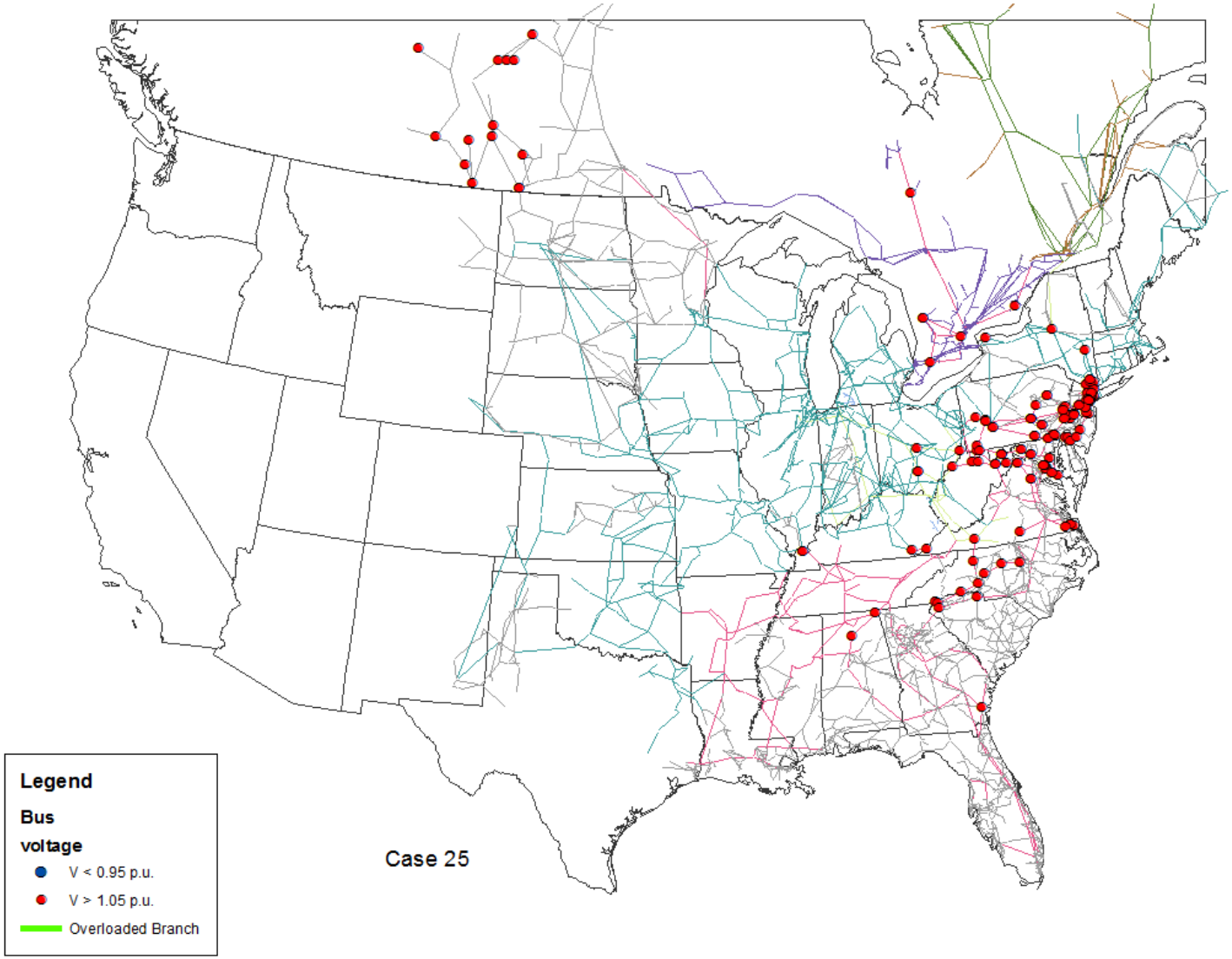


Fig. 7.13: 2013 gas base case

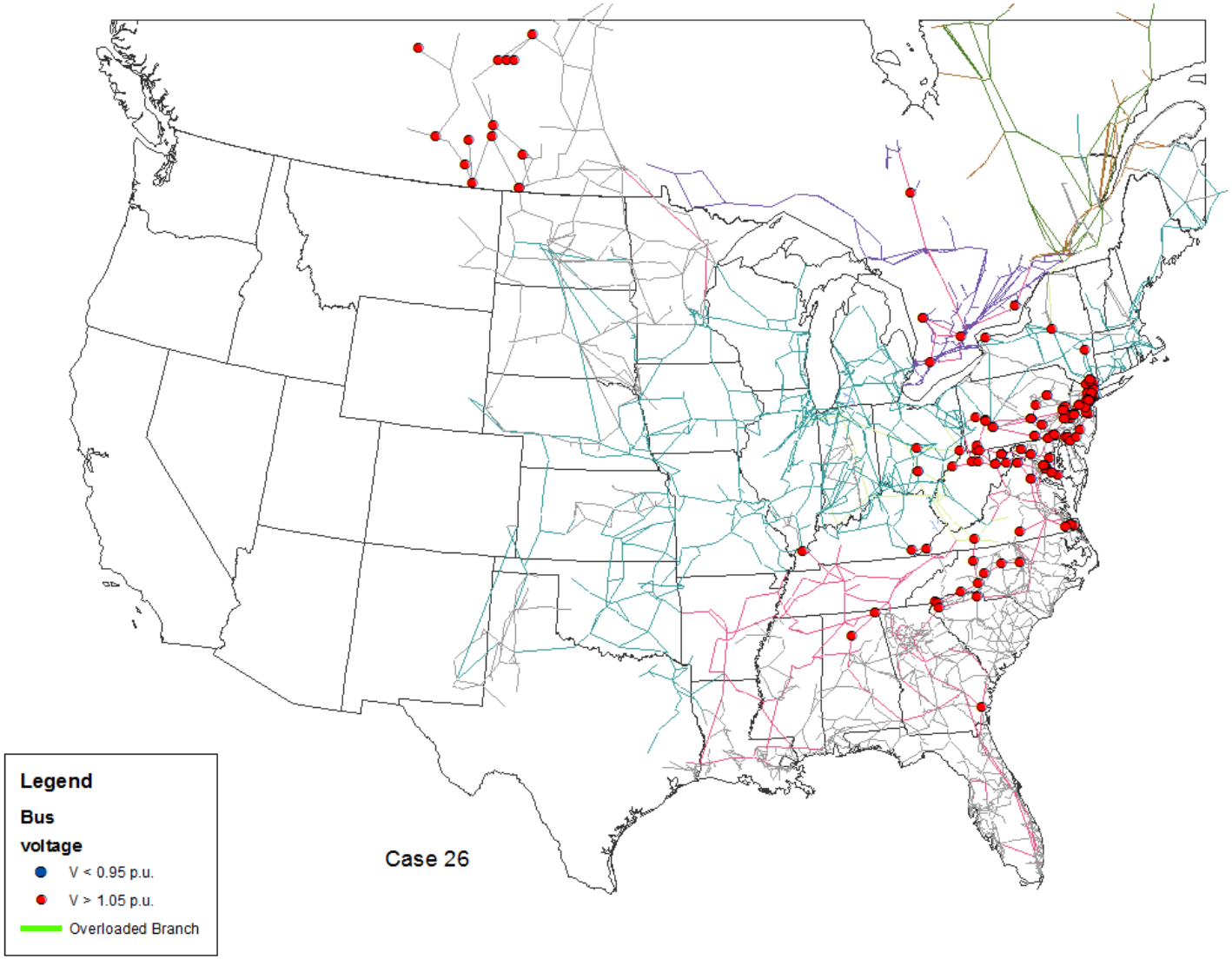


Fig. 7.14: 2013 gas base case with wind

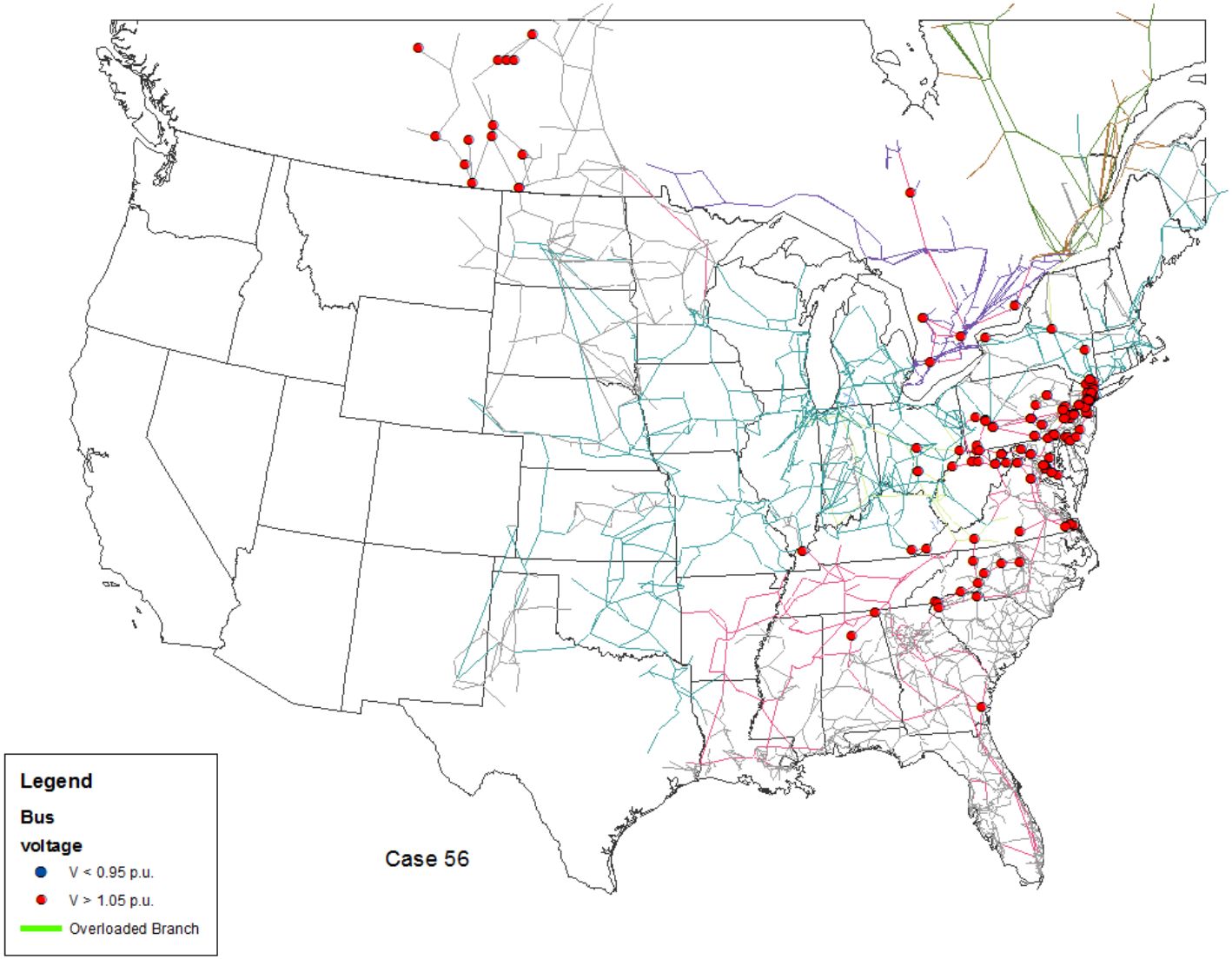


Fig. 7.15: 2013 gas base case with wind and synchronous condensers

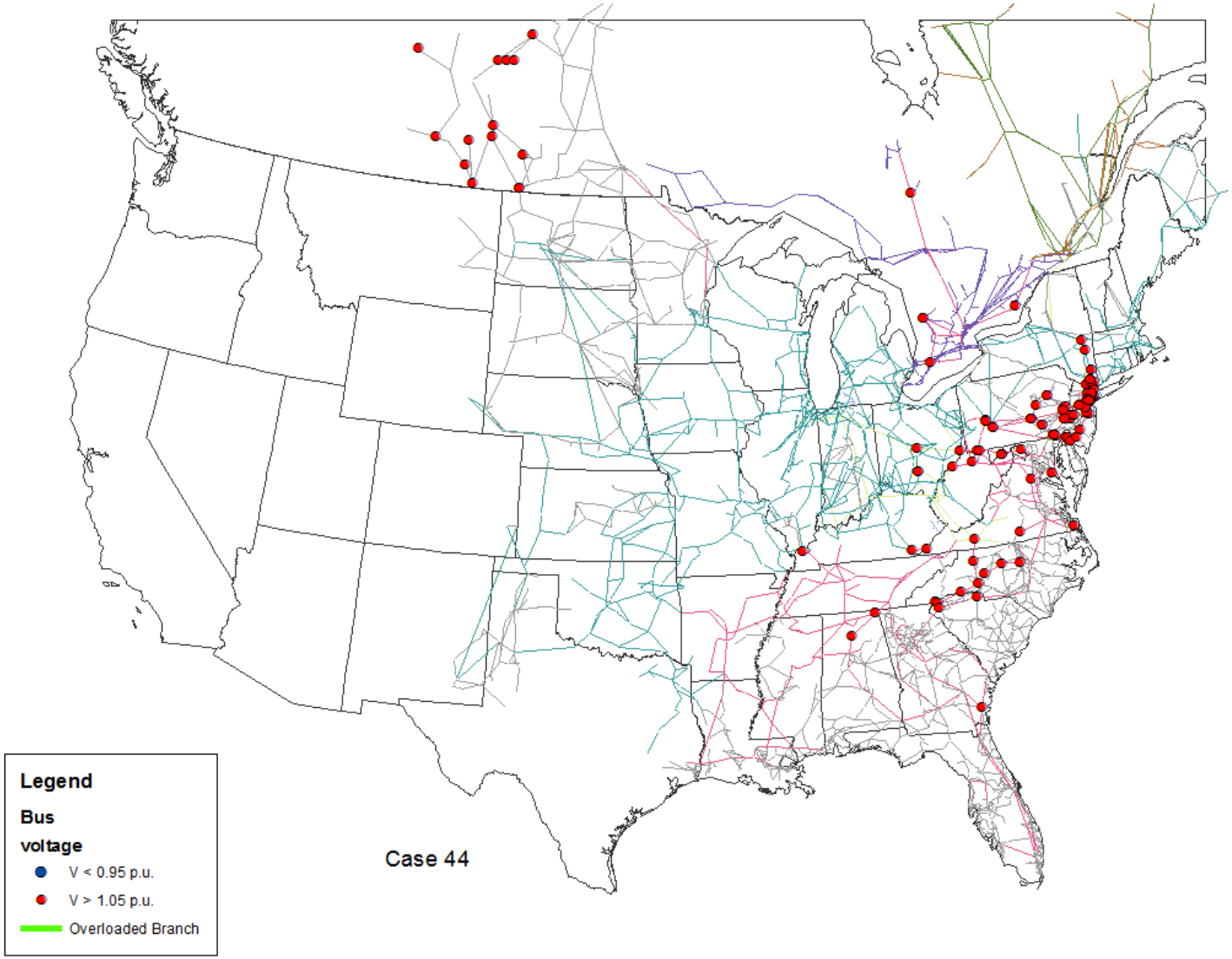


Fig. 7.16: 2013 gas and wind added, nuclear plants removed

2014

Simulation results for the 2014 base case are shown in Fig. 7.17. This case includes 1.5% load growth from the previous year. Due to convergence issues, Unit 2 of the V.C. Summers Nuclear Power Plant near Jenkinsville, South Carolina was left in the model, even though this generator is not scheduled to begin production until 2017. Sixty-three generators with a capacity of 6,539 MW are due to be deactivated in 2014; the results of the corresponding simulation are shown in Fig. 7.18. The removal of these generators resulted in 14 fewer buses being above their nominal limits. The generation scaling applied to the remaining units to make up for this loss appears to correct the low voltage condition found at two of the buses in the system, and also decreases the flow on four of the overloaded lines down to acceptable levels. Conversion of some of the deactivated units to synchronous condensers had a negligible effect on the system (Fig. 4.18).

Approximately 1,300 MW of gas-fired generating units are scheduled to be brought online in 2014. The simulation results show that this creates 23 more overvoltage buses (Fig. 7.20) than the scenario where the MATS/CSAPR-affected generators have been removed. The addition of 300 MW of wind power (Fig. 7.21) to the gas base case results in 11 fewer overvoltages buses. Synchronous condenser conversion of deactivated generators causes two additional buses to exceed their nominal voltage compared to the gas/wind case (Fig. 7.22). For the final scenario involving the removal of all nuclear generators, 28 out of the 92 reactors were successfully disabled before the simulation failed to converge. This resulted in several transmission lines in northeastern Maryland being overloaded, as shown in Fig. 7.23.

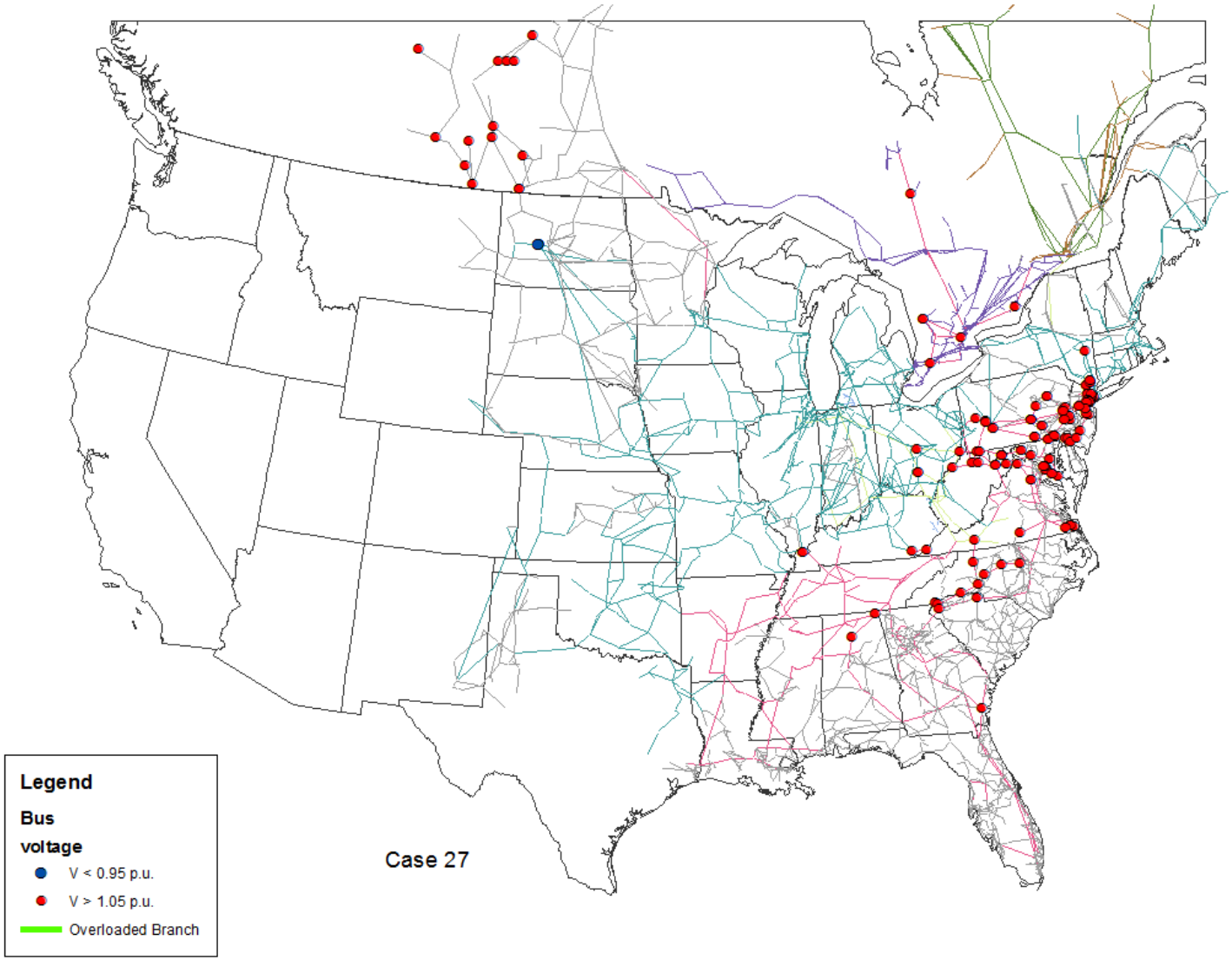


Fig. 7.17: 2014 base case

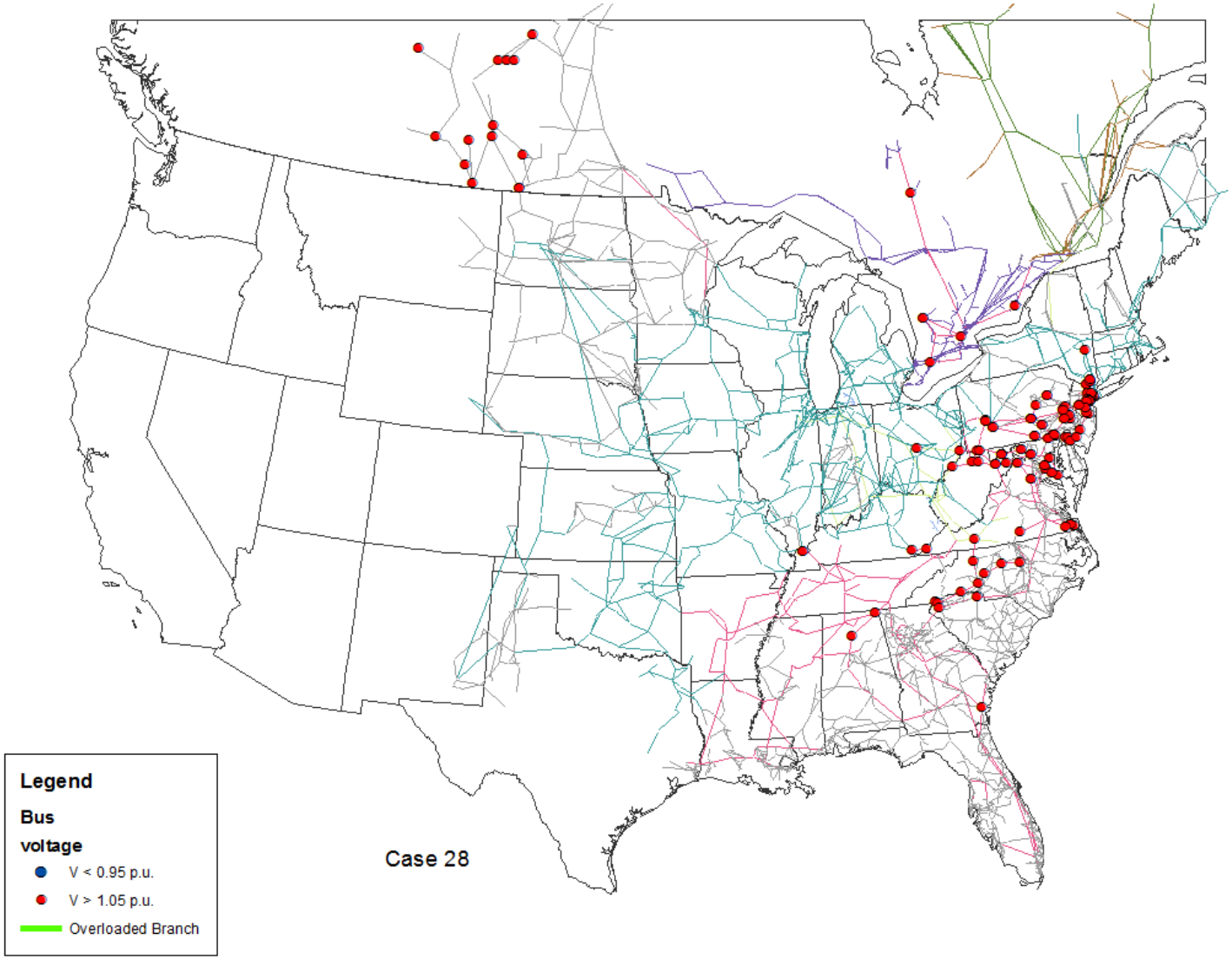


Fig. 7.18: 2014 base case with generators removed

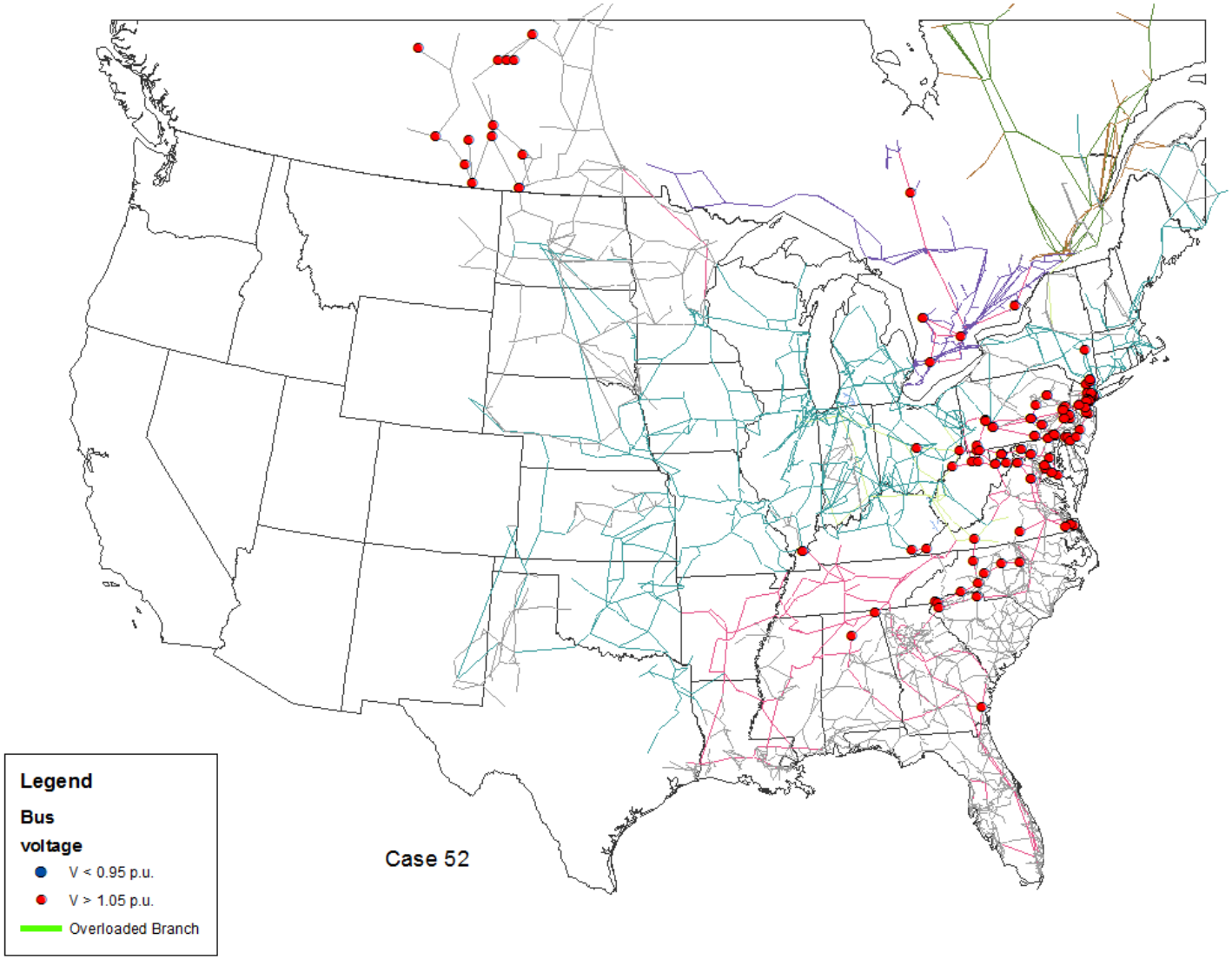


Fig. 7.19: 2014 base case with generators removed, synchronous condensers added

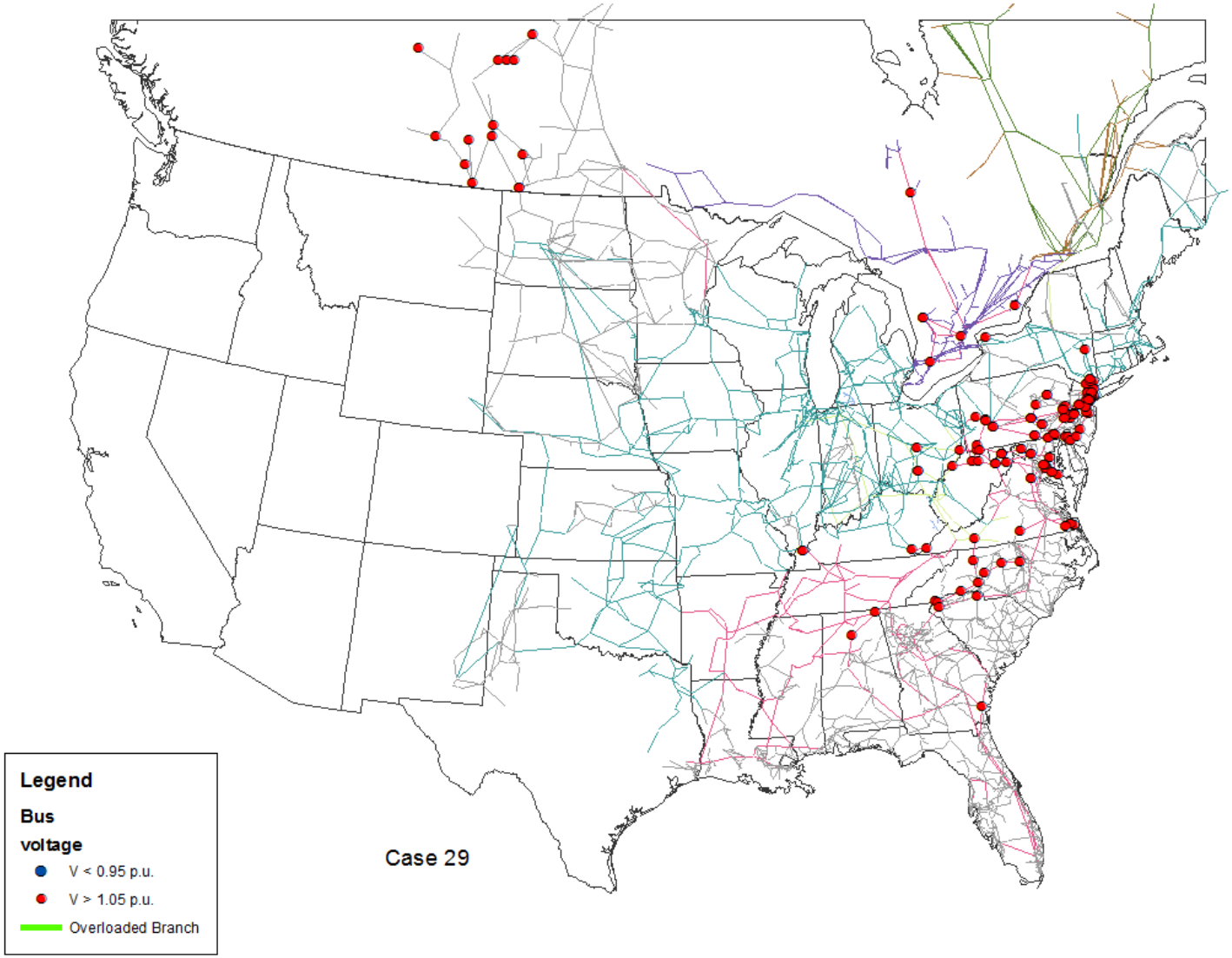


Fig. 7.20: 2014 gas base case

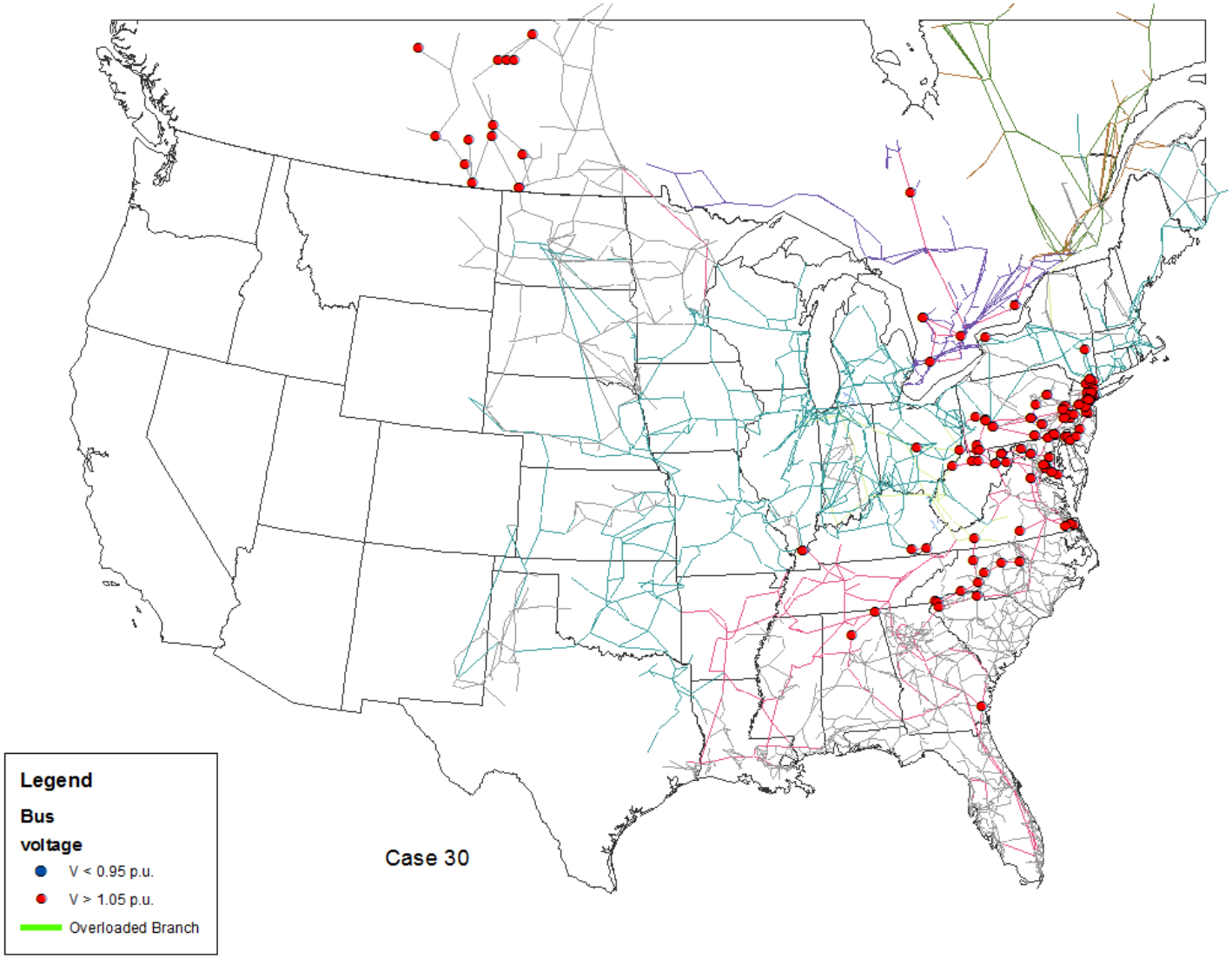


Fig. 7.21: 2014 gas base case with wind generation

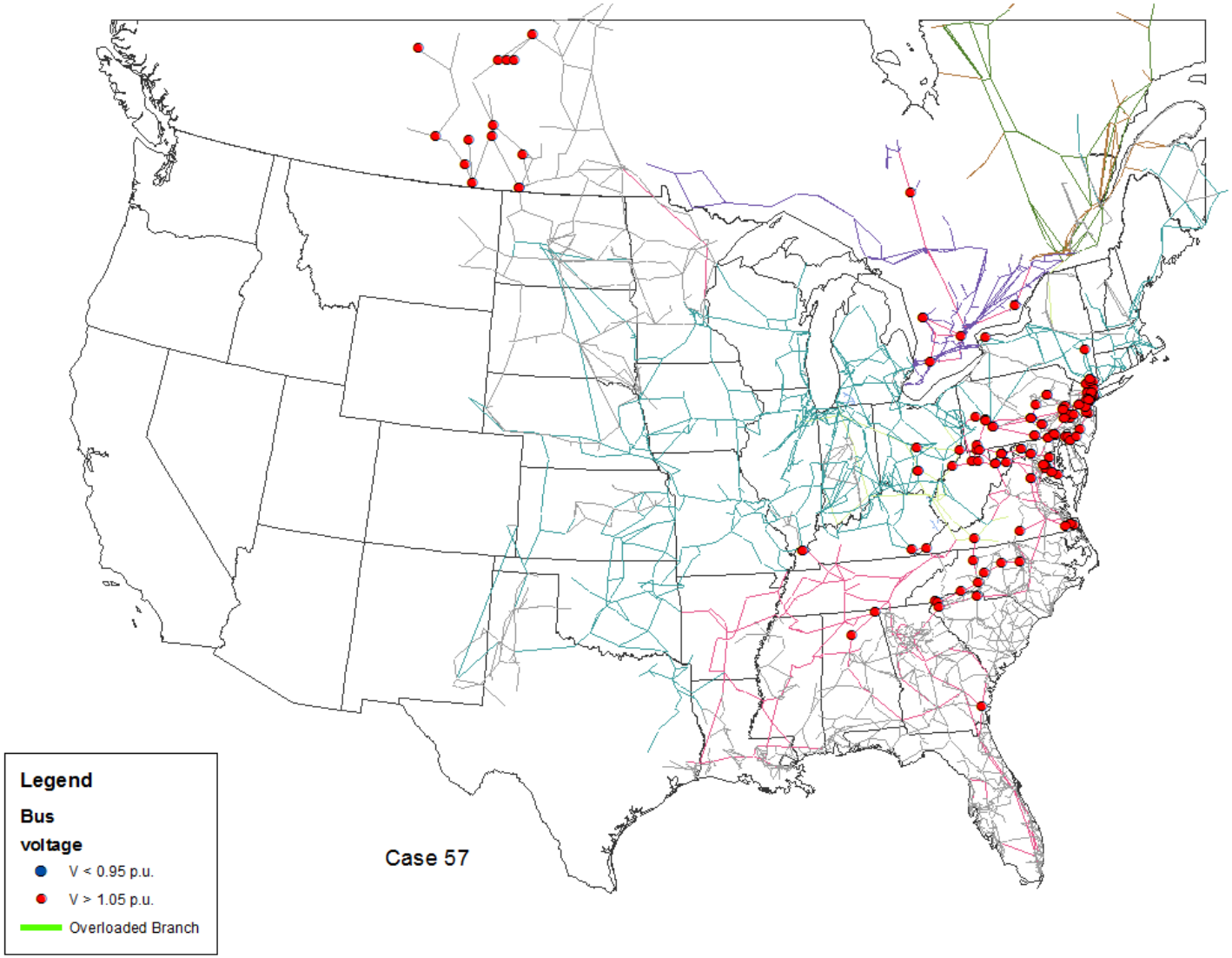


Fig. 7.22: 2014 gas base case with wind generation and synchronous condensers

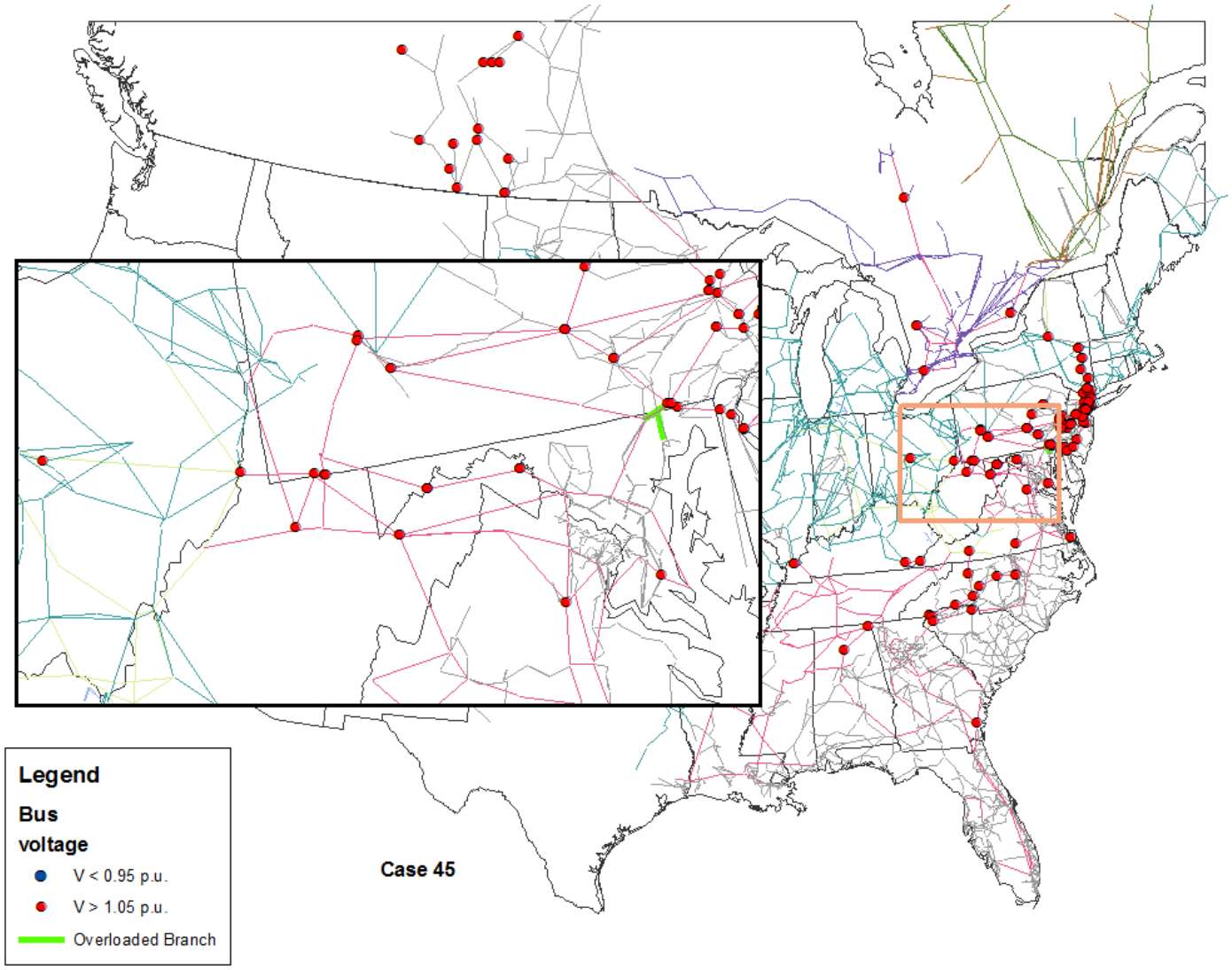


Fig. 7.23: 2014 gas and wind added, nuclear plants removed

2015

The 2015 scenarios are perhaps the most important ones in this study, since more generators (117) are to be taken offline that year than in the previous three years combined. The 2015 base case (Fig. 7.24) is fairly unremarkable in terms of the number of overvoltage buses (156) and overloaded lines (3). Once the nearly 16,000 MW of MATS/CSAPR-affected generators are removed (Fig. 7.25), two buses in northeastern Ohio drop below allowable voltage levels, and three additional lines become overloaded. It should be noted here that two yet-to-be-built generators (V.C. Summers, Unit 2, and Vogtle, Unit 3) had to be left in the model in order for these cases to converge. By converting several of the generators to synchronous condensers, the bus voltage issues were eliminated (Fig. 7.26).

The addition of 3,859 MW of natural gas-fired generation appears to create some undervoltage issues around the southeastern shore of Lake Erie, and a few overloaded lines in northeastern Maryland (Fig. 7.27), though several of the undervoltage buses are eliminated with the addition of planned wind generation (Fig. 7.28). For these two cases, the newly added generators allowed the previously mentioned Summers and Vogtle units to be successfully removed from the model. Conversion of some of the deactivated generators to synchronous condensers corrected the remaining voltage issues, but did not reduce the number of overloaded lines (Fig. 7.29).

In 2015, the Watts Bar Unit 2 reactor is scheduled to begin operation, the first nuclear reactor to do so in nearly twenty years. Because of its location, the plant does not appear to have much effect on the undervoltage buses and overloaded lines previously noted in the gas and wind cases for this year (Fig. 7.30 and Fig. 7.31). As before, synchronous condensers were helpful in alleviating some of these problems, as shown in Fig. 7.32. Only three nuclear power plants could be removed before the system failed to converge (Fig. 7.33).

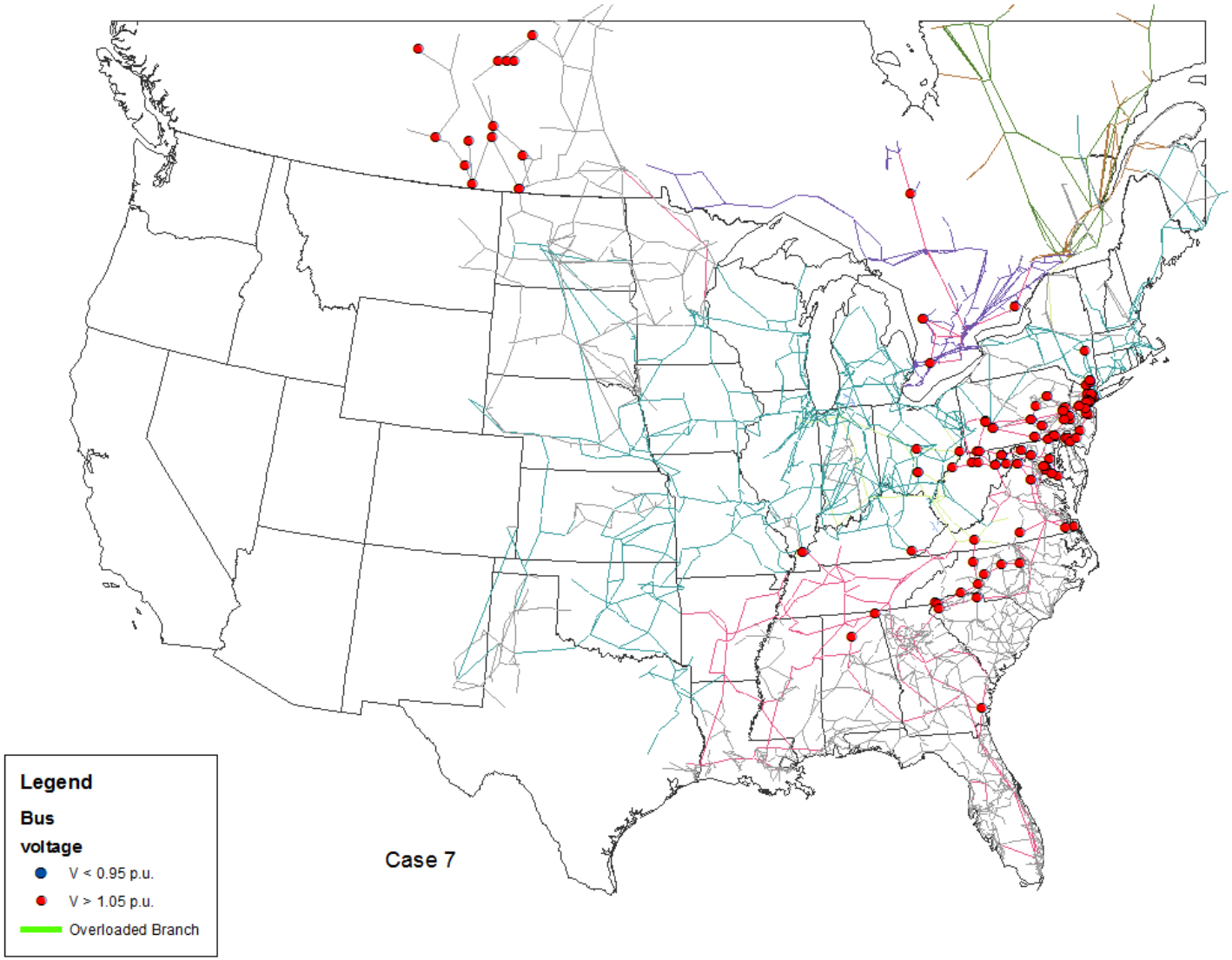


Fig. 7.24: 2015 base case

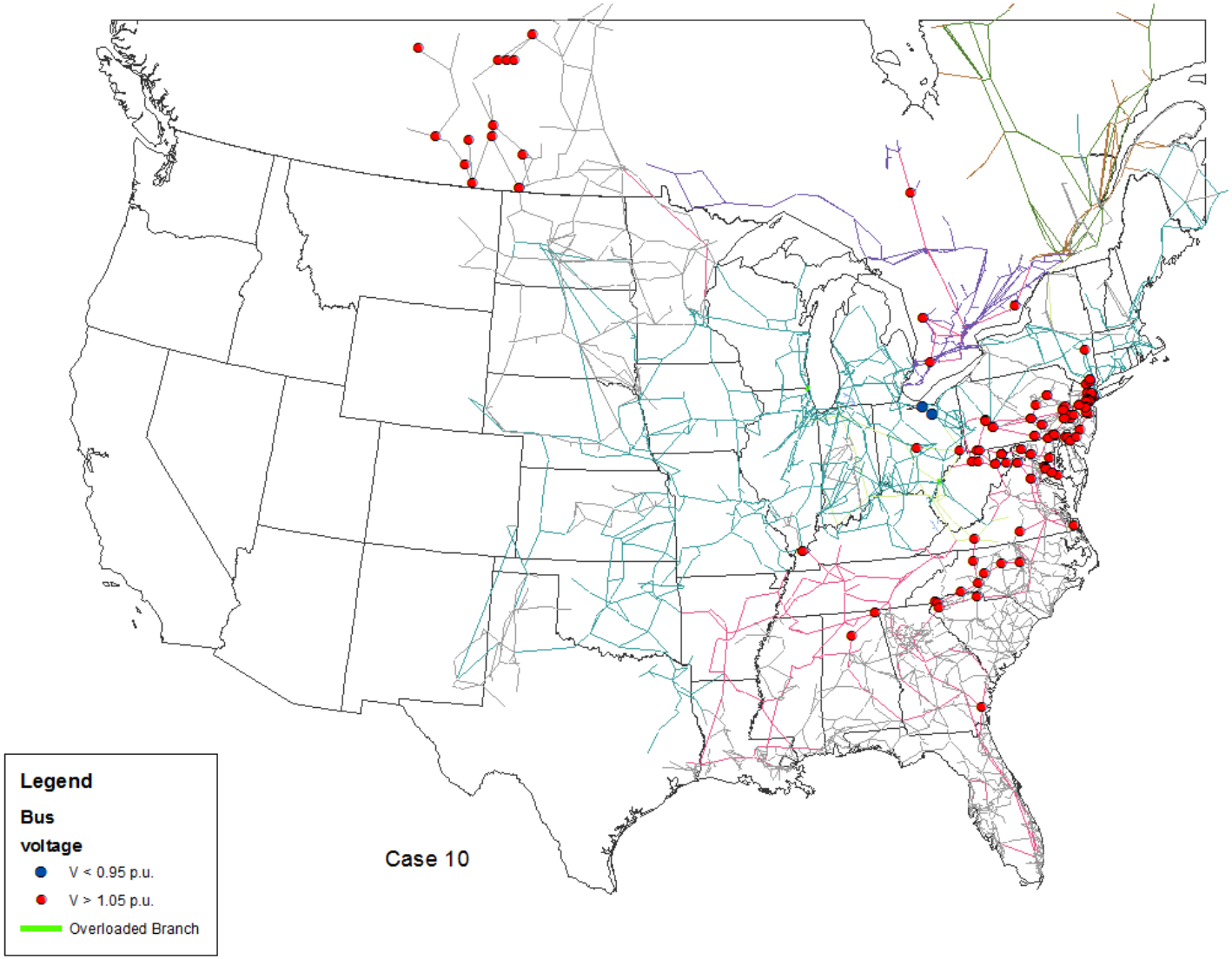


Fig. 7.25: 2015 base case with generators removed

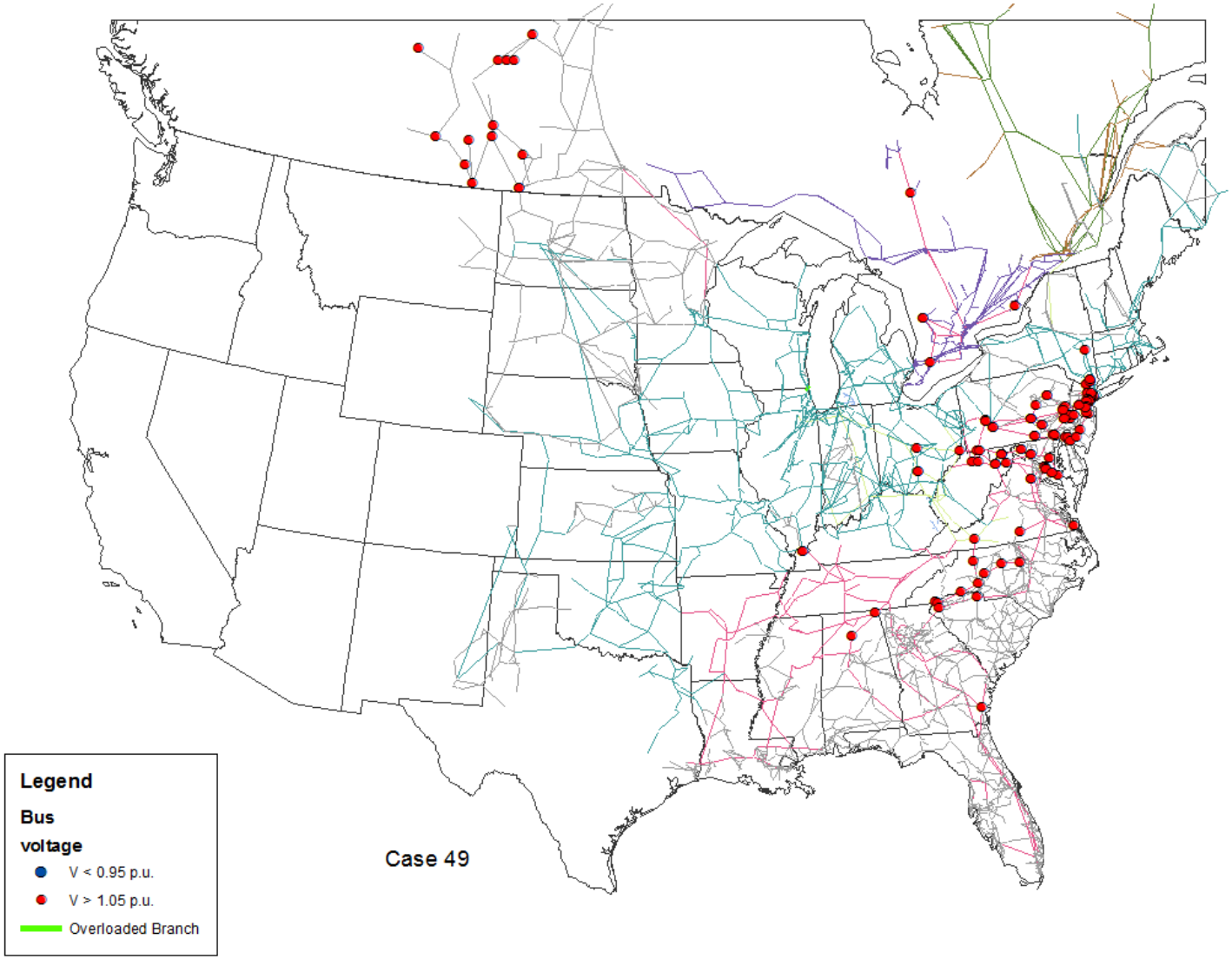


Fig. 7.26: 2015 base case with generators removed, synchronous condensers added

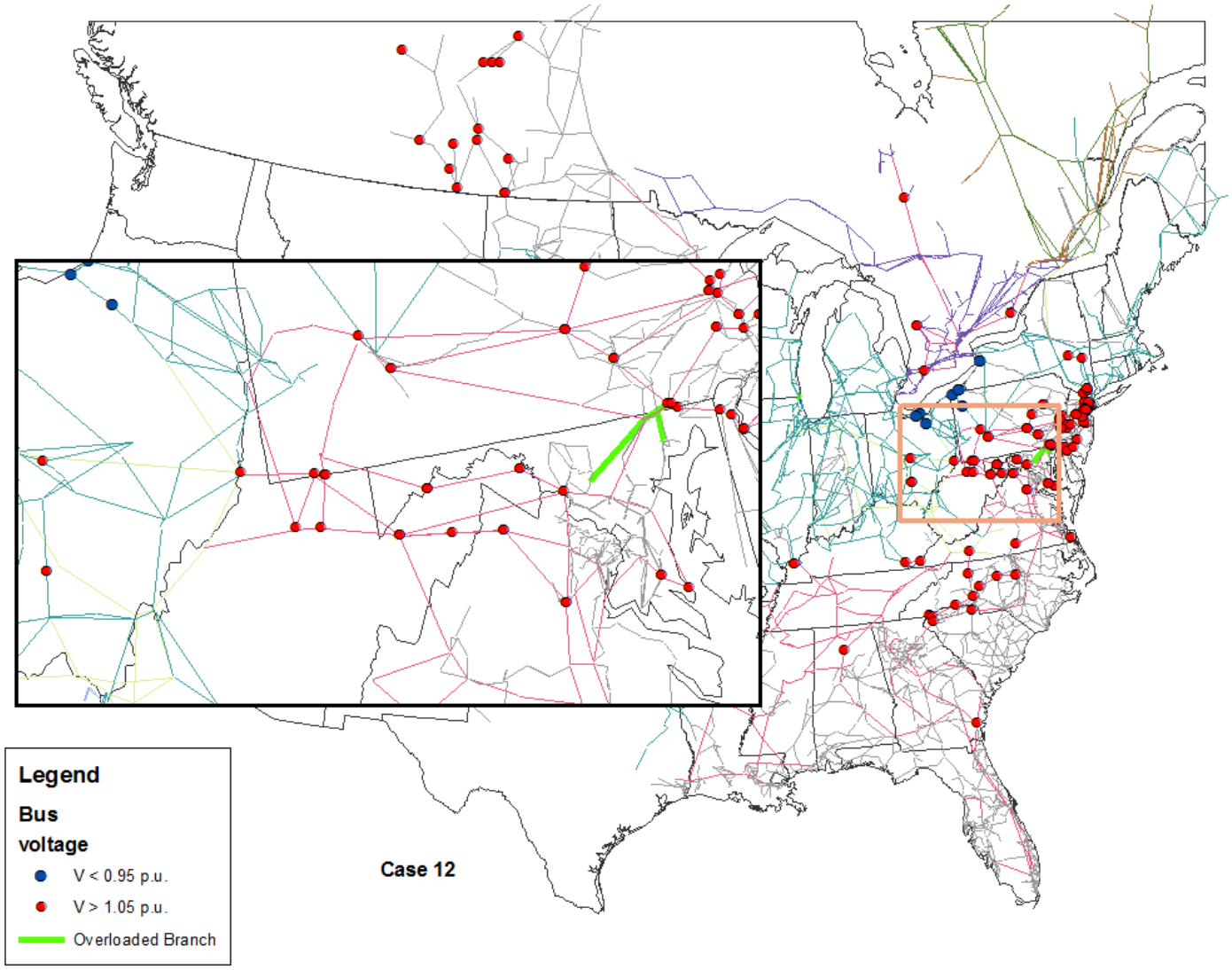


Fig. 7.27: 2015 gas base case

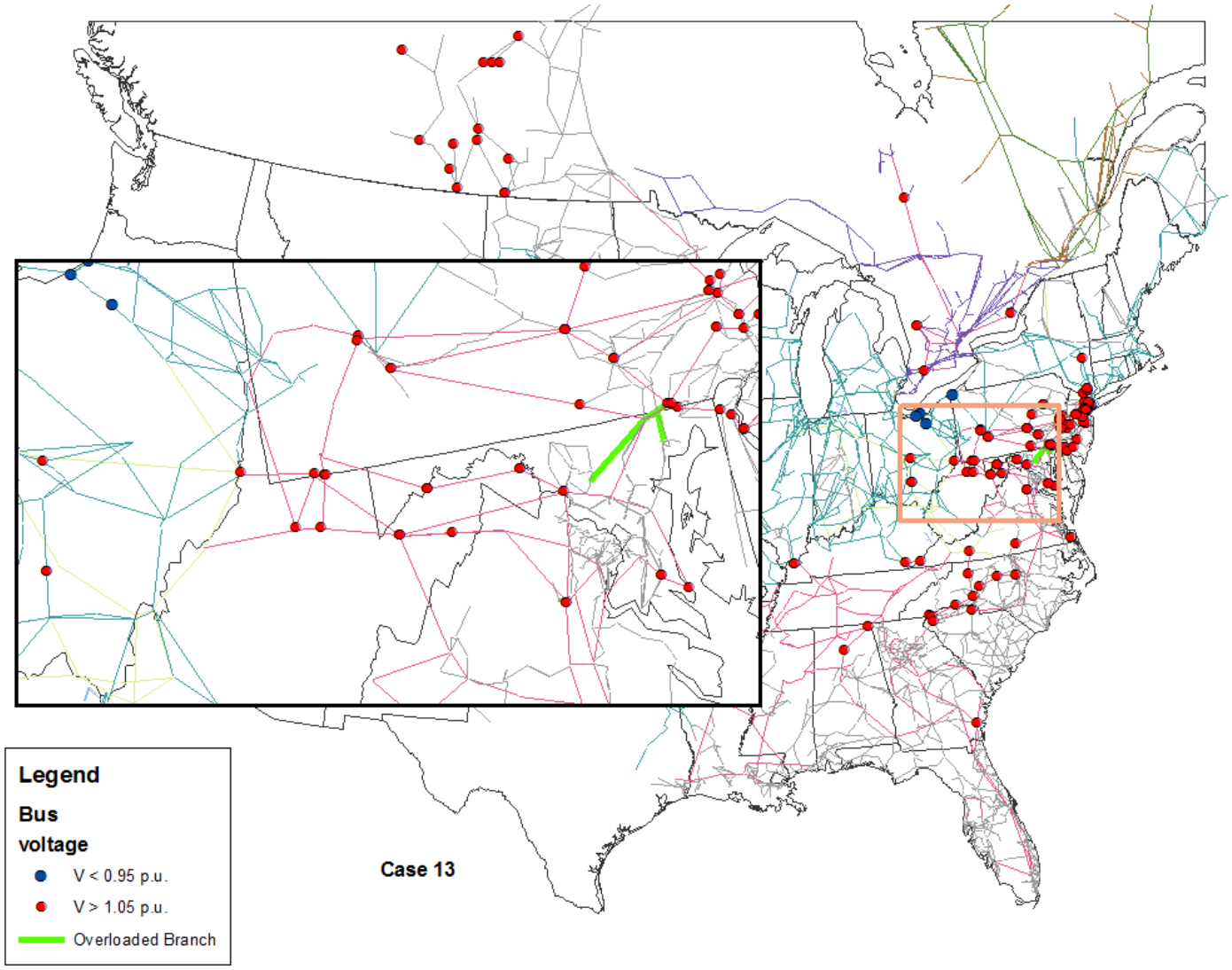


Fig. 7.28: 2015 gas base case with wind

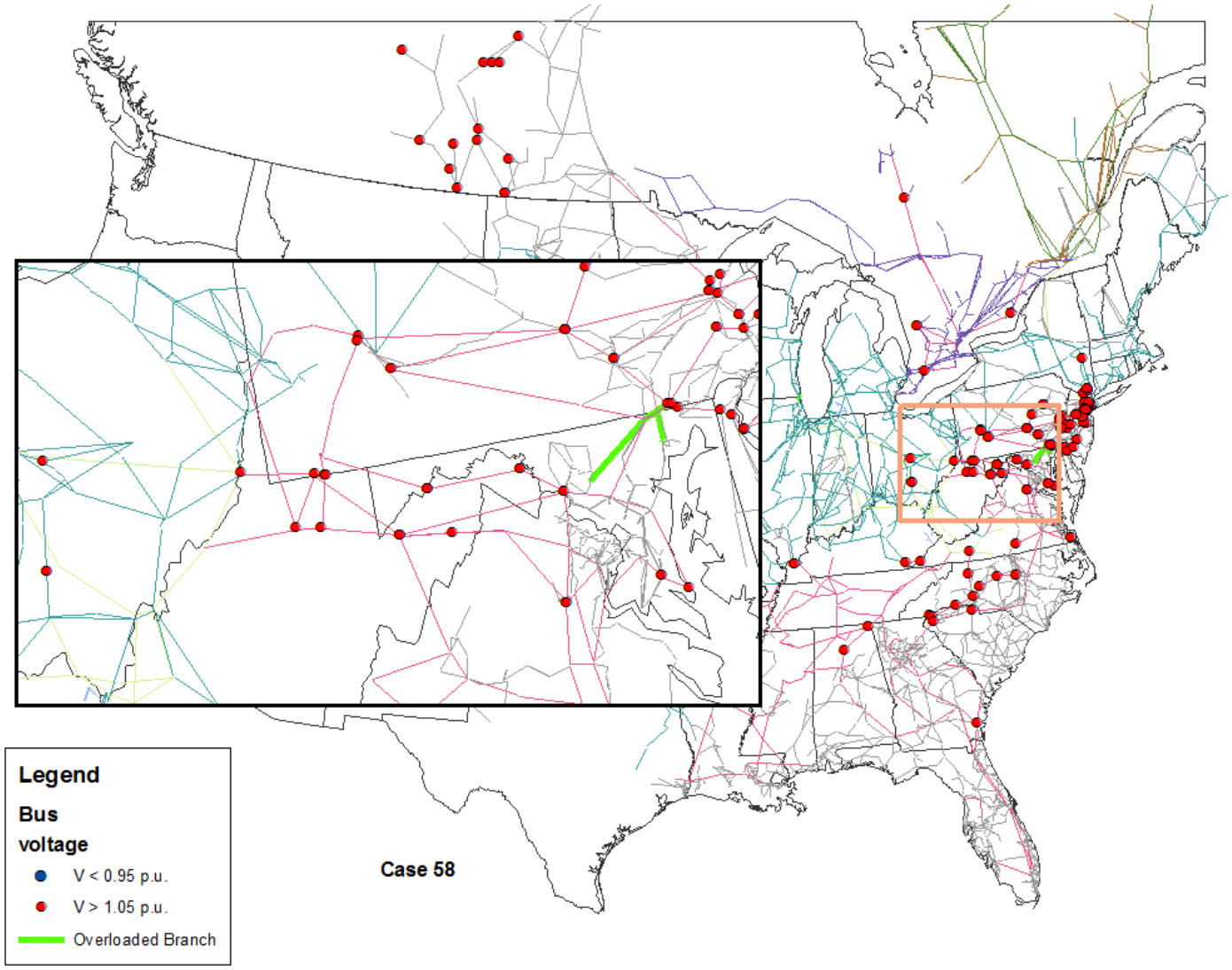


Fig. 7.29: 2015 gas base case with wind generation and synchronous condensers

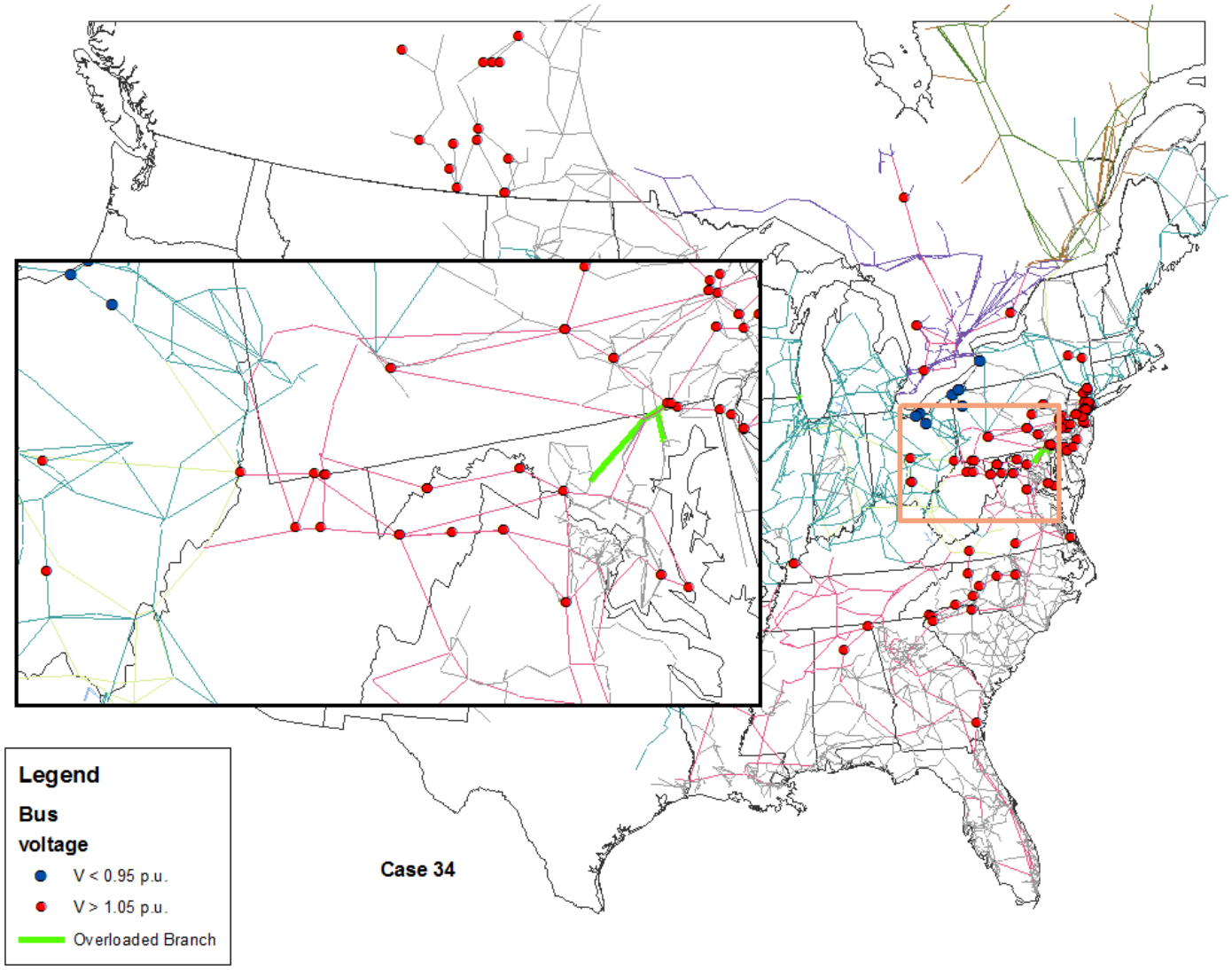


Fig. 7.30: 2015 gas base case with new nuclear units

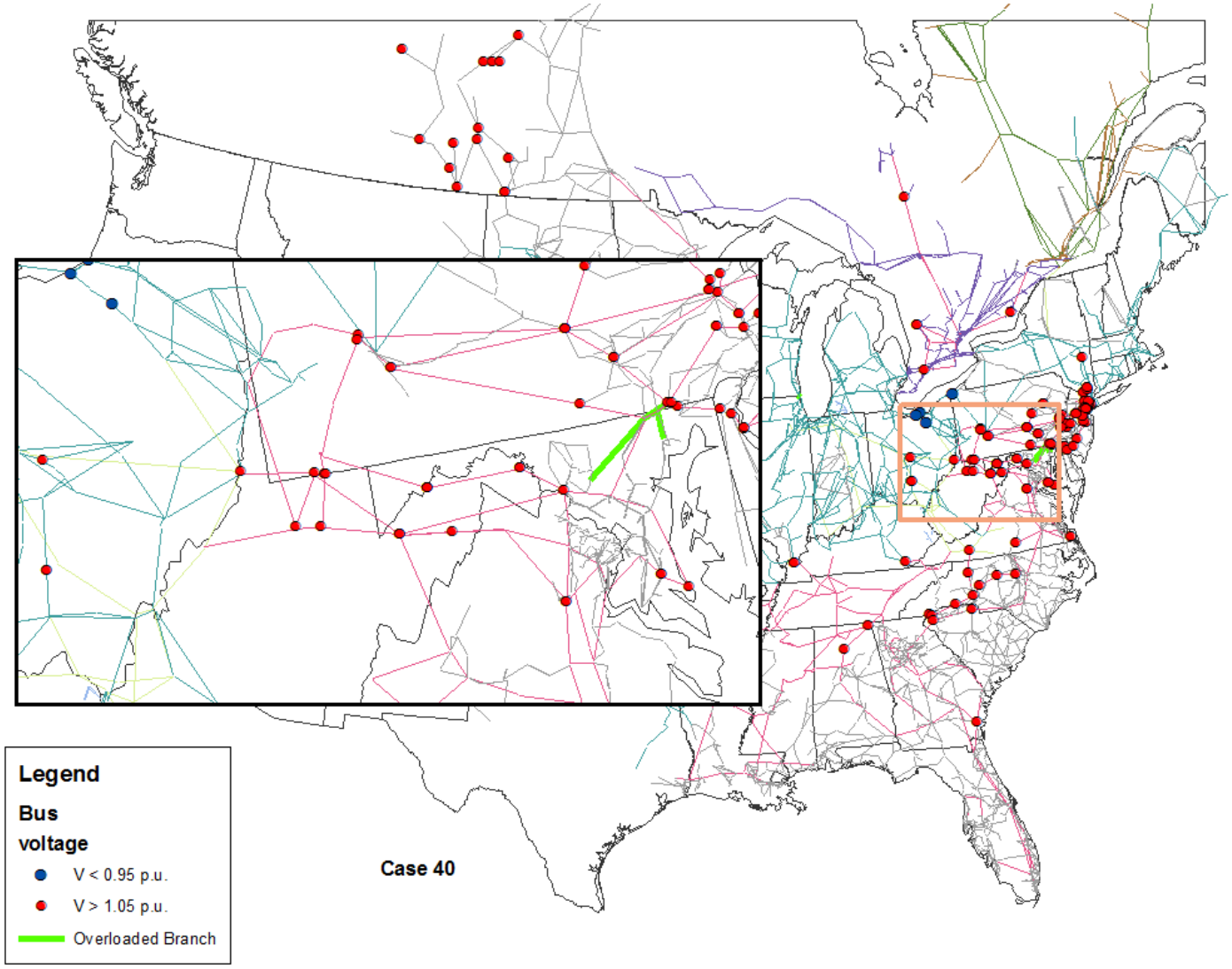


Fig. 7.31: 2015 gas, wind base case with new nuclear units

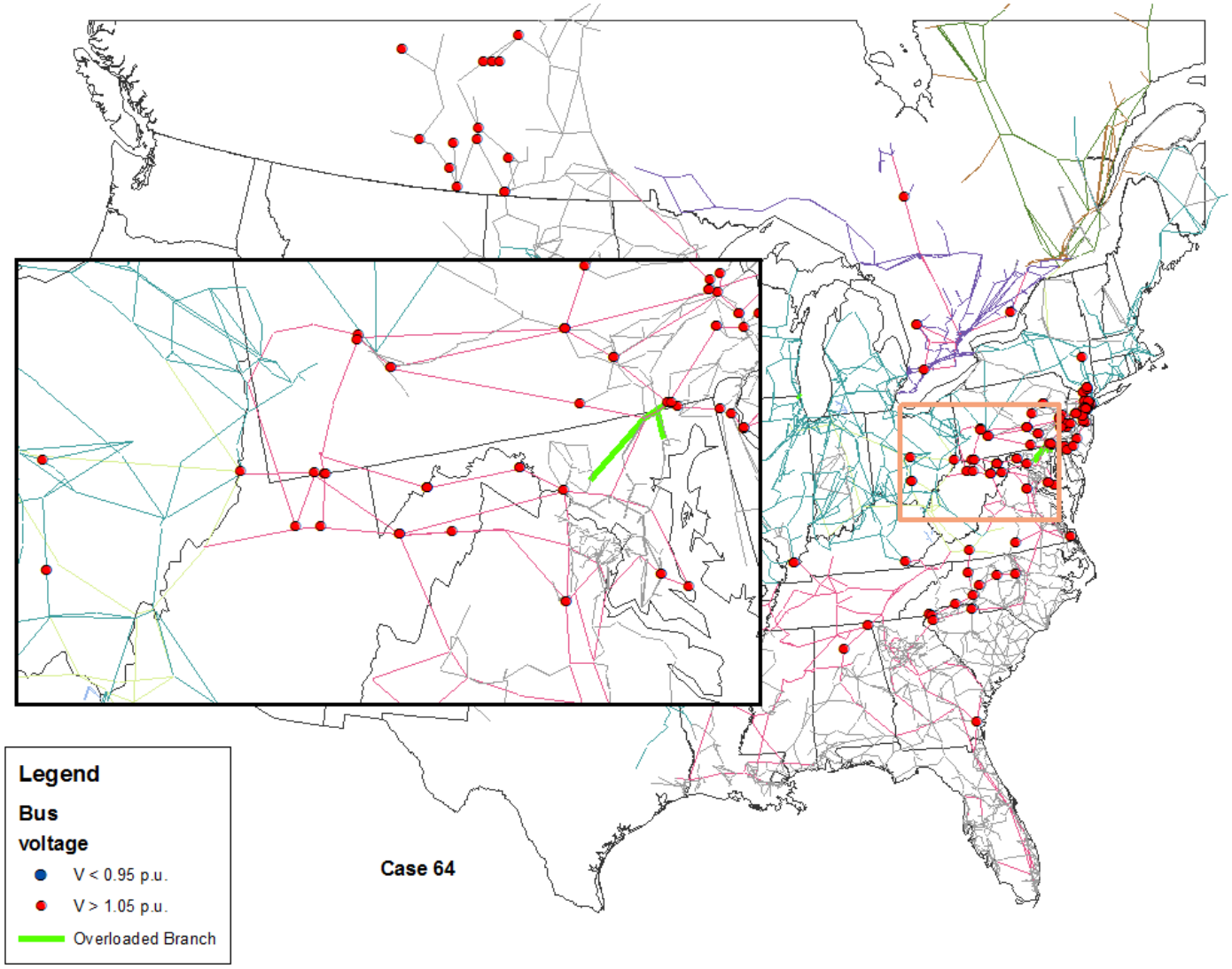


Fig. 7.32: 2015 gas, wind, new nuclear units, and synchronous condensers

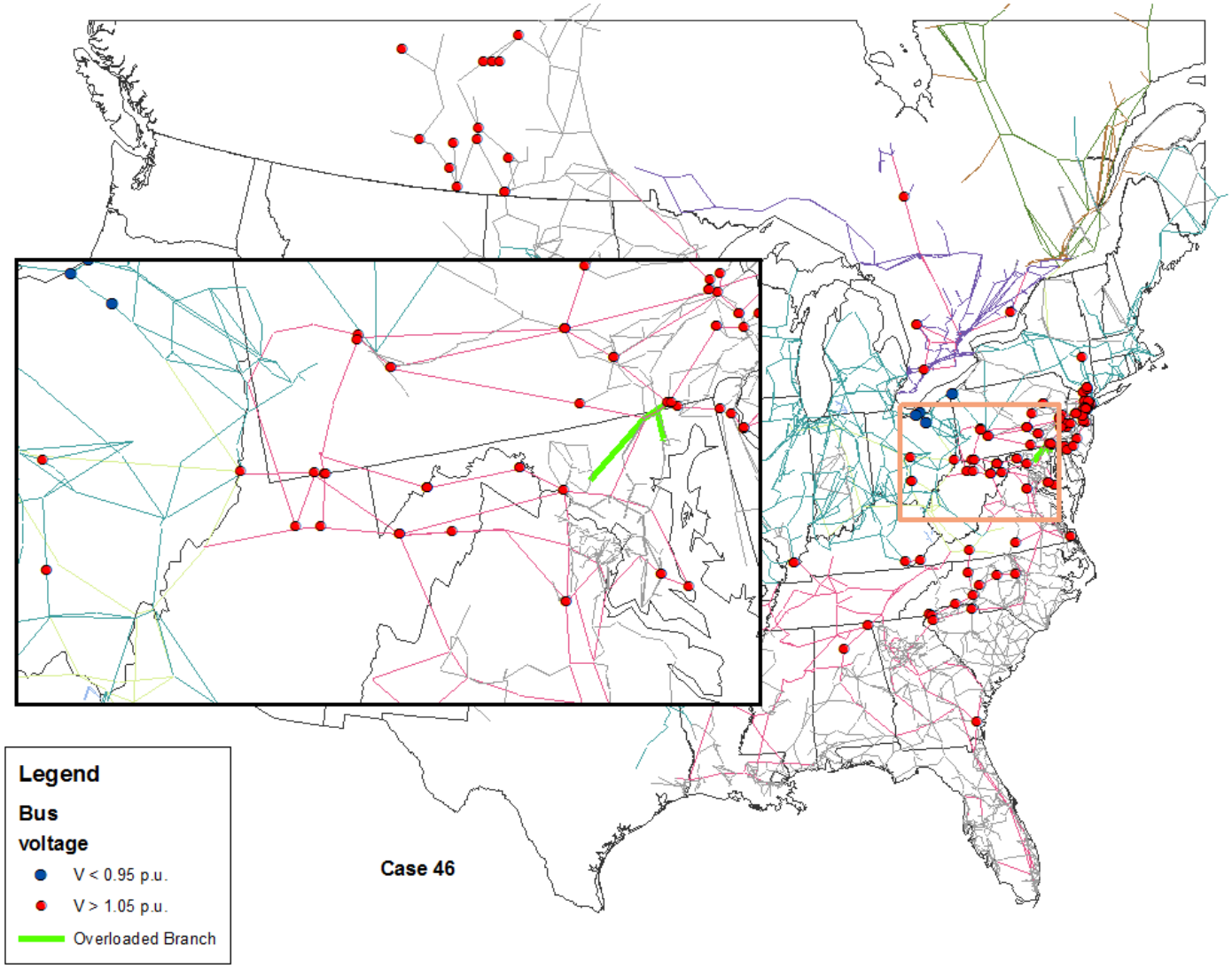


Fig. 7.33: 2015 gas and wind base case, nuclear plants removed

2016

Simulation results from the 2016 base case are shown in Fig. 7.34. This case contains 146 overvoltage buses and three overloaded lines. Removal of the 18 MATS/CSPAR-affected generators lowers the bus voltages such that only 126 buses exceed their voltage limits (Fig. 7.35). However, six additional lines become overloaded, most notably in the vicinity of the Kyger Creek Power Plant in West Virginia. This is at least partially due to the fact that the plant serves as the swing bus for the model being used. That is, its power output is adjusted by the power flow program to make up for the overall load-generation mismatch of the system once the remaining generators have been dispatched. For this particular simulation, the swing bus output was very high (8,166 MW), well above the actual 973 MW limit. Because this large amount of power must be sent out through the surrounding transmission lines, it is not surprising that they would become overloaded. Efforts made to lower the swing bus output to a more reasonable value by increasing the remaining generators' outputs generally resulted in a non-convergent model. Many of the subsequent cases for 2016 and 2017 also exhibited this phenomenon, and their results should be interpreted with caution. Application of synchronous condensers to the system (Fig. 7.36) made virtually no difference in the number of out-of-limit bus voltages or overloaded lines.

Undervoltage buses along the southeastern shore of Lake Erie and overloaded lines in northeastern Maryland were again noted once gas-fired generators were introduced into the model (Fig. 7.37). Given that the swing bus output for this case was within reasonable limits, this result likely provides a decent reflection of reality. However, the apparent improvement in bus voltages provided by the added wind generation (Fig. 7.38) may be illusory, since the swing bus output was very high for this simulation. This is also true for the synchronous condenser case (Fig. 7.39), which appeared to show additional improvements in the bus voltages.

Two new nuclear plants (V.C. Summers, Unit 2, and Vogtle, Unit 3) are scheduled to come online in 2016. The results of simulations reflecting these additions (Fig. 7.40, Fig. 7.41, and Fig. 7.42) essentially mirror those of the previous ones where no new nuclear reactors were added. The simulation of a total nuclear shutdown experienced significant difficulties, with only one generator being removed before the model failed to converge (Fig. 7.43).

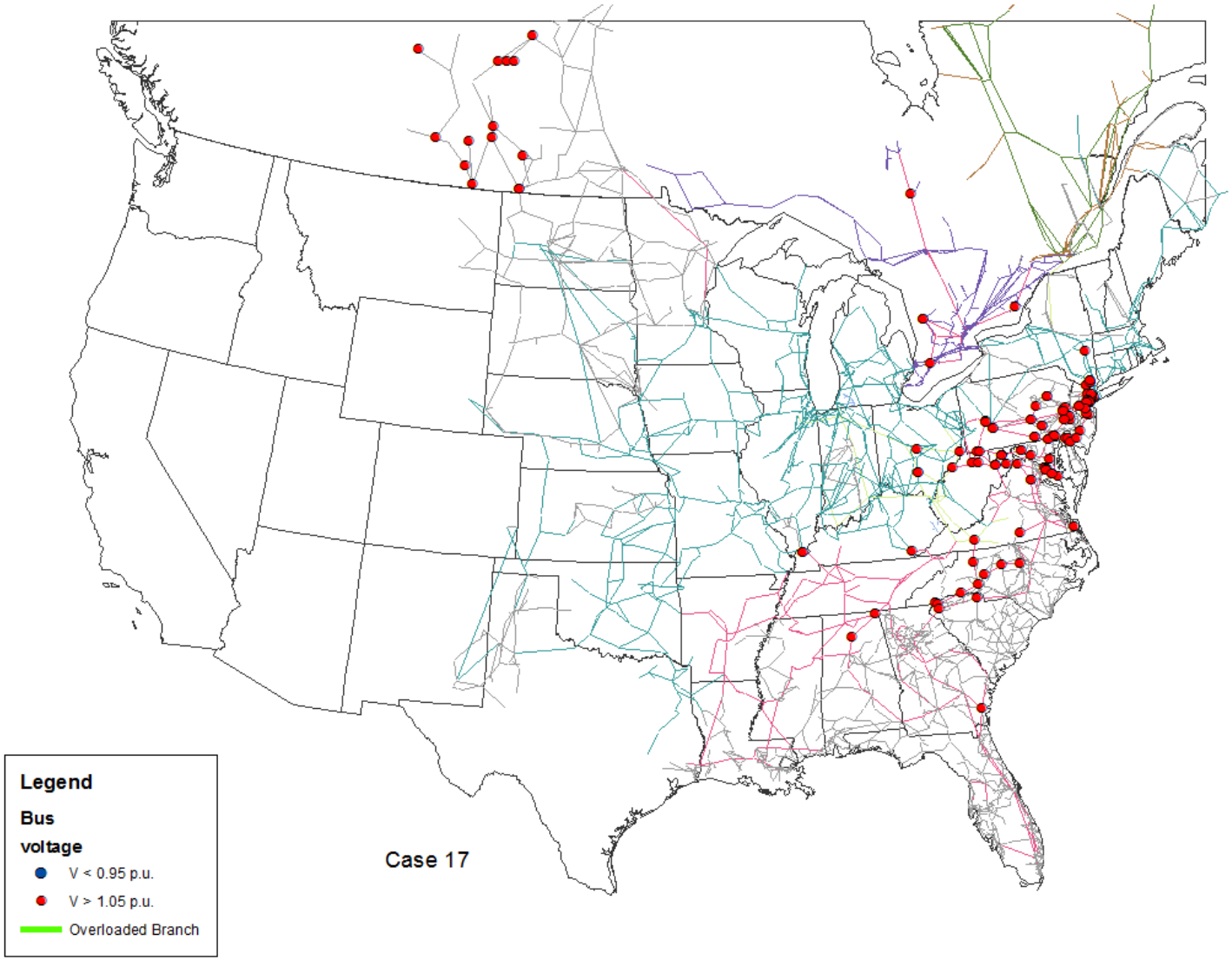


Fig. 7.34: 2016 base case

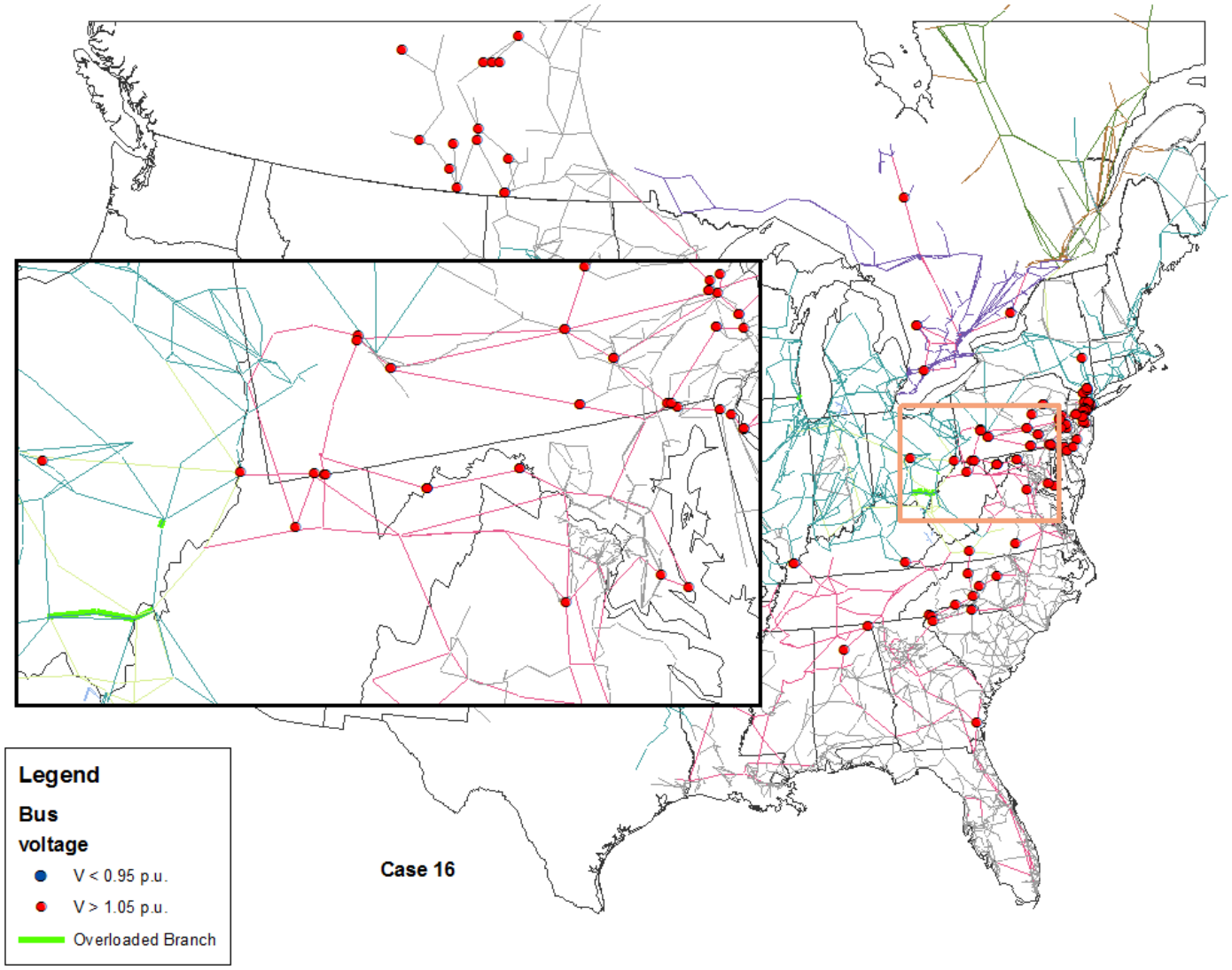


Fig. 7.35: 2016 base case with generators removed

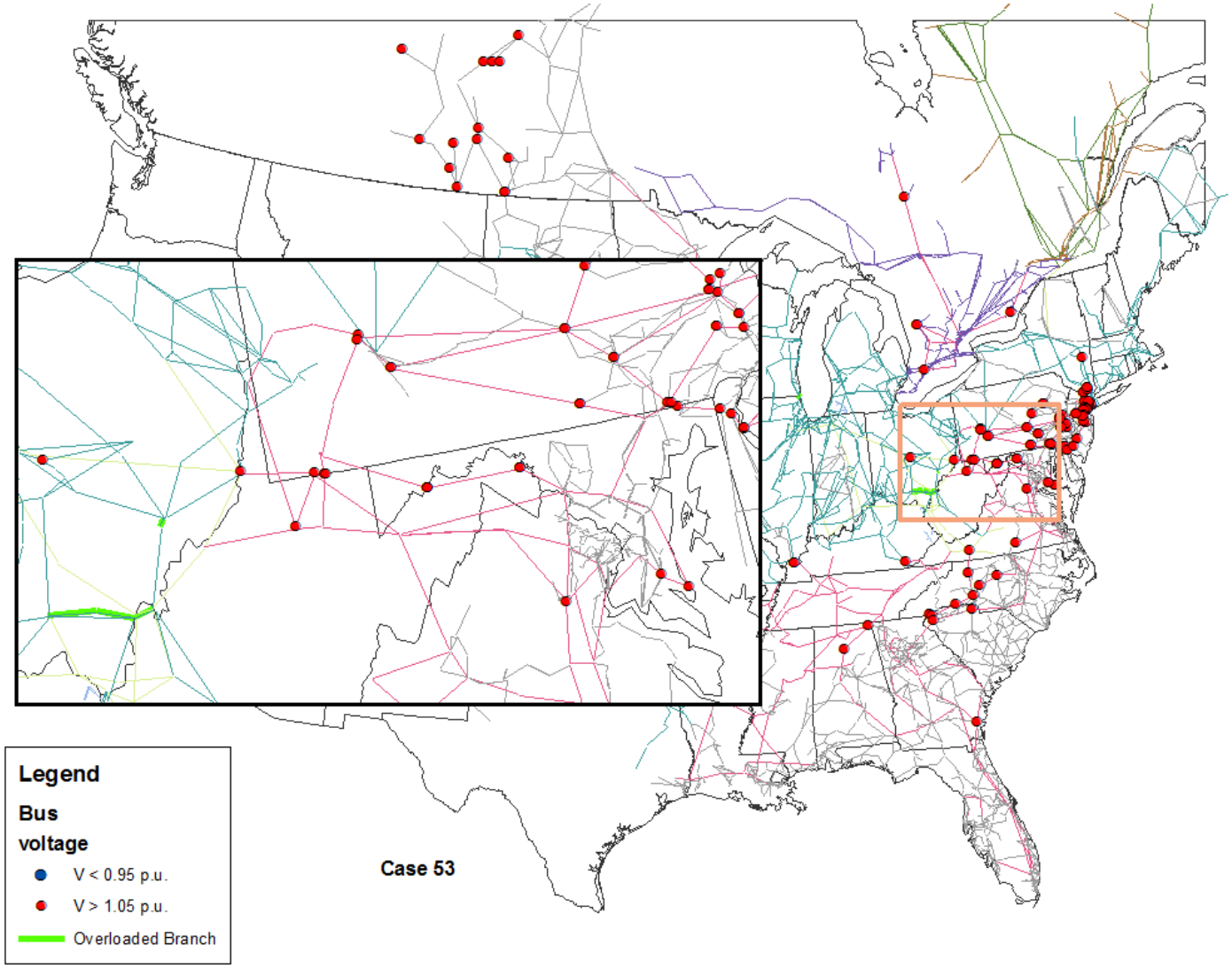


Fig. 7.36: 2016 base case with generators removed, synchronous condensers added

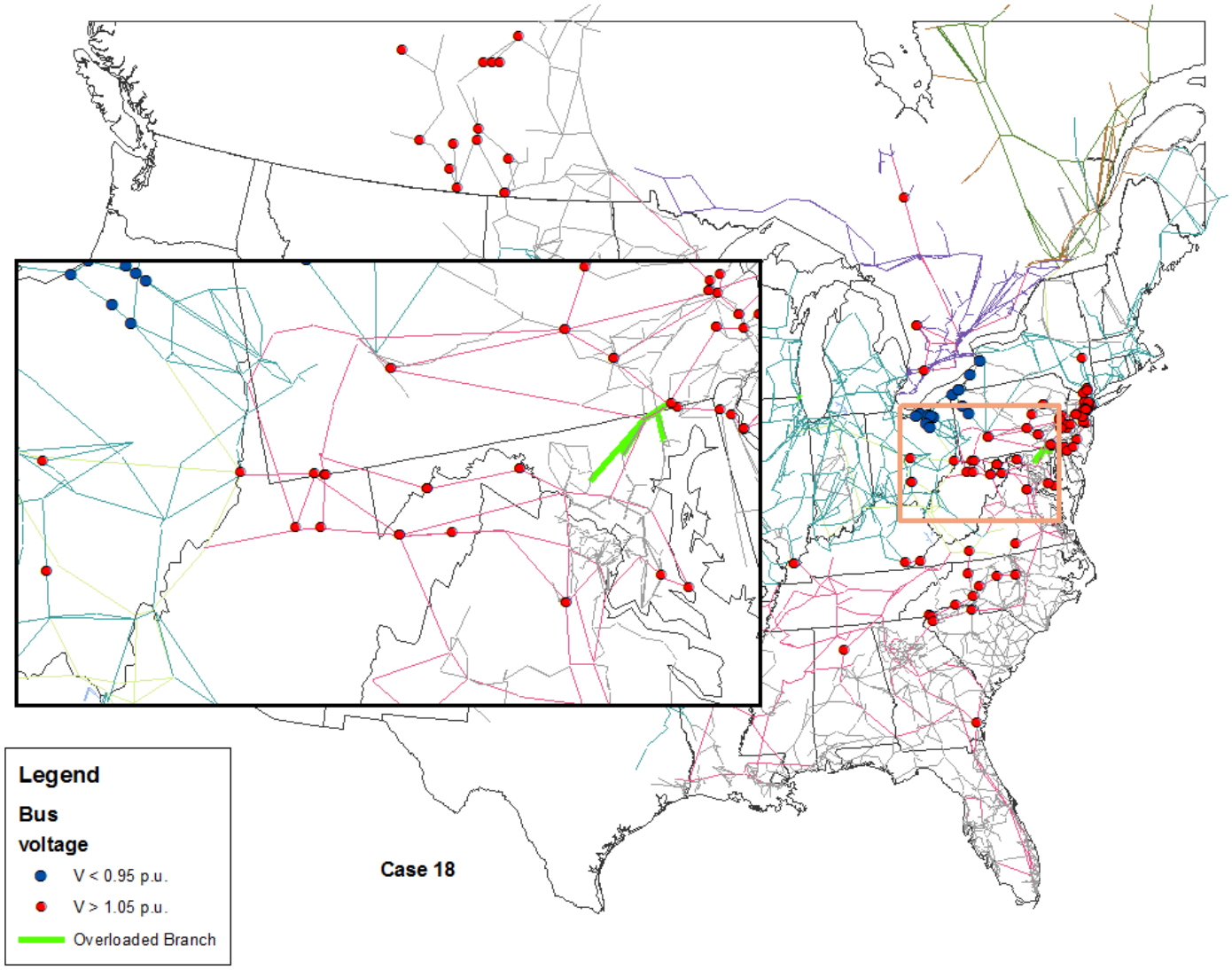


Fig. 7.37: 2016 gas base case

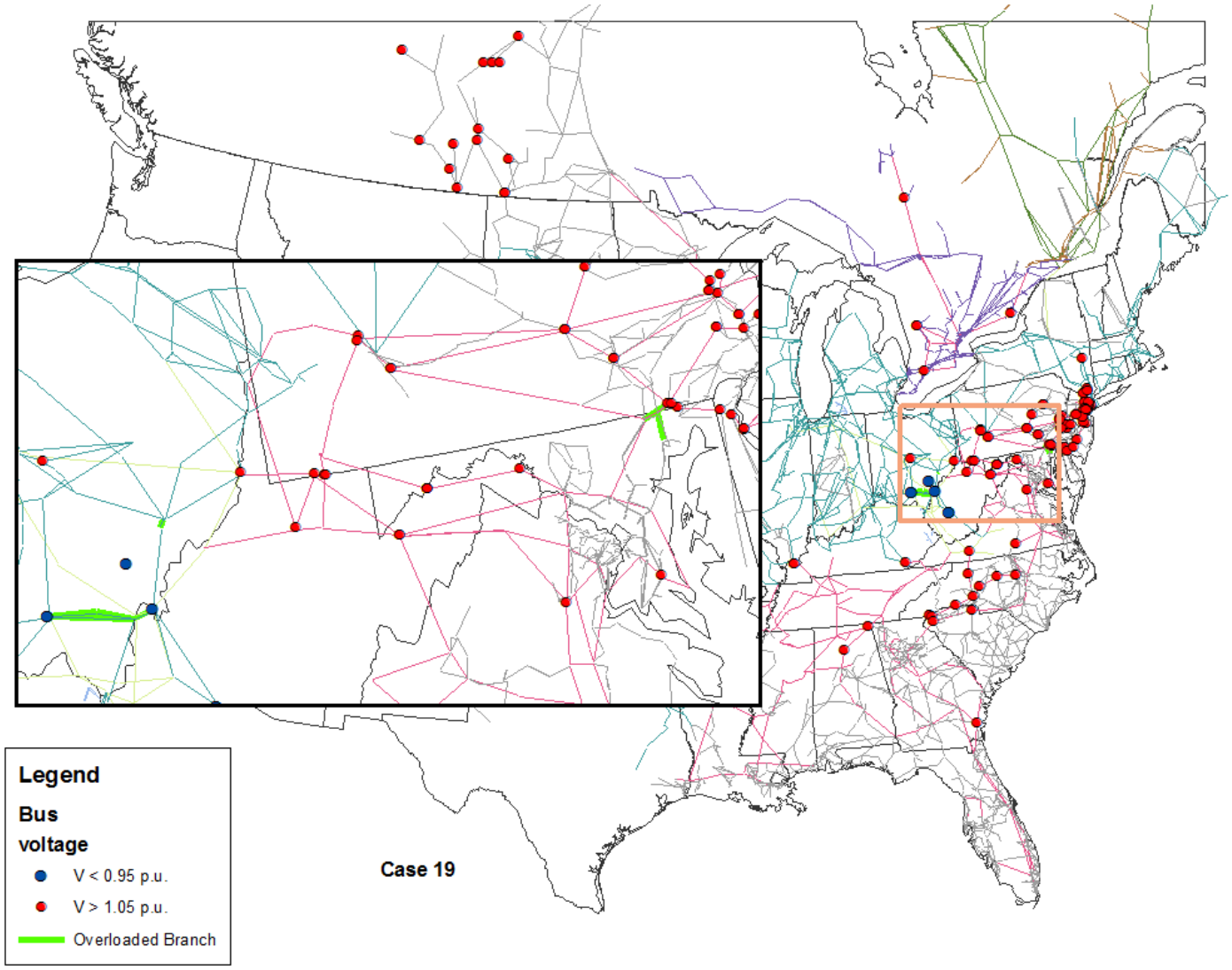


Fig. 7.38: 2016 gas base case with wind

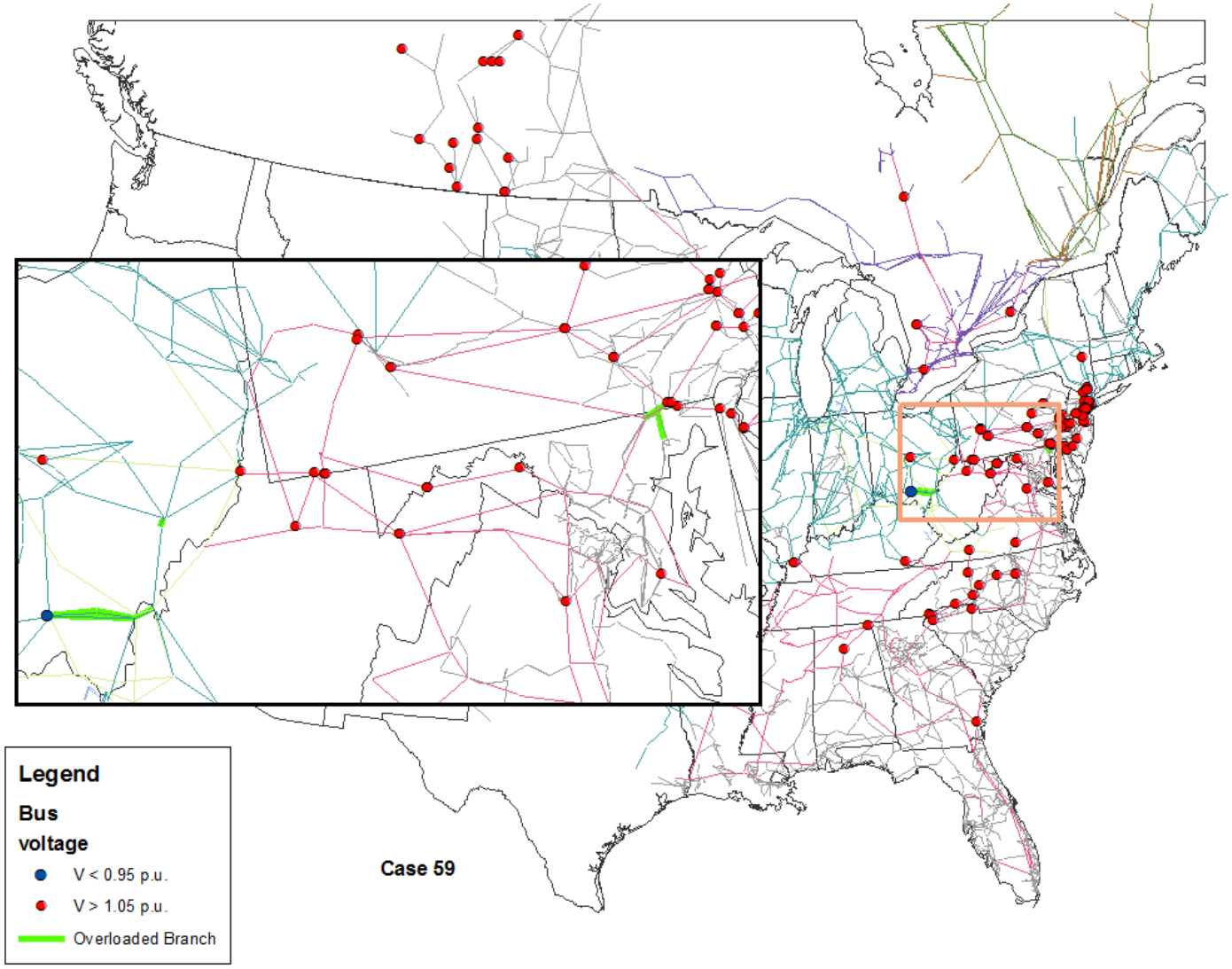


Fig. 7.39: 2016 gas base case with wind generation and synchronous condensers

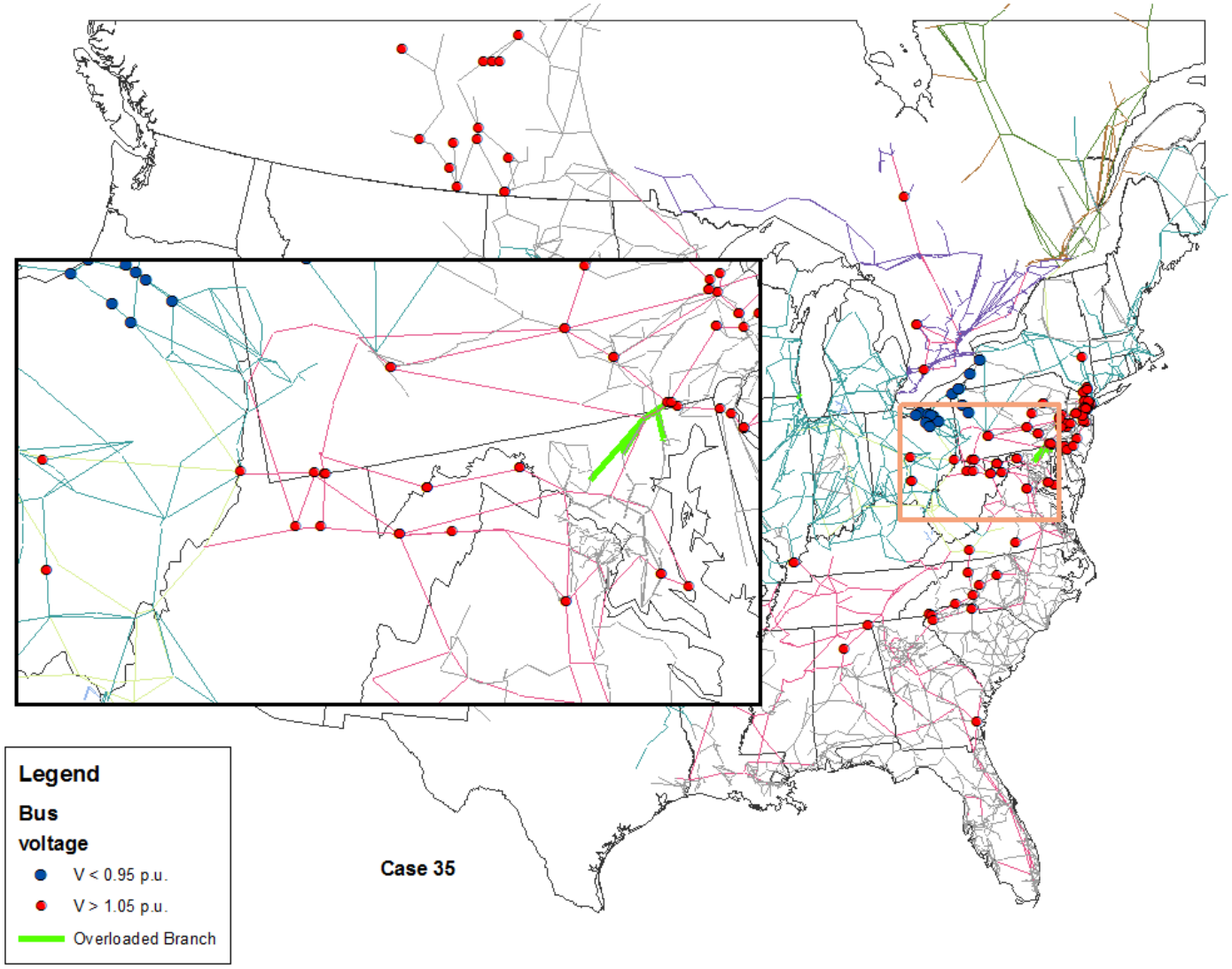


Fig. 7.40: 2016 gas base case with new nuclear units

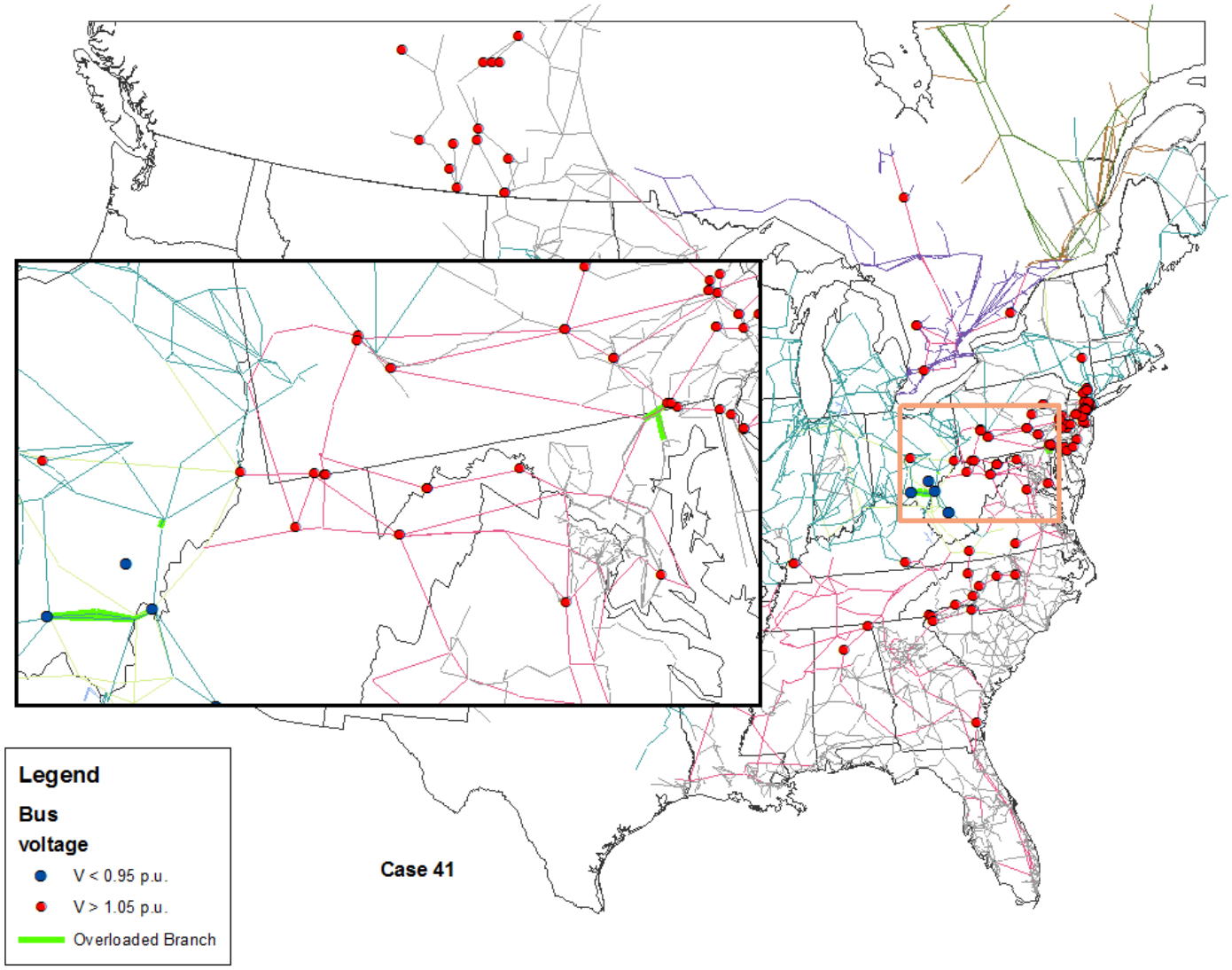


Fig. 7.41: 2016 gas and wind base case with new nuclear units

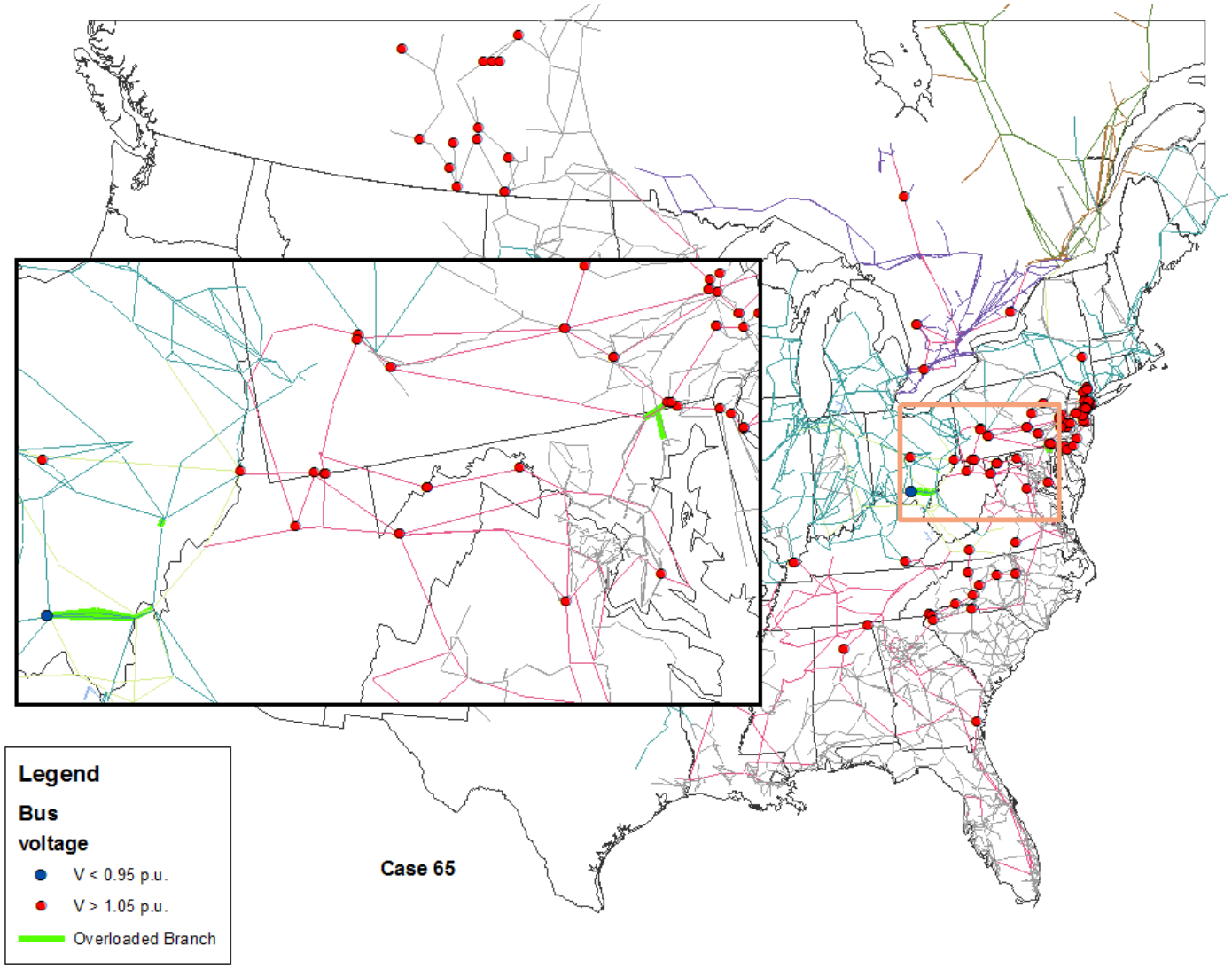


Fig. 7.42: 2016 gas, wind, new nuclear units, and synchronous condensers

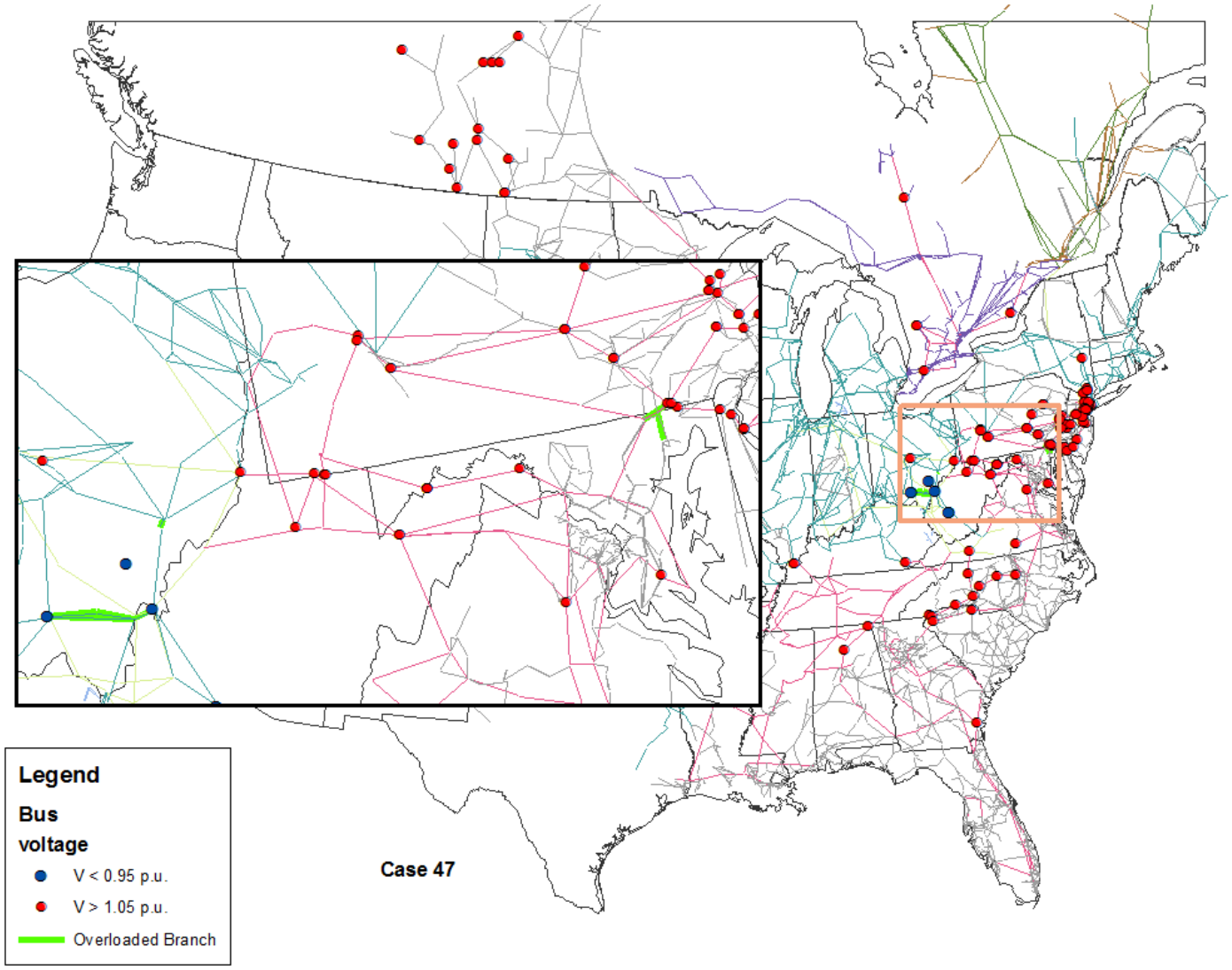


Fig. 7.43: 2016 gas and wind base case, nuclear plants removed

2017

The results of the 2017 base case simulation are shown in Fig. 7.44, and are similar to the base cases from previous years, however there are many fewer overvoltage buses (134). The swing bus output for this simulation was negative, indicating that it was absorbing rather than delivering real power. Normally, this would not happen in a real system. However, the magnitude of the absorbed power remained within the generator limits, so the results may still have some value. Once the MATS/CSAPR-affected generators were removed, the number of overvoltage buses dropped to 125, and two buses in Manitoba were below their acceptable voltage (Fig. 7.45). Also, four additional lines became overloaded. It should be noted that in this simulation, a large number of generators could not be removed successfully, and the swing bus output was several hundred MW above its actual limit. This was also observed for the synchronous condenser case (Fig. 7.46).

While the addition of gas-fired generation seemed to make the case slightly easier to solve, there were still 17 generators that could not be removed (Fig. 7.47). The swing bus output of the solved model was 7,378 MW, well above the actual limits, and this resulted in several additional overloaded lines. The introduction of wind generation (Fig. 7.48) seemed to create some difficulties, resulting in a large swing bus output and corresponding voltage problems. Synchronous condensers did not prove helpful in significantly improving bus voltages or alleviating overloaded lines, and resulted in a large swing bus output (Fig. 7.49).

Only one nuclear reactor (Vogtle, Unit 4) is scheduled to come online in 2017. Compared to the gas base case, the number of out-of-limit buses remained relatively constant, however 7 additional lines became overloaded (Fig. 7.50). As in many of the other 2017 cases, this was due in large part to the high swing bus power output required for the case to converge. Compared to the gas/wind case (Fig. 7.51), there were 9 fewer undervoltage buses and four fewer overloaded lines. Converting some of the removed generators to synchronous condensers appeared to improve the bus voltages in areas where they were too low (Fig. 7.52). For the nuclear shutdown case, the system failed to converge if more than one reactor was removed (Fig. 7.53).

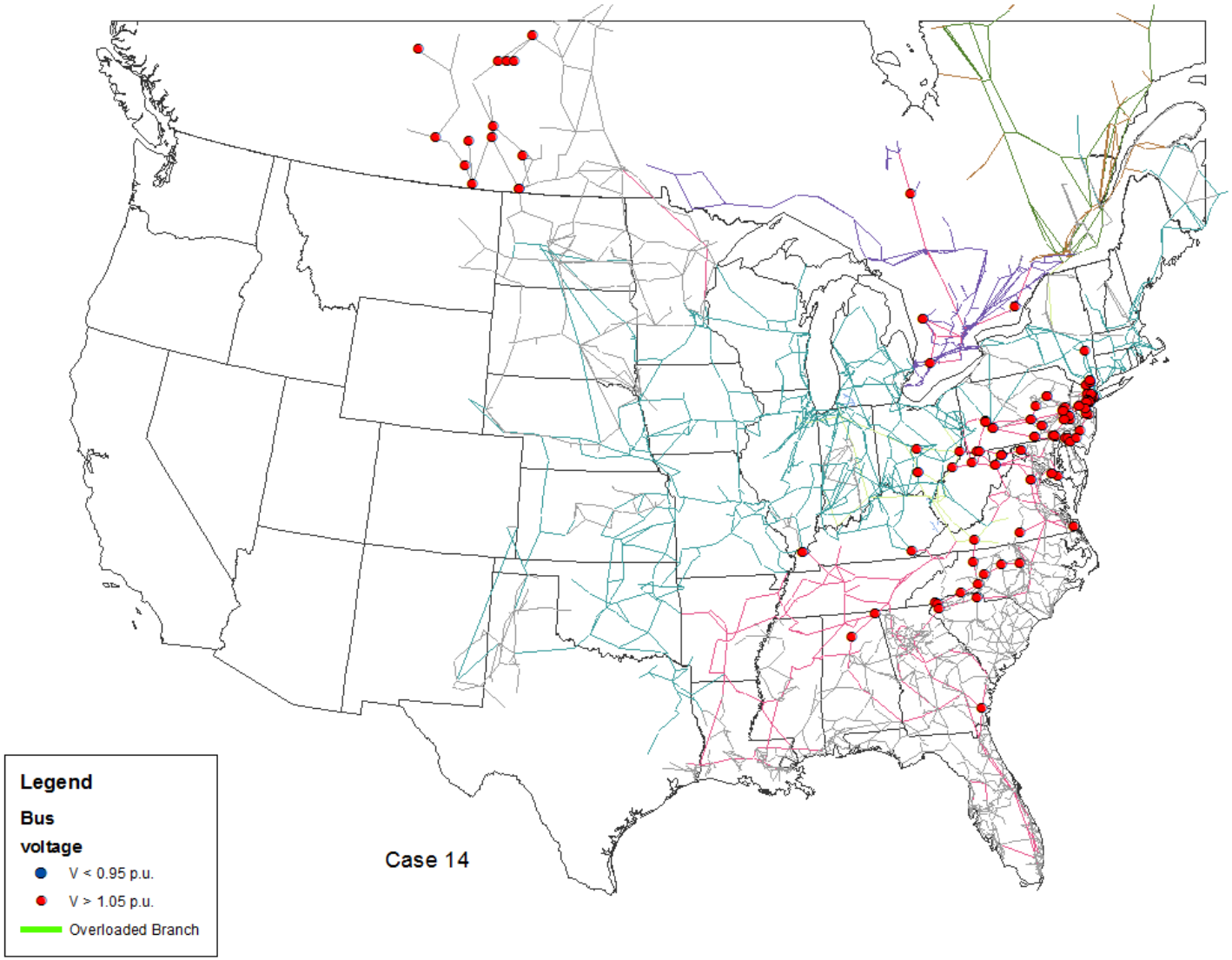


Fig. 7.44: 2017 base case

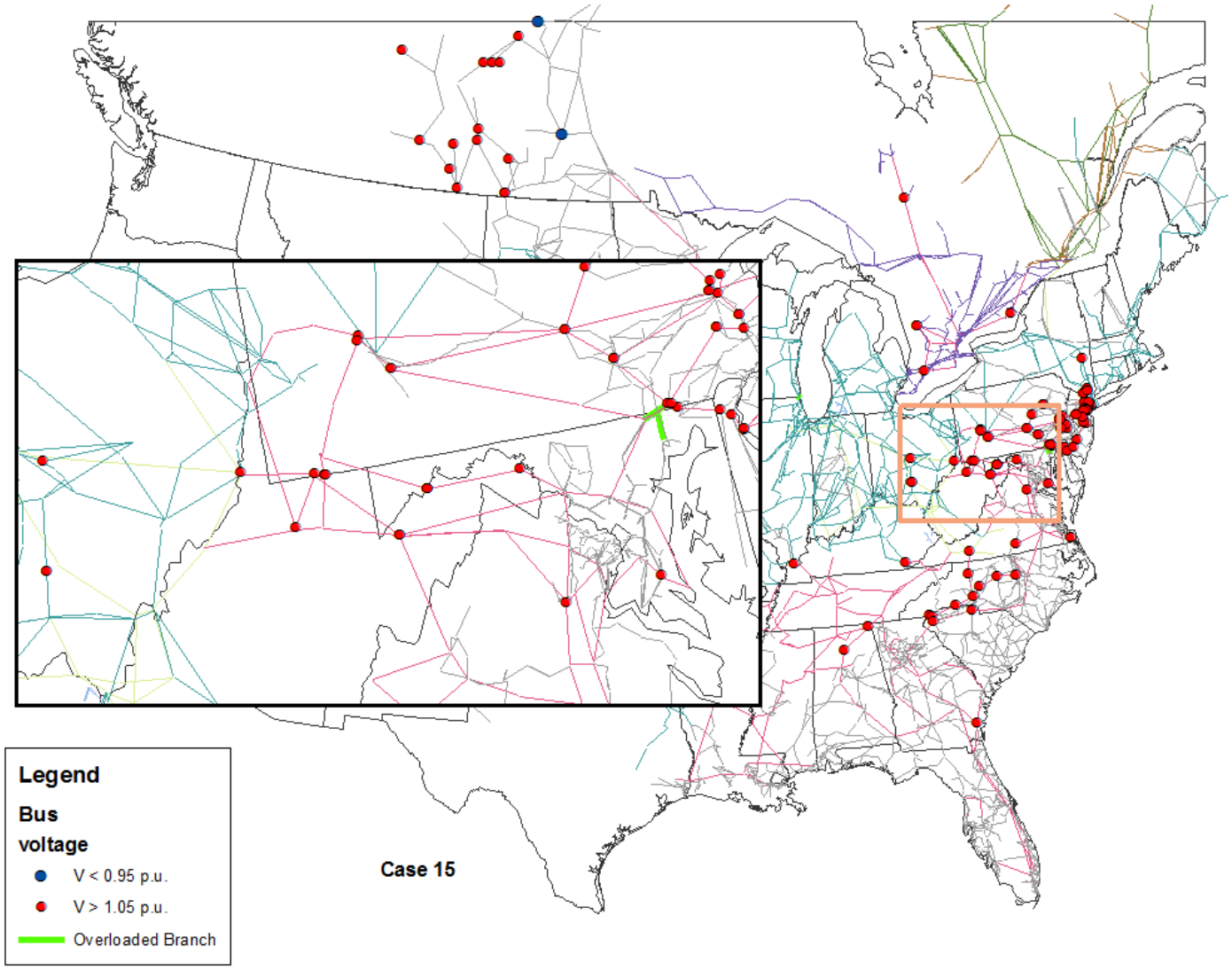


Fig. 7.45: 2017 base case with generators removed

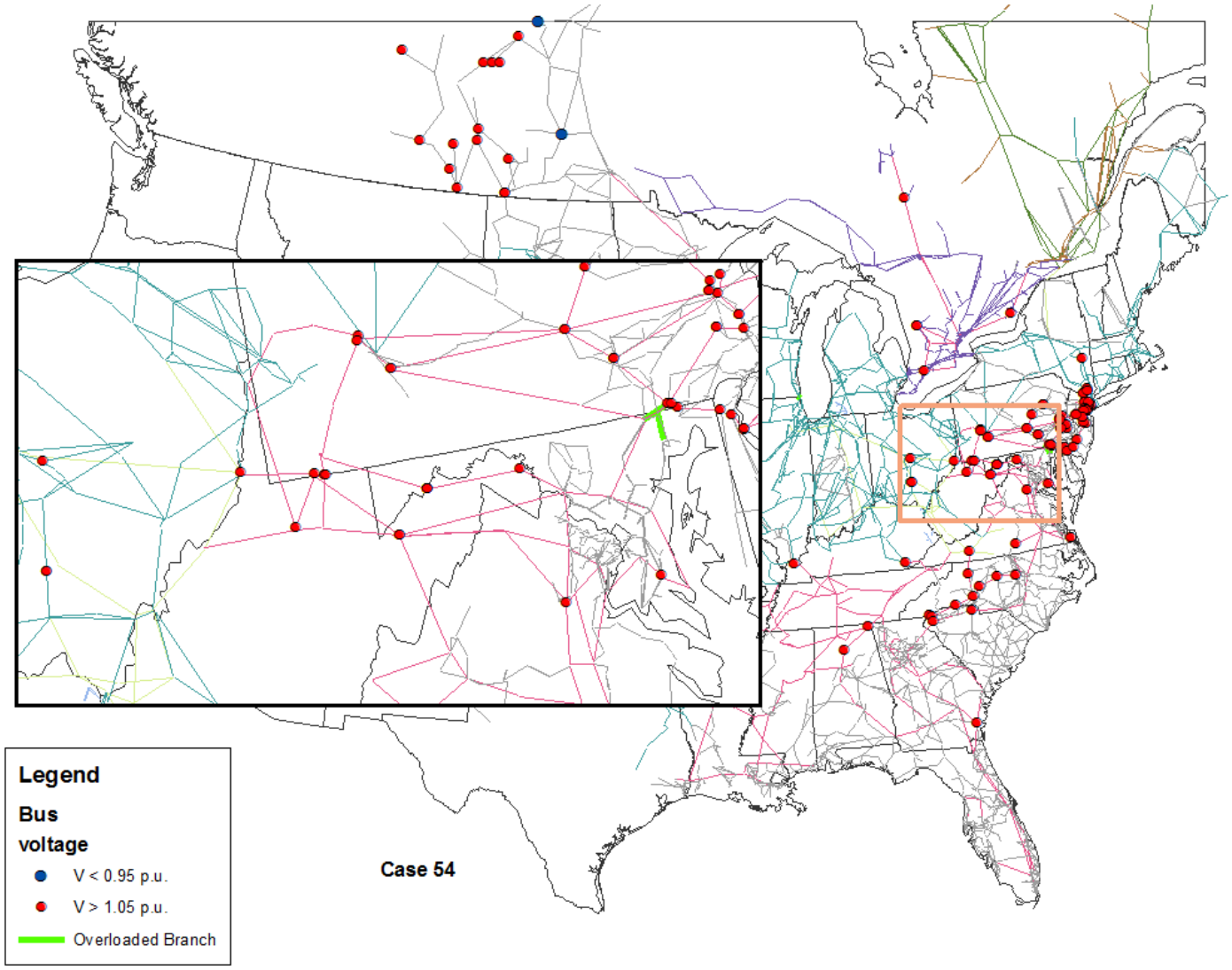


Fig. 7.46: 2017 base case with generators removed, synchronous condensers added

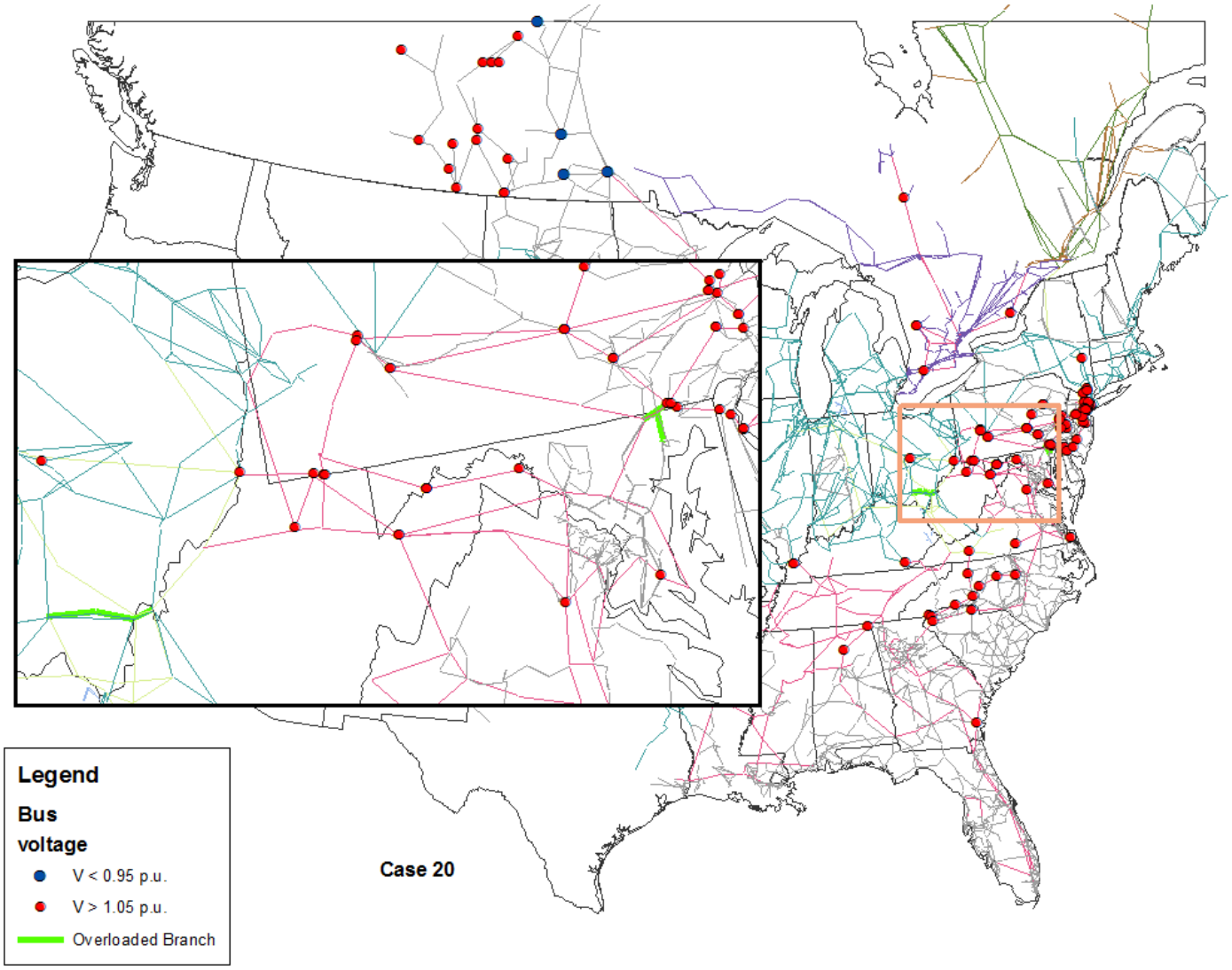


Fig. 7.47: 2017 gas base case

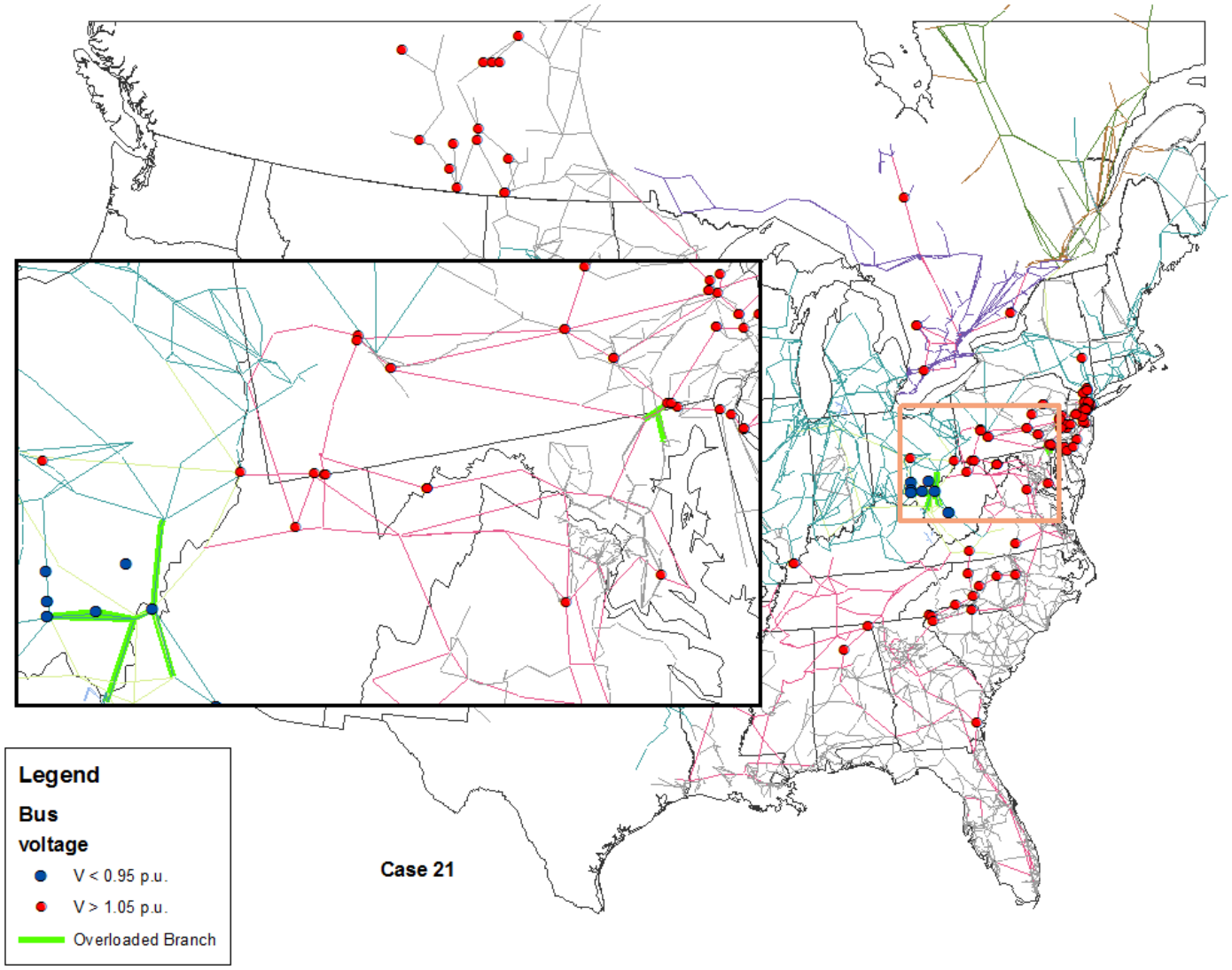


Fig. 7.48: 2017 gas base case with wind

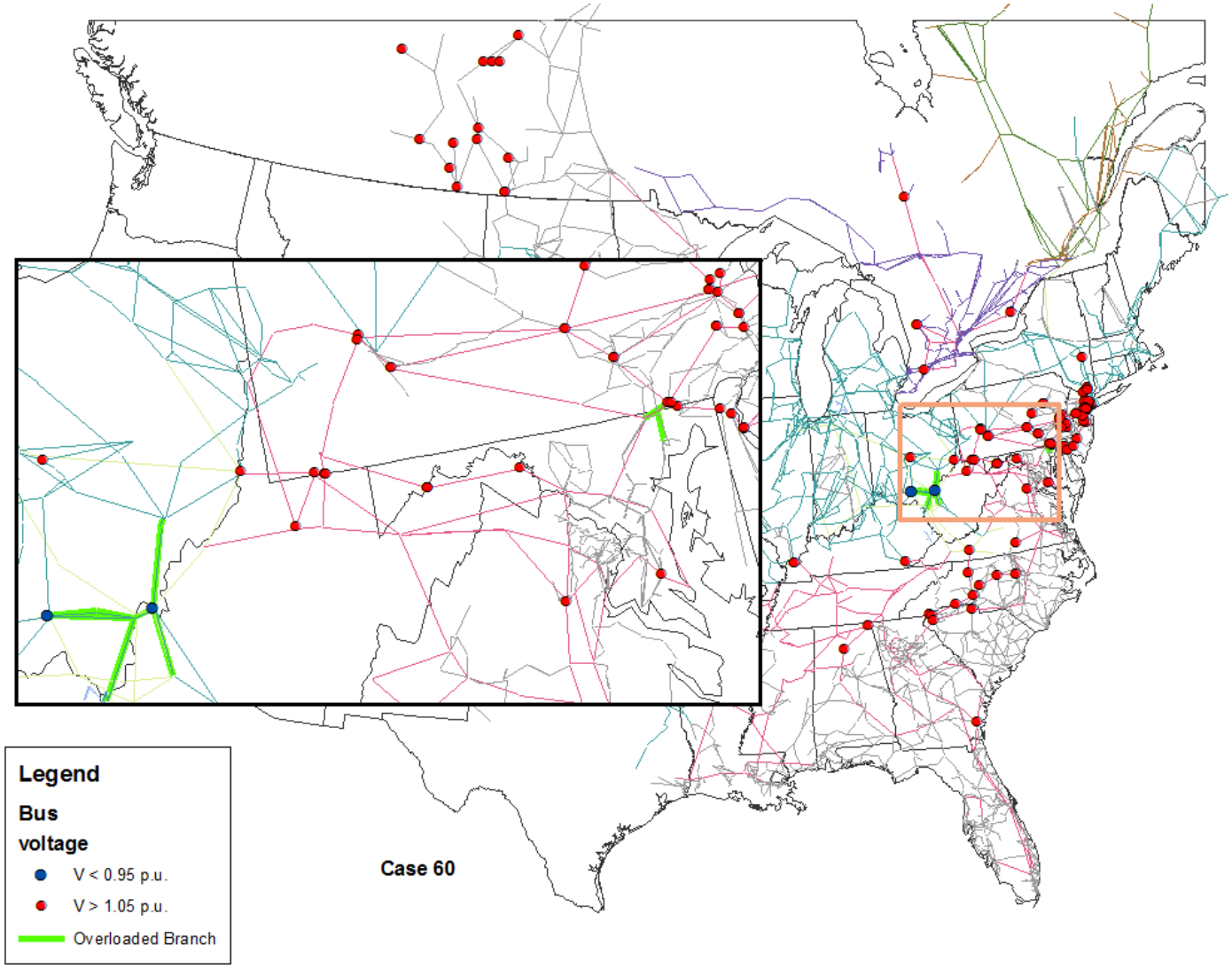


Fig. 7.49: 2017 gas base case with wind generation and synchronous condensers

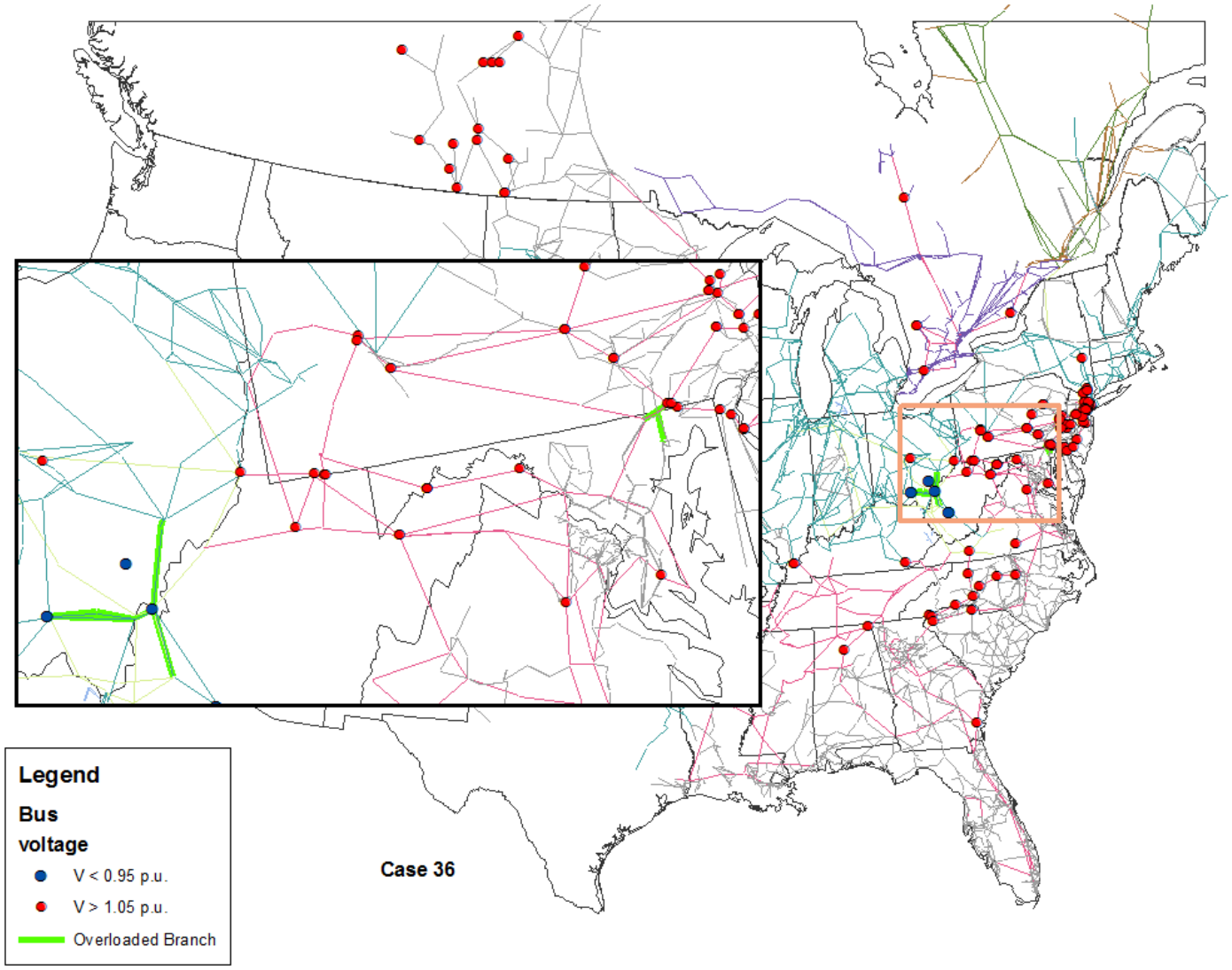


Fig. 7.50: 2017 gas base case with new nuclear units

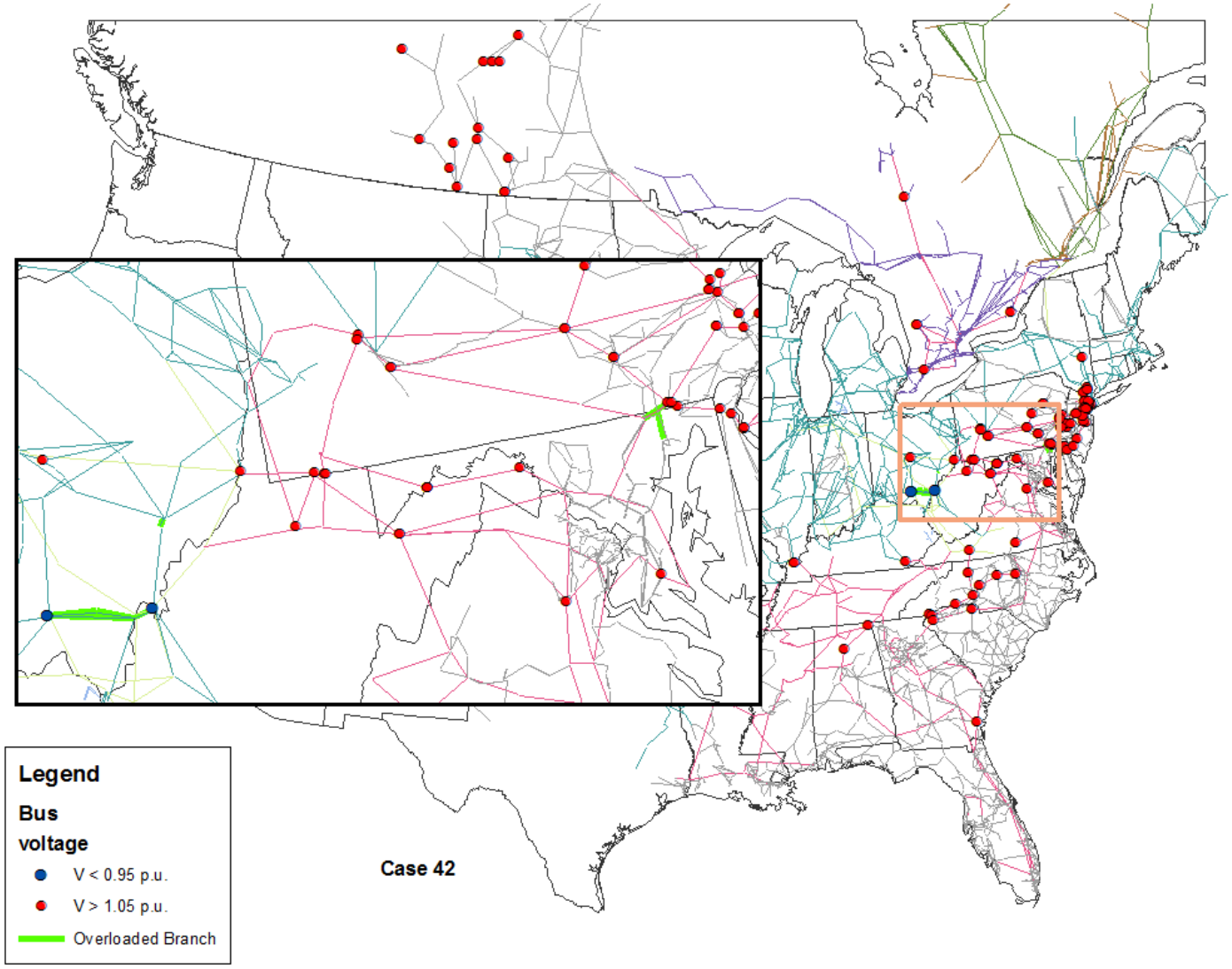


Fig. 7.51: 2017 gas and wind base case with new nuclear units

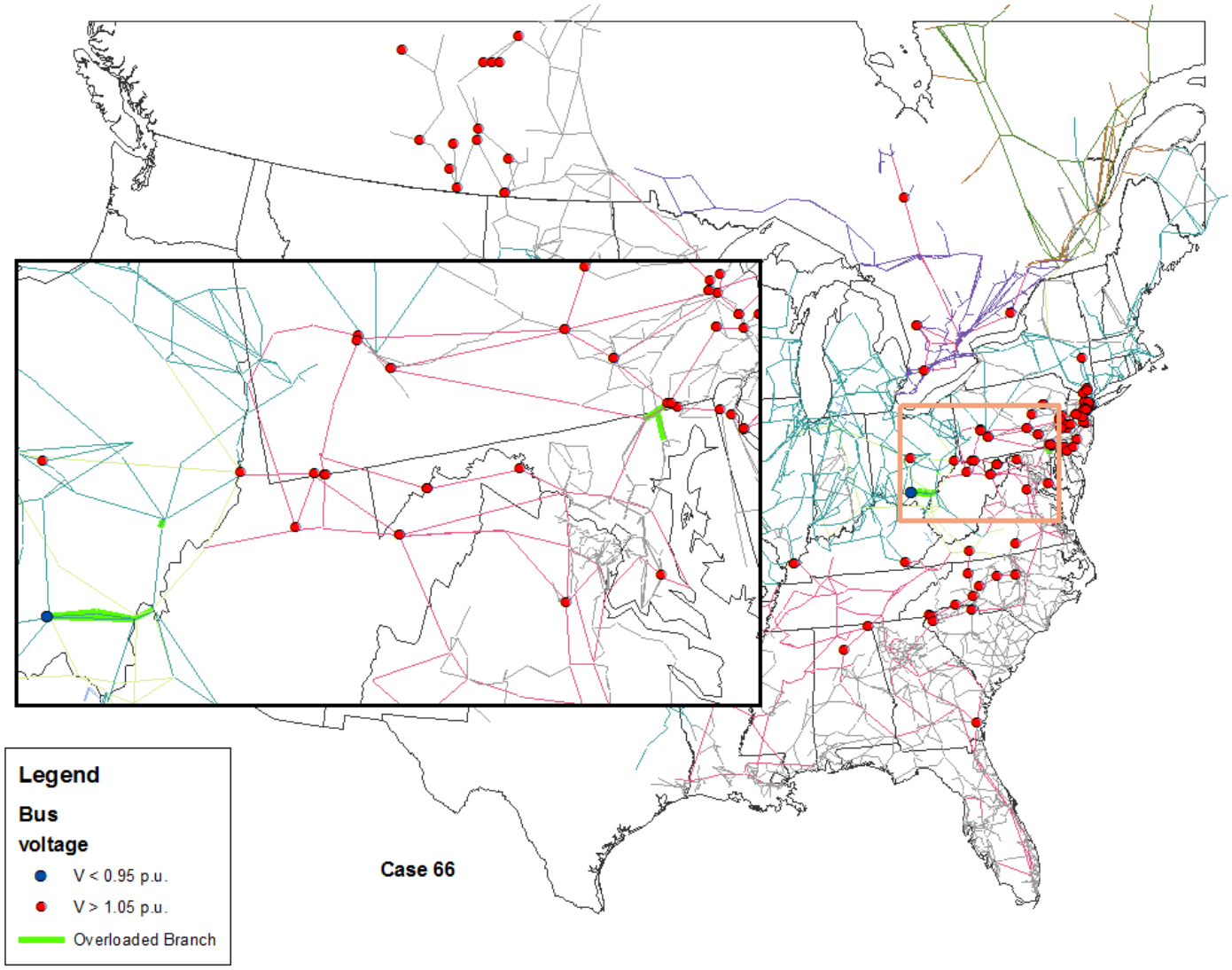


Fig. 7.52: 2017 gas and wind base case with new nuclear units and synchronous condensers

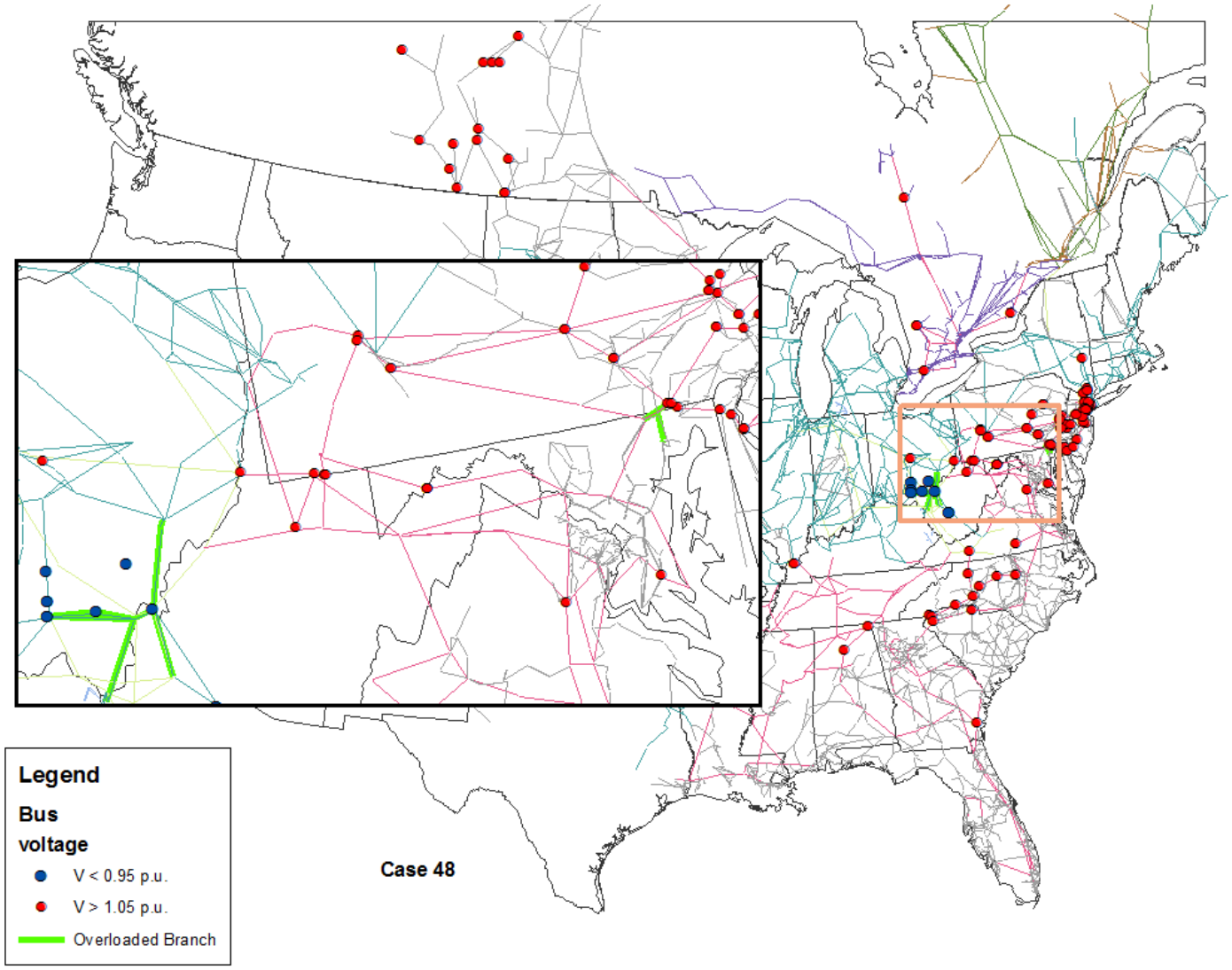


Fig. 7.53: 2017 gas and wind base case, nuclear plants removed

Conclusions

This chapter has presented a comprehensive study of the impacts of the Cross-State Air Pollution Rule and the Mercury and Air Toxics Standards on the Eastern Interconnection. Sixty power flow cases were constructed using a 29,000-bus PSS/E model. These cases were based upon planned generator deactivations and anticipated construction of new generating capacity, including gas, wind, and nuclear plants. Additionally, several cases were developed to examine what would happen if all nuclear reactors in the EI were to be shut down. The major conclusions of this study are as follows:

- It does not appear that there would be widespread voltage stability or line overloading issues in the high-voltage transmission system as a result of MATS and CSAPR. Compared to the overall size of the system, very few buses and lines would be significantly affected in a negative manner. This study did not examine the reliability impacts of these regulations, however, since this has been done by others.
- Low bus voltages were noted in the northeastern Ohio/Lake Erie region, particularly in 2015. Nearly all simulations in this study required that the Bay Shore and/or Eastlake power plants in this area be set as synchronous condensers in order for the model to converge properly after the MATS/CSAPR generators were removed. Since low bus voltages are indicative of a lack of reactive power, it is likely that remedial measures will need to be taken in this area. The results of this study seem to agree with recent announcements by the owner of the Eastlake plant indicating that some of its generators will need to remain operating for a few more years until additional transmission capacity can be built, while others will be converted to synchronous condensers [78, 79].
- Conversion of MATS/CSAPR-affected generators to synchronous condensers is probably not worthwhile, except where previously noted. However, this option may be appropriate for some utilities, depending on their own needs and planning requirements.
- Newly added gas-fired generation will not create major voltage stability problems or cause existing transmission lines to be overloaded. In some cases, the introduction of wind power may result in lowered bus voltages due to altered power flows and system operators will need to plan accordingly.
- A total shutdown of all nuclear power plants in the Eastern Interconnection could create significant voltage stability issues, not to mention the impact on reserve requirements. Such a scenario is highly unlikely, and should remain so.

8. Conclusion

The electric power industry is experiencing a transformation unlike anything seen in its 130-year history. Advances in computing and electronics technology have led to the creation of intelligent devices that will fundamentally alter how the grid is monitored and controlled. Phasor measurement units (PMUs) and related devices will form the basis of wide-area measurement systems (WAMS) that will play a critical role in the power grid's operation for the foreseeable future. This creates a new problem, since the enormous amount of data produced by these systems will need to be turned into actionable information quickly and reliably in order for them to be truly useful. Data mining and machine learning techniques will likely be an important part of the solution. At the same time, traditional sources of power generation in the United States are being replaced with greener, more efficient technologies, and this transition presents additional challenges to utilities and system operators. The research presented in this dissertation addresses both of these important topics.

Much of the work discussed here utilized data obtained from the FNET wide-area measurement system. FNET is unique among other WAMS in the sense that it covers entire interconnections, as opposed to a single utility's service territory. Thus, it provides a perspective that cannot be found elsewhere. Even more importantly, because Frequency Disturbance Recorders are essentially single-phase PMUs, the techniques developed in this research can easily be applied to PMU data.

A number of uses for FNET measurements have been developed in recent years, and this dissertation builds upon those contributions by applying methods borrowed from computer science and statistics. Clustering techniques allowed patterns of frequency extrema to be found in the FNET data that had previously gone unnoticed. Artificial neural networks provided a straightforward means for quickly and accurately identifying system disturbances over much shorter time intervals than was previously possible. Examination of the time delay of frequency decline at closely situated FDRs led to the creation of electromechanical transient speed maps that could improve FNET's disturbance triangulation accuracy. As an added benefit, the software developed for this research can serve as a framework for future application development.

A significant portion of this work was devoted to quantifying the effects of new environmental regulations on the Eastern Interconnection (EI). Despite significant changes in how and where power will be produced in the system, it appears that from a voltage stability perspective, few significant problems will result from the anticipated generator deactivations. In those areas identified as having a lack of reactive power, utilities have begun taking steps to forestall possible voltage issues. Other utilities may

need to take corrective action to accommodate the addition of new natural gas-fired generators and wind turbines, which could lower bus voltages slightly in some areas. Although a small number of generators may need to be operated as synchronous condensers in order to provide voltage support in certain areas, widespread conversion of deactivated generators appears unnecessary. Finally, attempts to simulate the removal of all nuclear power plants from the EI revealed that such an action would likely result in significant voltage stability problems.

Contributions

The major contributions of this research include:

- A survey of existing data mining and machine learning applications for wide-area measurement data from electric power systems.
- A review of recent efforts to assess the effects of environmental regulations on the power grid of North America.
- A survey of previous work on the development of electromechanical transient speed maps for power systems.
- An in-depth analysis of daily frequency extrema occurring in the Eastern and Western Interconnections, as well as an algorithm for locating the extrema in the FNET data.
- A neural-network based classifier for identifying different types of power system disturbances based solely on their frequency signatures, as well as the determination of the minimum signal length needed for such a classifier.
- A new method for calculating the propagation speeds of electromechanical transients in a power grid that does not rely upon knowing the exact time and location of the disturbance.
- A MATLAB-based software framework that greatly simplifies the process of developing FNET data analysis applications and allows for the creation of virtual, software-based FDRs. This framework could also be extended to support PMU data.
- A comprehensive study of the impacts of possible generator deactivations due to the Cross-State Air Pollution Rule and the Mercury and Air Toxics Standards on the Eastern Interconnection over the next five years.
- The development of Python scripts used to automate the process of modifying and solving an existing power system model in PSS/E using a database of scenario descriptions.

- The creation of power flow models simulating the possible evolution of the Eastern Interconnection from 2012-2017 with regards to load growth, unit deactivations, and new generation.

Future Work

- The analysis of frequency extrema in the EI and WECC has been repeated each year for the past two years as new data became available. Members of the power engineering community have expressed an interest in seeing this work continued for the foreseeable future, particularly as electrical markets become more closely aligned with those of natural gas.
- Artificial neural network-based disturbance classifiers should be implemented as part of the openPDC platform. Ideally, these could be made available as modules that other members of the synchrophasor community can download and install for their own use. Other types of disturbances, like fault-induced delayed voltage recovery (FIDVR) events, could also be added to the classifier.
- The electromechanical transient speed calculation technique can be modified in a number of ways. For example, instead of using the simple mean of speeds between a sensor and its neighbors, the median speed, or a distance-weighted average could be used. It may also be worthwhile to use angle measurements rather than frequency, since those contain less noise. Automatic determination of the cutoff frequency (or angle) used by the algorithm would allow it to be run unsupervised on a routine basis, which is necessary for the maps to be used for disturbance triangulation.
- A decline in the frequency response of the EI has been observed for the past few years. It is likely that this trend will continue as coal-fired generators with large inertia constants are replaced with relatively low-inertia gas turbines. The power flow cases developed for the Eastern Interconnection study in Chapter 7 could be used to build dynamics models to study the frequency response implications of generators being taken offline.

List of References

[1] J. N. Bank*, et al.*, "Generator Trip Identification Using Wide-Area Measurements and Historical Data Analysis," in *Power Systems Conference and Exposition, 2006. PSCE '06. 2006 IEEE PES*, 2006, pp. 1677-1681.
[2] R. M. Gardner*, et al.*, "Non-Parametric Power System Event Location Using Wide-Area Measurements," in *Power Systems Conference and Exposition, 2006. PSCE '06. 2006 IEEE PES*, 2006, pp. 1668-1675.
[3] R. M. Gardner*, et al.*, "Power system event location analysis using wide-area measurements," in *Power Engineering Society General Meeting, 2006. IEEE*, 2006, p. 7 pp.
[4] J. K. Wang*, et al.*, "Analysis of system oscillations using wide-area measurements," in *Power Engineering Society General Meeting, 2006. IEEE*, 2006, p. 6 pp.
[5] Z. Yuan*, et al.*, "Inter-area oscillation analysis using wide area voltage angle measurements from FNET," in *Power and Energy Society General Meeting, 2010 IEEE*, 2010, pp. 1-7.
[6] J. N. Bank*, et al.*, "Visualization of Wide-Area Frequency Measurement Information," in *Power Engineering Society General Meeting, 2007. IEEE*, 2007, pp. 1-8.
[7] T. Xia*, et al.*, "Wide-area Frequency Based Event Location Estimation," in *Power Engineering Society General Meeting, 2007. IEEE*, 2007, pp. 1-7.
[8] Y. Liu*, et al.*, "An accurate electrical network frequency reference database-FNET system," presented at the Scientific Working Group on Digital Evidence meeting, Denver, CO, 2009.
[9] Y. Zhang*, et al.*, "Wide-Area Frequency Monitoring Network (FNET) Architecture and Applications," *Smart Grid, IEEE Transactions on,* vol. 1, pp. 159-167, 2010.
[10] Z. Zhong*, et al.*, "Power system frequency monitoring network (FNET) implementation," *Power Systems, IEEE Transactions on,* vol. 20, pp. 1914-1921, 2005.
[11] Y. Liu, "A US-wide power systems frequency monitoring network," in *Power Engineering Society General Meeting, 2006. IEEE*, 2006, p. 8 pp.
[12] R. M. Gardner and Y. Liu, "FNET: A Quickly Deployable and Economic System to Monitor the Electric Grid," in *Technologies for Homeland Security, 2007 IEEE Conference on*, 2007, pp. 209-214.
[13] (17 September 2012). *SCADA History*. Available: http://energy.sandia.gov/?page_id=5508.
[14] N. Dahal*, et al.*, "Online dimension reduction of synchrophasor data," in *Transmission and Distribution Conference and Exposition (T&D), 2012 IEEE PES*, 2012, pp. 1-7.
[15] M. Kezunovic*, et al.*, "An expert system for transmission substation event analysis," *Power Delivery, IEEE Transactions on,* vol. 8, pp. 1942-1949, 1993.
[16] M. Kezunovic and I. Rikalo, "Detect and classify faults using neural nets," *Computer Applications in Power, IEEE,* vol. 9, pp. 42-47, 1996.
[17] A. Poeltl and K. Frohlich, "Two new methods for very fast fault type detection by means of parameter fitting and artificial neural networks," *Power Delivery, IEEE Transactions on,* vol. 14, pp. 1269-1275, 1999.
[18] P. Trachian, "Machine learning and windowed subsecond event detection on PMU data via Hadoop and the openPDC," in *Power and Energy Society General Meeting, 2010 IEEE*, 2010, pp. 1-5.
[19] E. E. Bernabeu*, et al.*, "Methodology for a Security/Dependability Adaptive Protection Scheme Based on Data Mining," *Power Delivery, IEEE Transactions on,* vol. 27, pp. 104-111, 2012.
[20] E. E. Bernabeu, "Methodology for a Security-Dependability Adaptive Protection Scheme based on Data Mining," Ph.D. Dissertation, Department of Electrical and Computer Engineering, Virginia Polytechnic Institute and State University, Blacksburg, VA, 2009.
[21] I. Kamwa*, et al.*, "On the Accuracy Versus Transparency Trade-Off of Data-Mining Models for Fast-Response PMU-Based Catastrophe Predictors," *Smart Grid, IEEE Transactions on,* vol. 3, pp. 152-161, 2012.

[22] M. He*, et al.*, "A data mining framework for online dynamic security assessment: Decision trees, boosting, and complexity analysis," in *Innovative Smart Grid Technologies (ISGT), 2012 IEEE PES*, 2012, pp. 1-8.
[23] K. Sun*, et al.*, "An online dynamic security assessment scheme using phasor measurements and decision trees," in *Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century, 2008 IEEE*, 2008, pp. 1-1.
[24] K. Sun*, et al.*, "An Online Dynamic Security Assessment Scheme Using Phasor Measurements and Decision Trees," *Power Systems, IEEE Transactions on,* vol. 22, pp. 1935-1943, 2007.
[25] R. Diao*, et al.*, "Design of a Real-Time Security Assessment Tool for Situational Awareness Enhancement in Modern Power Systems," *Power Systems, IEEE Transactions on,* vol. 25, pp. 957-965, 2010.
[26] V. Vittal*, et al.*, "A Tool for On-line Stability Determination and Control for Coordinated Operations between Regional Entities Using PMUs," Power Systems Engineering Research Center, January 2008.
[27] F. Hashiesh*, et al.*, "A wide area synchrophasor based ANN transient stability predictor for the Egyptian Power System," in *Innovative Smart Grid Technologies Conference Europe (ISGT Europe), 2010 IEEE PES*, 2010, pp. 1-7.
[28] F. R. G. Lezama, "Prediction and Control of Transient Instability Using Wide Area Phasor Measurements," Ph.d. Dissertation, Department of Electrical and Computer Engineering, University of Manitoba, 2011.
[29] (31 January 2012). *Cross-State Air Pollution Rule (CSAPR) - Basic Information*. Available: http://epa.gov/airtransport/basic.html.
[30] "Fact Sheet: Mercury and Air Toxics Rule for Power Plants," U. S. E. P. Agency, Ed., ed, 2011.
[31] "Update on the Impact of EPA's Regulatory Assault on Power plants: New Regulations to Take 30 GW of Electricity Generation Offline and the Plant Closing Announcements Keep Coming…," Institute for Energy Research, Washington, D.C.
[32] "2010 Special Reliability Scenario Assessment: Resource Adequacy Impacts of Potential U.S. Environmental Regulations," North American Electric Reliability Corporation.
[33] "2012 Summer Reliability Assessment," North American Electric Reliability Corporation, Atlanta, Georgia, May 2012.
[34] "Resource Adequacy Implications of Forthcoming EPA Air Quality Regulations," U.S. Department of Energy, Washington, D.C., December 2011.
[35] K. S. Kook, "Dynamic Model Based Novel Findings in Power Systems Analysis and Frequency Measurement Verification," Ph.D. Dissertation, Department of Electrical and Computer Engineering, Virginia Polytechnic Institute and State University, Blacksburg, VA, 2007.
[36] S. S. Tsai*, et al.*, "Study of global frequency dynamic behavior of large power systems," in *Power Systems Conference and Exposition, 2004. IEEE PES*, 2004, pp. 328-335 vol.1.
[37] A. J. Arana*, et al.*, "Estimating Speed of Frequency Disturbance Propagation Through Transmission and Distribution Systems," in *Power Systems Conference and Exposition, 2006. PSCE '06. 2006 IEEE PES*, 2006, pp. 1286-1290.
[38] J. Bank, "Propagation of electromechanical disturbances across large interconnected power systems and extraction of associated modal content from measurement data," Ph.D. Dissertation, Department of Electrical and Computer Engineering, Virginia Polytechnic Institute and State University, Blacksburg, 2009.
[39] J. S. Thorp*, et al.*, "Electromechanical wave propagation in large electric power systems," *Circuits and Systems I: Fundamental Theory and Applications, IEEE Transactions on,* vol. 45, pp. 614-622, 1998.

[40] R. M. Gardner, "A Wide-Area Perspective on Power System Operation and Dynamics," Ph.D. Dissertation, Department of Electrical and Computer Engineering, Virginia Polytechnic Institute and State University, Blacksburg, VA, 2008.
[41] S. Backhaus and Y. Liu, "Electromechanical Wave Green's Function Estimation from Ambient Electrical Grid Frequency Noise," in *System Science (HICSS), 2012 45th Hawaii International Conference on*, 2012, pp. 2054-2061.
[42] P. Kundur, *Power System Stability and Control*: McGraw-Hill, 1994.
[43] (March 6). *Access Specifications*. Available: http://office.microsoft.com/en-us/access-help/access-specifications-HP005186808.aspx.
[44] S. S. Tsai*, et al.*, "Analysis of wide-area frequency measurement of bulk power systems," in *Power Engineering Society General Meeting, 2006. IEEE*, 2006, p. 8 pp.
[45] P. N. Markham*, et al.*, "Analysis of frequency extrema in the eastern and western interconnections," in *Power and Energy Society General Meeting, 2011 IEEE*, 2011, pp. 1-8.
[46] L. Kaufman and P. Rousseeuw, *Finding Groups in Data: An Introduction to Cluster Analysis*. Hoboken, NJ: Wiley, 2005.
[47] P.-N. Tan*, et al.*, *Introduction to Data Mining*: Addison-Wesley, 2006.
[48] J. W. Ingleson and E. Allen, "Tracking the Eastern Interconnection frequency governing characteristic," in *Power and Energy Society General Meeting, 2010 IEEE*, 2010, pp. 1-6.
[49] B. J. Kirby*, et al.*, "Frequency Control Concerns In the North American Electric Power System," Oak Ridge National Laboratory, December 2002.
[50] "“22:00” Frequency Excursions (Final Report)," Princeton, NJ, August 28, 2002.
[51] H. Innah and T. Hiyama, "Neural network method based on PMU data for voltage stability assessment and visualization," in *TENCON 2011 - 2011 IEEE Region 10 Conference*, 2011, pp. 822-827.
[52] K. Seethalekshmi*, et al.*, "Adaptive distance relaying scheme in presence of UPFC using WAMS," in *Power Systems Conference and Exposition, 2009. PSCE '09. IEEE/PES*, 2009, pp. 1-6.
[53] N. P. Padhy, *Artificial Neural Networks*. New Delhi, India: Oxford UP, 2005.
[54] T. M. Mitchell, *Artificial Neural Networks*: McGraw-Hill, 1997.
[55] (2010). *PSSE 32.0.5 Online Documentation* [Adobe PDF].
[56] (2010). *MATLAB R2010b Help*.
[57] J. Adams*, et al.*, *C++: An Introduction to Computing*, 2nd ed. Upper Saddle River, NJ: Prentice-Hall, 2001.
[58] (3 March). *Cleaner Power Plants*. Available: http://www.epa.gov/mats/powerplants.html.
[59] K. Leonard. "3 more power plants set to close in W.Pa.," *Pittsburgh Tribune-Review*, Available: http://www.pittsburghlive.com/x/pittsburghtrib/business/s_784221.html.
[60] (March 8, 2012). *GenOn Reports 2011 Results and Announces Expected Deactivation of Generation Units*. Available: http://phx.corporate-ir.net/phoenix.zhtml?c=124294&p=irol-newsArticle&ID=1667152&highlight=.
[61] (8 March). *FirstEnergy, Citing Impact of Environmental Regulations, Will Retire Six Coal-Fired Power Plants*. Available: https://www.firstenergycorp.com/newsroom/featured_stories/Coal_Plant_Retirements0.html.
[62] (8 March). *State By State Breakdown*. Available: http://www.aep.com/environmental/NewEPARules/StateByState.aspx.
[63] M. McHenry. (2011). *AEP Shares Plan For Compliance With Proposed EPA Regulations* [Press Release]. Available: http://www.aep.com/newsroom/newsreleases/?id=1697.
[64] R. C. Kuether*, et al.*, "Electricity Transfers and Reliability," North American Electric Reliability Council, October 1989.

[65] K. Begos. "Electric plants turn to natural gas as fuel," *USA Today*, 17 January 2012. Available: http://www.usatoday.com/USCP/PNI/NEWS/2012-01-17-BCUSGas-DrillingElectricity1st-LdWritethru_ST_U.htm.
[66] (19 March). *Institute for Energy Research*. Available: http://www.sourcewatch.org/index.php?title=Institute_for_Energy_Research.
[67] "Annual Electric Generator Report," Washington, D.C., Available: http://www.eia.gov/cneaf/electricity/page/eia860.html
[68] "U.S. Commercial Nuclear Power Reactors," U. S. N. R. Agency, Ed., ed. Washington, D.C.
[69] W. E. Winkler, "String Comparator Metrics and Enhanced Decision Rules in the Fellegi-Sunter Model of Record Linkage," *Proceedings of the Section on Survey Research Methods (American Statistical Association),* pp. 354–359, 1990.
[70] "2008–2017 Regional & National Peak Demand and Energy Forecasts Bandwidths," North American Electric Reliability Corporation, Princeton, NJ, August 2008.
[71] "Annual Electric Generator Report," ed. Washington, D.C.: U.S. Energy Information Administration, 2010.
[72] (23 March). *The Nuclear Renaissance*. Available: http://www.world-nuclear.org/info/inf104.html.
[73] "Expected New Nuclear Power Plant Applications," Nuclear Regulatory Commission, Washington, D.C., 6 October 2011.
[74] E. Marcum. "NRC sets meeting on delayed Watts Bar reactor," *Knoxville News-Sentinel*, 6 April 2012. Available: http://www.knoxnews.com/news/2012/apr/06/tva-says-watts-bar-construction-to-cost-more/.
[75] (23 September 2012). *Germany: Nuclear power plants to close by 2022*. Available: http://www.bbc.co.uk/news/world-europe-13592208.
[76] J. Kanter. "Switzerland Decides on Nuclear Phase-Out," *New York Times*, 25 May 2011. Available: http://www.nytimes.com/2011/05/26/business/global/26nuclear.html?_r=2.
[77] R. D. Bottoms. "RE: question." Personal e-mail (7 September 2012).
[78] (1 September 2012). *FirstEnergy to delay planned coal plant retirements at PJM's request*. Available: http://www.platts.com/RSSFeedDetailedNews/RSSFeed/Coal/6256235.
[79] J. Funk. (30 September 2012). *FirstEnergy will keep older power plants open until 2015, launch nearly $1 billion in transmission upgrades*. Available: http://www.cleveland.com/business/index.ssf/2012/05/firstenergy_will_keep_some_of.html.

Appendixes

Appendix A – List of Currently Deployed FDRs in the EI

Unit Number	FDR Name	City	State
523	UsTNKnoxville523	Knoxville	TN
601	UsVaRichmond601	Richmond	VA
616	UsMoKansasCity616	Kansas City	MO
620	UsIlChicago620	Chicago	IL
621	UsMiWayneState621	Detroit	MI
623	UsFlPlantcity623	Plant City	FL
624	UsMoMst624	Rolla	MO
639	UsVaVtech639	Blacksburg	VA
647	UsLaEntergy647	New Orleans	LA
661	UsVaAri661	Alexandria	VA
663	UsFlUfl663	Gainesville	FL
664	UsNdBismarck664	Bismarck	ND
665	CaMbWinnipeg665	Winnipeg	MB
666	UsCtDanbury666	Danbury	CT
667	UsNyLeroy667	Leroy	NY
668	UsVaNewportNews668	Newport News	VA
669	UsScFarris669	Simpsonville	SC
670	UsOhChillicothe670	Gahanna	OH
671	UsAlSoco671	Birmingham	AL
672	UsMsGolfport672	Gulfport	MS
673	UsAlMontgomery673	Montgomery	AL
674	UsFlPensacola674	Pensacola	FL
675	UsGaAtlanta675	Atlanta	GA
678	UsNjNerc678	Princeton	NJ
679	UsMiCalvin679	Grand Rapids	MI
681	UsIaIsu681	Ames	IA
682	UsMaNeiso682	Holyoke	MA
683	UsTxTexasTech683	Lubbock	TX
684	UsMaBoston684	Waltham	MA
686	UsVaRvcs686	Roanoke	TX
688	UsWvCharleston688	Charleston	WV
690	UsVaBlacksburg690	Blacksburg	VA
692	UsTnUtk692	Knoxville	TN
696	UsOhCleveland696	Cleveland	OH
703	CaOnToronto703	Markham	ON
704	UsFlFsu704	Tallahassee	FL
705	UsPaPjm705	Norristown	PA

706	UsNcNcat706	Greensboro	NC
707	UsNyRpi707	Troy	NY
710	UsInImpa710	Carmel	IN
712	UsMiMitech712	Houghton	MI
713	UsInNipsc713	Hammond	IN
714	UsIaMidaec714	Urbandale	IA
715	CaMbGrandrapids715	Grand Rapids	MB
718	CaMbKelsey718	Kelsey	MB
719	CaMbThompson719	Thompson	MB
720	UsMnOtpc720	Fergus Falls	MN
722	UsTnORNL722	Oak Ridge	TN
726	UsNdMduwilliston726	Williston	ND
728	UsVaDoe728	Fredericksburg	VA
729	UsNdMdudickinson729	Dickinson	ND
730	UsNdMduglendive730	Glendive	MT
733	UsMeBangor733	Bangor	ME
739	UsSdBigstone739	Big Stone	SD
740	UsNdDevilslake740	Devil's Lake	ND
744	UsWvWvu744	Morgantown	WV
747	UsMnCrystal747	Crystal	MN
749	UsMoGape749	Gape	MO
750	UsMiAtc750	Kingsford	MI
753	UsNcHampstead753	Hampstead	NC
754	UsIaCedarfalls754	Cedar Falls	IA
755	UsIlUiuc755	Urbana	IL
756	UsMoFranklin756	Union	MO
759	UsOkNorman759	Norman	OK
760	UsMnElkriver760	Elk River	MN
762	UsSCCharleston762	Charleston	SC
767	UsIlMarion767	Marion	IL
770	UsTnKnoxsolar770	Knoxville	TN
777	UsNeLes777	Lincoln	NE
778	UsIlMatton778	Matton	IL
781	UsMoKirksville781	Kirksville	MO
785	UsVaBlacksburg785	Blacksburg	VA
786	UsFlFsu786	Tallahassee	FL
790	UsMnGre790	Elk River	MN
796	UsFLNPalmBeach796	North Palm Beach	FL
797	UsInNipsco797	Hammond	IN
803	UsOhAep803	Gahanna	OH
830	UsOhDuke830	Cincinnati	OH

832	UsMnMISOStPaul832	St. Paul	MN
835	UsNyClarkson835	Potsdam	NY
856	UsFlUsf856	Tampa	FL
857	UsVaFredericksburg857	Fredericksburg	VA
861	UsMDFrederick861	Frederick	MD
862	UsNjAtlanticCity862	Atlantic City	NJ
868	UsTnGallatin868	Gallatin	TN
875	UsMiUm875	Ann Arbor	MI
877	UsMnDodgeCenter877	Dodge Center	MN
936	UsTNChattanooga936	Chattanooga	TN

Appendix B – FNET Software Framework Sample Programs

speedmap2.m

```
clear;
% This file holds the connection parameters for the FNET Application
% server database.
host = 'localhost';
dbname = 'research';
user = 'poweritlab';
password = '';

% Create the connection URL.
conurl = ['jdbc:mysql://' host '/' dbname];

% Connect to the database.
con = database(dbname,user,password,'com.mysql.jdbc.Driver', conurl);

for caseId=1:1:15
    cutoffs = [59.998 59.992 59.990 59.982 59.99 59.985 59.9895 60.024 59.995
59.978 60.002 59.984 59.998 59.984 59.97];
    %cutoffFreq = 59.995;
    % Create an fdrlist object.
    A = fdrlist();
    % Load the extracted data files.
    A.buildListFromFiles(['.\\case' num2str(caseId) '\\'])
    % Trim the data vectors to have the same starting and ending times.
    A.trimData();
    % Check for and fix timestamp duplication.
    A.checkTimestampDuplication(); % This must be run before interpolation
    A.interpolateMissingData();
    A.medianFilter(5);
    A.getArrivalTime('frequency',cutoffs(caseId),'below');

    A.relativeArrivalTime();
    A.makeDistanceMatrix();
    % Do the speed calculation. This goes through the fdrList in order.
    % Then, it goes through the distance matrix in an L-shaped path to
    % determine which FDRs are less than 300 miles away.
    maxSpeed = 0;
%     figure;
%     hold all;
%     for x=1:1:length(A.units)
%         plot(A.units{x}.time,A.units{x}.frequency);
%     end
%     datetick('x','HH:MM:SS.FFF');
    for x=1:1:length(A.units)
        fdrID = A.units{x}.id;
        fprintf('FDR %d:\n',fdrID);
        for i=2:1:length(A.units)+1
           if(A.distMatrix(i,1) == fdrID)
               % Traverse Column
               for a=2:1:i
                   tempDist = A.distMatrix(a,i);
                   if (tempDist <= 300 && tempDist > 0)
```

```
                        for c=1:1:length(A.units)
                           if (A.units{c}.id == A.distMatrix(a,1))
                              deltaT = sqrt((A.units{x}.arrivalTime -
A.units{c}.arrivalTime)^2);
                              if deltaT ~= 0
                                 speed = tempDist/deltaT;
                                 fprintf('(%.2f) ==> ',tempDist);
                                 fprintf('%.2f mi/s\n',speed);
                                 if(speed > maxSpeed)
                                    maxSpeed = speed;
                                 end
                                 A.units{x}.speedArray =
[A.units{x}.speedArray speed];
                              else
                                  fprintf('deltaT is zero!\n');

                              end
                           end
                        end
                    end
                end
                % Traverse Row
                for a=i+1:1:length(A.units)+1
                    tempDist = A.distMatrix(i,a);
                    if (tempDist <= 300 && tempDist > 0)
                        %fprintf('FDR %d is close by.\n',A.distMatrix(1,a));

                        for c=1:1:length(A.units)
                           if (A.units{c}.id == A.distMatrix(1,a))
                              deltaT = sqrt((A.units{x}.arrivalTime -
A.units{c}.arrivalTime)^2);
                              if deltaT ~= 0
                                 speed = tempDist/deltaT;
                                 fprintf('(%.2f) ==> ',tempDist);
                                 fprintf('%.2f mi/s\n',speed);
                                 if(speed > maxSpeed)
                                    maxSpeed = speed;
                                 end
                                 A.units{x}.speedArray =
[A.units{x}.speedArray speed];
                              else
                                  fprintf('deltaT is zero!\n');
                              end
                           end
                        end
                    end
                end
            end
        end

    end
    fprintf('Normalized speeds:\n');
    for i=1:1:length(A.units)
        fprintf('FDR %d (%s): ',A.units{i}.id,char(A.units{i}.name));
        mySpeed = mean(A.units{i}.speedArray/maxSpeed);
        realSpeed = mean(A.units{i}.speedArray);
        if ~isnan(mySpeed)
```

```
%            query = ['INSERT INTO speedmap(caseId,fdrId,speed) VALUES' ...
%                '(' num2str(caseId) ',' num2str(A.units{i}.id) ','
num2str(mySpeed) ')'];
           query = sprintf('INSERT INTO
speedmap(caseId,fdrId,normSpeed,realSpeed) VALUES
(%d,%d,%f,%f)',caseId,A.units{i}.id,mySpeed,realSpeed);
           %fprintf(query);
           e = exec(con,query);
        end
        fprintf('%.2f\n',mean(A.units{i}.speedArray/maxSpeed));
    end
end
close(con);
```

nercfreq.m

```
clear;
% Create an fdrlist object.
A = fdrlist();
% Load the extracted data files.
A.buildListFromFiles('.\\case4-edge\\')
% Trim the data vectors to have the same starting and ending times.
A.trimData();
% Check for and fix timestamp duplication.
A.checkTimestampDuplication(); % This must be run before interpolation
A.interpolateMissingData();
A.medianFilter(5);

figure;
subplot(2,1,1);
plotNames = cell(A.getSize(),1);
hold all;
for i=1:1:length(A.units)
    plot(A.units{i}.time,A.units{i}.frequency);
    plotNames{i} = char(A.units{i}.name);
end
title('Frequency vs. Time');
xlabel('Time (UTC)');
ylabel('Frequency, Hz');
datetick('x','HH:MM:SS.FFF','keepticks');
%legend(plotNames);

dataLength = length(A.units{1}.time);
medianFreq = zeros(dataLength,1);
temp = zeros(A.getSize(),1);

for i=1:1:dataLength
    for j=1:1:A.getSize()
        temp(j) = A.units{j}.frequency(i);
    end
    medianFreq(i) = median(temp);
end
subplot(2,1,2);
plot(A.units{1}.time,medianFreq,'LineWidth',2);
title('Median Frequency vs. Time');
xlabel('Time (UTC)');
ylabel('Frequency, Hz');
datetick('x','HH:MM:SS.FFF','keepticks');
```

virtualfdrexample.m

```
clear;
clc;
host = 'localhost';
myFDR = fdr('Penn',688,34.5,-120,'Knoxville','TN');
%myFDR.loadDataFile('.\\');
myFDR.loadDataFromAccess('2012-02-29','21:01:11',45,'second');
myFDR.port = 9688;
myFDR.numSats = 5;
myFDR.sendData(host,'fnet')
myFDR.exportToFile('.\\export\\','mat');
```

Appendix C – Generators Deactivated due to MATS/CSAPR in the EI

Plant Name	Unit ID	Capacity (MW)	Fuel Type	Bus Number	Year	City	State	In model
Philip Sporn	5	220	Coal	242808	2011	Graham Station	WV	1
Philip Sporn	6	220	Coal	242808	2011	Graham Station	WV	1
Albright Power Station 1	1	69	Coal	235564	2012	Albright	WV	1
Albright Power Station 2	1	69	Coal	235565	2012	Albright	WV	1
Albright Power Station 3	1	140	Coal	235566	2012	Albright	WV	1
Alma	3	15	Coal	681543	2012	Alma	WI	0
Alma	2	15	Coal	681543	2012	Alma	WI	0
Alma	1	15	Coal	681543	2012	Alma	WI	0
Bay Shore	Z	641	Coal	238567	2012	Oregon	OH	1
Elrama Power Plant	1	100	Coal	254014	2012	Elrama	PA	0
Elrama Power Plant	4	185	Coal	254010	2012	Elrama	PA	1
Elrama Power Plant	3	125	Coal	254014	2012	Elrama	PA	0
Elrama Power Plant	2	100	Coal	254014	2012	Elrama	PA	0
Hutsonville	4	75	Coal	347272	2012	Hutsonville	IL	1
Hutsonville	3	75	Coal	347271	2012	Hutsonville	IL	1
Meredosia	4	166	Oil	347680	2012	Meredosia	IL	0
Meredosia	5	203	Coal	347680	2012	Meredosia	IL	0
Monticello	2	593	Coal	508337	2012	Mount Pleasant	TX	0
Monticello	1	593	Coal	508337	2012	Mount Pleasant	TX	0
Niles	2	133	Coal	239008	2012	Niles	OH	0
Niles	1	133	Coal	239008	2012	Niles	OH	0
State Line3	L	180	Coal	274679	2012	Hammond	IN	1
State Line3	H	318	Coal	274679	2012	Hammond	IN	1
State Line4	L	100	Coal	274680	2012	Hammond	IN	1
State Line4	H	197	Coal	274680	2012	Hammond	IN	1
Blue Valley	1	51	Coal	548806	2014	Independence	MO	1
Brayton Point	5	435	Natural Gas	129475	2014	Somerset	MA	0
Buck	6	38	Coal	306022	2014	Salisbury	NC	1
Buck	1	38	Coal	306309	2014	Salisbury	NC	0
Buck	7	38	Coal	306309	2014	Salisbury	NC	0
Chamois	1	49	Coal	300019	2014	Chamois	MO	1
Dale	1	75	Coal	341443	2014	Winchester	KY	1
Dale	1	75	Coal	341440	2014	Winchester	KY	1
Dale	1	27	Coal	341436	2014	Winchester	KY	1
Dale	1	27	Coal	341433	2014	Winchester	KY	1
Endicott Station	4	55	Coal	256228	2014	Litchfield	MI	0
James De Young	1	27	Coal	256002	2014	Holland	MI	0

John Sevier	1	176	Coal	100	2014	Rogersville	TN	0
Johnsonville	1	106	Coal	4142	2014	New Johnsonville	TN	1
Johnsonville	1	141	Coal	4148	2014	New Johnsonville	TN	1
Johnsonville	1	106	Coal	4141	2014	New Johnsonville	TN	1
Johnsonville	1	141	Coal	4147	2014	New Johnsonville	TN	1
Johnsonville	1	106	Coal	4146	2014	New Johnsonville	TN	1
Johnsonville	1	106	Coal	4145	2014	New Johnsonville	TN	1
Johnsonville	1	106	Coal	4144	2014	New Johnsonville	TN	1
Johnsonville	1	141	Coal	4150	2014	New Johnsonville	TN	1
Johnsonville	1	106	Coal	4143	2014	New Johnsonville	TN	1
Johnsonville	1	141	Coal	4149	2014	New Johnsonville	TN	1
Lone Star	1	50	Natural Gas	508297	2014	Lone Star	TX	0
Marion	4	170	Coal	350234	2014	Marion	IL	1
New Castle	2A	138	Coal	242940	2014	West Pittsburg	PA	0
New Castle	5	138	Coal	238812	2014	West Pittsburg	PA	0
Philip Sporn	6	41	Coal	242807	2014	Graham Station	WV	1
Philip Sporn	8	41	Coal	242807	2014	Graham Station	WV	1
Philip Sporn	2	105	Coal	242807	2014	Graham Station	WV	1
Philip Sporn	4	105	Coal	242807	2014	Graham Station	WV	1
Philip Sporn	5	41	Coal	242807	2014	Graham Station	WV	1
Philip Sporn	7	41	Coal	242807	2014	Graham Station	WV	1
Philip Sporn	1	105	Coal	242807	2014	Graham Station	WV	1
Philip Sporn	3	105	Coal	242807	2014	Graham Station	WV	1
Plant Mitchell	C	42	Oil	383783	2014	Albany	GA	1
R E Burger	6	47	Coal	238583	2014	Shadyside	OH	0
R E Burger	5	47	Coal	238583	2014	Shadyside	OH	0
Riverbend	8	94	Coal	306154	2014	Mount Holly	NC	0
Riverbend	7	94	Coal	306040	2014	Mount Holly	NC	1
Riverton	Z	54	Coal	547644	2014	Riverton	KS	1
Riverton	39	38	Coal	547469	2014	Riverton	KS	0
Rivesville 5	1	35	Coal	235575	2014	Rivesville	WV	1
Rivesville 6	1	75	Coal	235576	2014	Rivesville	WV	1
Robert A Reid	1	65	Coal	340572	2014	Robards	KY	1
Salem Harbor	1	82	Coal	221125	2014	Salem	MA	0
Salem Harbor	4	476	Coal	221125	2014	Salem	MA	0
Salem Harbor	3	166	Coal	221125	2014	Salem	MA	0
Salem Harbor	2	82	Coal	221125	2014	Salem	MA	0
Sibley	2	54	Coal	541152	2014	Sibley	MO	1
Sibley	1	54	Coal	541153	2014	Sibley	MO	1

Sunbury Generation LP	2B	40	Coal	200021	2014	Shamokin Dam	PA	0
Sunbury Generation LP	2A	40	Coal	200021	2014	Shamokin Dam	PA	0
Sunbury Generation LP	1	128	Coal	209017	2014	Shamokin Dam	PA	1
Sunbury Generation LP4	3	94	Coal	200021	2014	Shamokin Dam	PA	0
Valley	Z	267	Coal	699506	2014	Milwaukee	WI	1
Wabash River	6	387	Coal	251893	2014	Terre Haute	IN	1
Wabash River	5	125	Coal	251892	2014	Terre Haute	IN	1
Wabash River	4	113	Coal	251890	2014	Terre Haute	IN	1
Wabash River	3	123	Coal	251889	2014	Terre Haute	IN	1
Wabash River	2	113	Coal	251888	2014	Terre Haute	IN	1
Willow Island	1	163	Coal	235578	2014	Willow Island	WV	1
Willow Island	1	50	Coal	235577	2014	Willow Island	WV	1
Armstrong	1	163	Coal	235569	2015	Adrian	PA	1
Armstrong	1	163	Coal	235567	2015	Adrian	PA	1
Ashtubula	5	256	Coal	239036	2015	Ashtabula Township	OH	0
Avon Lake	Z	94	Coal	238554	2015	Avon Lake	OH	1
Avon Lake	Z	640	Coal	238555	2015	Avon Lake	OH	1
B.C. Cobb	4	156	Coal	256108	2015	Muskegon	MI	0
B.C. Cobb	2	156	Coal	256108	2015	Muskegon	MI	0
Black Dog	4	180	Coal	603066	2015	Burnsville	MN	0
Black Dog	3	114	Coal	603066	2015	Burnsville	MN	0
Blount Street	8	49	Coal	699168	2015	Madison	WI	0
Blount Street	9	48	Coal	699168	2015	Madison	WI	0
Canadys Steam	1	105	Coal	370812	2015	Walterboro	SC	1
Cape Fear	1	175	Coal	304881	2015	Moncure	NC	1
Cape Fear	1	148	Coal	304880	2015	Moncure	NC	1
Clifton	1	73	Natural Gas	539655	2015	Clifton	KS	1
Clinch River	3L	104	Coal	242904	2015	Cleveland	VA	1
Clinch River	3H	126	Coal	242903	2015	Cleveland	VA	1
Conesville	3	165	Coal	243654	2015	Conesville	OH	1
D.E. Karn	2	260	Coal	256007	2015	Essexville	MI	0
D.E. Karn	1	255	Coal	256007	2015	Essexville	MI	0
Dubuque	4	30	Coal	630290	2015	Dubuque	IA	1
Dubuque	3	35	Coal	630290	2015	Dubuque	IA	1
Eagle Valley	4	56	Coal	249613	2015	Martinsville	IN	0
Eagle Valley	3	43	Coal	249613	2015	Martinsville	IN	0
Eastlake	Z	1257	Coal	238683	2015	Eastlake	OH	1
Frank E. Ratts1	1	117	Coal	248903	2015	Petersburg	IN	1
Frank E. Ratts2	1	117	Coal	248904	2015	Petersburg	IN	1

Glen Gardner	Z	80	Coal	206333	2015	Glen Gardner	NJ	1
Glen Gardner	Z	80	Coal	206331	2015	Glen Gardner	NJ	1
Glen Lyn	6	108	Coal	242651	2015	Glen Lyn	VA	1
Glen Lyn	5	90	Coal	242650	2015	Glen Lyn	VA	1
Glen Lyn	7	127	Coal	242651	2015	Glen Lyn	VA	1
Green River	4	95	Coal	324022	2015	Central City	KY	1
Green River	3	68	Coal	324021	2015	Central City	KY	1
Harllee Branch	2	319	Coal	383692	2015	Milledgeville	GA	1
Harllee Branch	1	262	Coal	383691	2015	Milledgeville	GA	1
Hutchinson Energy Center	T1	51	Natural Gas	533441	2015	Hutchinson	KS	0
Hutchinson Energy Center	T4	77	Natural Gas	533441	2015	Hutchinson	KS	0
Hutchinson Energy Center	T3	56	Natural Gas	533441	2015	Hutchinson	KS	0
Hutchinson Energy Center	T2	55	Natural Gas	533441	2015	Hutchinson	KS	0
J.R. Whiting (All Units)	Z	328	Coal	256368	2015	Erie	MI	1
Kammer	1L	92	Coal	243193	2015	Captina	WV	1
Kammer	1H	108	Coal	243192	2015	Captina	WV	1
Kammer	2L	92	Coal	243195	2015	Oroville	WV	1
Kammer	2H	108	Coal	243194	2015	Oroville	WV	1
Kammer	3L	92	Coal	243197	2015	Captina	WV	1
Kammer	3H	108	Coal	243196	2015	Captina	WV	1
Kanawha	1H	123	Coal	242895	2015	Glasgow	WV	1
Kanawha	1L	72	Coal	242896	2015	Glasgow	WV	1
Kanawha	2L	123	Coal	242898	2015	Glasgow	WV	1
Kanawha	2H	72	Coal	242897	2015	Glasgow	WV	1
Kraft	1	48	Coal	389008	2015	Port Wentworth	GA	1
Lake Shore	18	256	Coal	238637	2015	Cleveland	OH	0
Lawrence Energy Center	4	110	Coal	532853	2015	Lawrence	KS	0
Lawrence Energy Center	3	48	Coal	532853	2015	Lawrence	KS	0
Meramec	2	138	Coal	345140	2015	St. Louis	MO	1
Meramec	L	170	Coal	345156	2015	St. Louis	MO	1
Meramec	1	138	Coal	345132	2015	St. Louis	MO	1
Meramec	L	140	Coal	345148	2015	St. Louis	MO	1
Meramec	H	190	Coal	345156	2015	St. Louis	MO	1
Meramec	H	140	Coal	345148	2015	St. Louis	MO	1
Miami Fort	6	163	Coal	251949	2015	North Bend	OH	1
Muskingum River	A	70	Coal	243045	2015	Beverly	OH	1
Muskingum River	4	92	Coal	242940	2015	Beverly	OH	1
Muskingum River	3	92	Coal	243045	2015	Beverly	OH	1
Muskingum River	D	113	Coal	242940	2015	Beverly	OH	1

Muskingum River	2	120	Coal	242940	2015	Beverly	OH	1
Muskingum River	C	112	Coal	243045	2015	Beverly	OH	1
Muskingum River	1	120	Coal	243045	2015	Beverly	OH	1
Muskingum River	B	70	Coal	242940	2015	Beverly	OH	1
New Castle	4	96	Coal	238812	2015	West Pittsburg	PA	0
P H Glatfelter	Z	36	Coal	204639	2015	Spring Grove	PA	1
Picway	1	100	Coal	243522	2015	Lockbourne	OH	0
Portland	Z	172	Coal	204661	2015	Mt. Bethel	PA	1
Portland	Z	255	Coal	204651	2015	Mt. Bethel	PA	1
Potomac River	5	110	Coal	314053	2015	Alexandria	VA	0
Potomac River	4	110	Coal	314053	2015	Alexandria	VA	0
Potomac River	3	110	Coal	314053	2015	Alexandria	VA	0
Potomac River	2	88	Coal	314053	2015	Alexandria	VA	0
Potomac River	1	88	Coal	314053	2015	Alexandria	VA	0
Quindaro	T3	46	Natural Gas	530592	2015	Kansas	KS	0
Quindaro	T2	56	Natural Gas	530592	2015	Kansas	KS	0
R Gallagher	1	140	Coal	251857	2015	New Albany	IN	1
R Gallagher	3	140	Coal	251859	2015	New Albany	IN	1
R. Paul Smith	11	75	Coal	235509	2015	Williamsport	MD	0
R. Paul Smith	9	35	Coal	235509	2015	Williamsport	MD	0
Rumford Cogeneration	7	43	Coal	204614	2015	Rumford	ME	0
Rumford Cogeneration	6	43	Coal	204614	2015	Rumford	ME	0
Scholz	2	49	Coal	386752	2015	Sneeds	FL	1
Scholz	1	49	Coal	386751	2015	Sneeds	FL	1
Shawville 1	Z	125	Coal	200715	2015	Shawville	PA	1
Shawville 2	2	125	Coal	200722	2015	Shawville	PA	1
Shawville 3	3	188	Coal	200665	2015	Shawville	PA	1
Shawville 4	4	188	Coal	200666	2015	Shawville	PA	1
Tanners Creek	3	153	Coal	243233	2015	Lawrenceburg	IN	1
Tanners Creek	C	145	Coal	243233	2015	Lawrenceburg	IN	1
Tanners Creek	4	215	Coal	243233	2015	Lawrenceburg	IN	1
Tecumseh Energy Center	Z	74	Coal	532671	2015	Tecumseh	KS	1
Titus	3	75	Coal	204512	2015	Birsboro	PA	0
Titus	2	75	Coal	204512	2015	Birsboro	PA	0
Titus	1	75	Coal	204512	2015	Birsboro	PA	0
Titus	5	18	Coal	204512	2015	Birsboro	PA	0
Titus	4	18	Coal	204512	2015	Birsboro	PA	0
WC Beckjord	4	163	Coal	251936	2015	New Richmond	OH	1
WC Beckjord	3	125	Coal	251935	2015	New Richmond	OH	1

WC Beckjord	2	113	Coal	251934	2015	New Richmond	OH	1
WC Beckjord	1	115	Coal	251939	2015	New Richmond	OH	1
WC Beckjord	6	461	Coal	251938	2015	New Richmond	OH	1
WC Beckjord	5	245	Coal	251937	2015	New Richmond	OH	1
Weatherspoon	A	49	Coal	304924	2015	Lumberton	NC	1
Weatherspoon	A	79	Coal	304927	2015	Lumberton	NC	1
Weatherspoon	A	49	Coal	304925	2015	Lumberton	NC	1
Welsh 2	1	528	Coal	509405	2015	Pittsburg	TX	1
WPS Power Niagara	1	53	Coal	135415	2015	Niagara Falls	NY	0
Yates	1	99	Coal	383641	2015	Newnan	GA	1
Yorktown	2	188	Coal	315091	2015	Yorktown	VA	1
Yorktown	1	188	Coal	315090	2015	Yorktown	VA	1
Cane Run	5	209	Coal	324011	2016	Louisville	KY	1
Cane Run	4	163	Coal	324010	2016	Louisville	KY	1
Cane Run	6	272	Coal	324012	2016	Louisville	KY	1
Chesapeake	J	16	Coal	315101	2016	Chesapeake	VA	1
Chesapeake	G	24	Coal	315100	2016	Chesapeake	VA	1
Chesapeake	B	19	Coal	315099	2016	Chesapeake	VA	1
Chesapeake	D	24	Coal	315099	2016	Chesapeake	VA	1
Chesapeake	H	24	Coal	315100	2016	Chesapeake	VA	1
Chesapeake	F	16	Coal	315099	2016	Chesapeake	VA	1
Chesapeake	I	16	Coal	315101	2016	Chesapeake	VA	1
Chesapeake 1	1	113	Coal	315094	2016	Chesapeake	VA	1
Chesapeake 2	2	113	Coal	315095	2016	Chesapeake	VA	1
Chesapeake 3	3	185	Coal	315096	2016	Chesapeake	VA	1
Chesapeake 4	4	239	Coal	315097	2016	Chesapeake	VA	1
Green River	1	75	Coal	324144	2016	Central City	KY	0
Green River	2	114	Coal	324144	2016	Central City	KY	0
Northeast Station	3	473	Coal	510396	2016	Oologah	OK	0
Tyrone	3	135	Coal	324042	2016	Versailles	KY	1

Appendix D – Planned Gas-Fired Power Plants in the EI

The following is a list of planned gas-fired power plants with winter capacity greater than 100 MW.

Plant Name	Unit ID	Capacity (MW)	Fuel Type	Bus Number	Year	City	State	In Model
Astoria Energy II	T4	156	Natural Gas	126295	2011	Astoria	NY	0
Astoria Energy II	T3	156	Natural Gas	126295	2011	Astoria	NY	0
Astoria Energy II	T2	228	Natural Gas	126295	2011	Astoria	NY	0
Bear Garden	S1	254	Natural Gas	315193	2011	New Canton	VA	1
Bear Garden	G2	170	Natural Gas	315192	2011	New Canton	VA	1
Bear Garden	G1	165	Natural Gas	315191	2011	New Canton	VA	1
Buck	10	163	Natural Gas	306119	2011	Salisbury	NC	0
Buck	11	163	Natural Gas	306119	2011	Salisbury	NC	0
Cane Island	4	160	Natural Gas		2011	Intercession City	FL	0
Fremont Energy Center	2	175	Natural Gas	238602	2011	Fremont	OH	1
Fremont Energy Center	3	325	Natural Gas	238603	2011	Fremont	OH	1
Fremont Energy Center	1	175	Natural Gas	238601	2011	Fremont	OH	1
Gillette SBMC	G3	7	Natural Gas		2011	Boston	MA	0
Greenland Energy Center	1	148	Natural Gas	200581	2011	Jacksonville	FL	0
Greenland Energy Center	2	148	Natural Gas	200581	2011	Jacksonville	FL	0
Hunlock Power Station	5	49	Natural Gas	234251	2011	Hunlock Creek	PA	0
Hunlock Power Station	6	49	Natural Gas	234251	2011	Hunlock Creek	PA	0
Kleen Energy Systems Project	ST	274	Natural Gas		2011	Middletown	CT	0
Kleen Energy Systems Project	U1	177	Natural Gas		2011	Middletown	CT	0
Kleen Energy Systems Project	U2	177	Natural Gas		2011	Middletown	CT	0
Marshfield Utilities Gas Plant	M1	55	Natural Gas	699244	2011	Marshfield	WI	0
Oneida Energy	E1	1	Natural Gas	699359	2011	Green Bay	WI	0
Richmond	A	200	Natural Gas	304978	2011	Hamlet	NC	1
Richmond	B	200	Natural Gas	304978	2011	Hamlet	NC	1
Richmond	C	252	Natural Gas	304978	2011	Hamlet	NC	1
Teche	4	33	Natural Gas	335567	2011	Baldwin	LA	0
West County Energy Center	3A	232	Natural Gas		2011	Loxahatchee	FL	0
West County Energy Center	ST	523	Natural Gas		2011	Loxahatchee	FL	0
West County Energy Center	3C	232	Natural Gas		2011	Loxahatchee	FL	0
West County Energy Center	3B	244	Natural Gas		2011	Loxahatchee	FL	0
York Energy Center	G4	188	Natural Gas	200122	2011	Peach Bottom	PA	0
York Energy Center	G2	122	Natural Gas	200122	2011	Peach Bottom	PA	0
York Energy Center	G3	122	Natural Gas	200122	2011	Peach Bottom	PA	0

York Energy Center	G1	113	Natural Gas	200122	2011	Peach Bottom	PA	0
Bayonne Energy Center	T1	57	Natural Gas	126285	2012	Bayonne	NJ	0
Bayonne Energy Center	T4	57	Natural Gas	126285	2012	Bayonne	NJ	0
Bayonne Energy Center	T2	57	Natural Gas	126285	2012	Bayonne	NJ	0
Bayonne Energy Center	T5	57	Natural Gas	126285	2012	Bayonne	NJ	0
Bayonne Energy Center	T3	57	Natural Gas	126285	2012	Bayonne	NJ	0
Bayonne Energy Center	T6	57	Natural Gas	126285	2012	Bayonne	NJ	0
Bayonne Energy Center	T8	57	Natural Gas	126285	2012	Bayonne	NJ	0
Bayonne Energy Center	T7	57	Natural Gas	126285	2012	Bayonne	NJ	0
Cleveland County Generating Facility	1	180	Natural Gas	306578	2012	Grover	NC	1
Cleveland County Generating Facility	4	180	Natural Gas	306581	2012	Grover	NC	1
Cleveland County Generating Facility	3	180	Natural Gas	306580	2012	Grover	NC	1
Cleveland County Generating Facility	2	180	Natural Gas	306579	2012	Grover	NC	1
Dan River	3	263	Natural Gas	306572	2012	Eden	NC	1
Dan River	2	163	Natural Gas	306571	2012	Eden	NC	1
Dan River	1	163	Natural Gas	306570	2012	Eden	NC	1
Deer Creek Station	1	300	Natural Gas	659285	2012	Elkton	SD	1
Dresden Energy Facility	1S	223	Natural Gas	246770	2012	Dresden	OH	1
Dresden Energy Facility	1B	158	Natural Gas	246770	2012	Dresden	OH	1
Dresden Energy Facility	1A	158	Natural Gas	246770	2012	Dresden	OH	1
Elkins Generating Center	C	20	Natural Gas	506983	2012	Elkins	AR	0
Howard Down	11	56	Natural Gas	228207	2012	Vineland	NJ	0
Jack McDonough	5A	240	Natural Gas	383962	2012	Smyrna	GA	1
Jack McDonough	5	373	Natural Gas	383961	2012	Smyrna	GA	1
Jack McDonough	6B	240	Natural Gas	383885	2012	Smyrna	GA	1
Jack McDonough	6	375	Natural Gas	383883	2012	Smyrna	GA	1
Jack McDonough	4A	240	Natural Gas	383879	2012	Smyrna	GA	1
Jack McDonough	6A	240	Natural Gas	383884	2012	Smyrna	GA	1
Jack McDonough	4	380	Natural Gas	383878	2012	Smyrna	GA	1
Jack McDonough	4B	240	Natural Gas	383880	2012	Smyrna	GA	1
JackMcDonough	5B	240	Natural Gas	383963	2012	Smyrna	GA	1
John Sevier	3	165	Natural Gas	4323	2012	Rogersville	TN	1
John Sevier	2	165	Natural Gas	4322	2012	Rogersville	TN	1
John Sevier	4	383	Natural Gas	4324	2012	Rogersville	TN	1
John Sevier	1	165	Natural Gas	4321	2012	Rogersville	TN	1
New Haven Harbor	3	44	Natural Gas		2012	New Haven	CT	0
New Haven Harbor	2	44	Natural Gas		2012	New Haven	CT	0

New Haven Harbor	4	44	Natural Gas		2012	New Haven	CT	0
Oneida Energy	E3	1	Natural Gas	699359	2012	Green Bay	WI	0
Oneida Energy	E2	1	Natural Gas	699359	2012	Green Bay	WI	0
PSEG Kearny Generating Station	33	44	Natural Gas	217000	2012	Kearny	NJ	0
PSEG Kearny Generating Station	31	44	Natural Gas	217000	2012	Kearny	NJ	0
PSEG Kearny Generating Station	42	44	Natural Gas	217000	2012	Kearny	NJ	0
PSEG Kearny Generating Station	41	44	Natural Gas	217000	2012	Kearny	NJ	0
PSEG Kearny Generating Station	34	44	Natural Gas	217000	2012	Kearny	NJ	0
PSEG Kearny Generating Station	32	44	Natural Gas	217000	2012	Kearny	NJ	0
Warren F Sam Beasley Generation Station	2	50	Natural Gas	232002	2012	Smyrna	DE	0
Waterloo	13	6	Natural Gas	348776	2012	Waterloo	IL	0
Big Bend	T5	56	Natural Gas		2013	Apollo Beach	FL	0
CPV Valley Energy Center	1	281	Natural Gas	148998	2013	Wawayanda	NY	1
CPV Valley Energy Center	1	175	Natural Gas	148997	2013	Wawayanda	NY	1
CPV Valley Energy Center	1	175	Natural Gas	148996	2013	Wawayanda	NY	1
Gowanus Gas Turbines Generating	SS	90	Natural Gas	126277	2013	Brooklyn	NY	0
H L Culbreath Bayside Power Station	8	56	Natural Gas		2013	Tampa Bay	FL	0
H L Culbreath Bayside Power Station	7	56	Natural Gas		2013	Tampa Bay	FL	0
Hamlet Generating Facility	S6	56	Natural Gas	304355	2013	Hamlet	NC	0
Wayne County	A	170	Natural Gas	304960	2013	Goldsboro	NC	1
Wayne County	A	170	Natural Gas	304956	2013	Goldsboro	NC	1
Wayne County	A	170	Natural Gas	304957	2013	Goldsboro	NC	1
Wayne County	A	170	Natural Gas	304959	2013	Goldsboro	NC	1
Wayne County	A	170	Natural Gas	304958	2013	Goldsboro	NC	1
West Deptford Energy Station	1	308	Natural Gas	219121	2013	West Deptford	NJ	0
Big Bend	T6	56	Natural Gas		2014	Apollo Beach	FL	0
Garrison Energy Center LLC	T1	150	Natural Gas	232003	2014	Dover	DE	0
Towantic Energy LLC	G1	165	Natural Gas	126281	2014	Oxford	CT	0
Towantic Energy LLC	G2	161	Natural Gas	126281	2014	Oxford	CT	0
Towantic Energy LLC	G1	161	Natural Gas	126281	2014	Oxford	CT	0
West Deptford Energy Station	2	308	Natural Gas	219121	2014	West Deptford	NJ	0
Zion Energy Center	G4	152	Natural Gas	270940	2014	Zion	IL	0
Zion Energy Center	G5	152	Natural Gas	270940	2014	Zion	IL	0
CPV Warren, LLC	1	180	Natural Gas	235110	2015	Front Royal	VA	0
CPV Warren, LLC	2	180	Natural Gas	235110	2015	Front Royal	VA	0
CPV Warren, LLC	1	105	Natural Gas	235110	2015	Front Royal	VA	0
CPV Warren, LLC	2	105	Natural Gas	235110	2015	Front Royal	VA	0

Cricket Valley Energy	1	346	Natural Gas	126294	2015	Dover	NY	0
Cricket Valley Energy	2	346	Natural Gas	126294	2015	Dover	NY	0
Cricket Valley Energy	3	346	Natural Gas	126294	2015	Dover	NY	0
Garrison Energy Center LLC	T2	150	Natural Gas	232003	2015	Dover	DE	0
Gibson County Generation Station	1	371	Natural Gas	141	2015	Rutherford	TN	0
Lima Energy	T2	240	Natural Gas	242909	2015	Lima	OH	0
Lima Energy	T1	240	Natural Gas	242909	2015	Lima	OH	0
Live Oaks Power Plant	1A	170	Natural Gas	386039	2015	Brunswick	GA	1
Live Oaks Power Plant	1B	170	Natural Gas	386040	2015	Brunswick	GA	1
Live Oaks Power Plant	1	250	Natural Gas	386038	2015	Brunswick	GA	1
Nearman Creek	T5	45	Natural Gas	542976	2015	Kansas City	KS	0
Tampa Electric Co NA 2	1	56	Natural Gas		2015	Tampa Bay	FL	0
Washington Parish Energy Center	T1	215	Natural Gas	336130	2015	Bogalusa	LA	0
Washington Parish Energy Center	G1	172	Natural Gas	336130	2015	Bogalusa	LA	0
Washington Parish Energy Center	G2	172	Natural Gas	336130	2015	Bogalusa	LA	0
Stony Brook	3A	289	Natural Gas	137455	2016	Ludlow	MA	0
Tampa Electric Co NA 2	2	56	Natural Gas		2016	Tampa Bay	FL	0
Trigen Trenton Energy	2	1	Natural Gas	219200	2016	Trenton	NJ	0
Trigen Trenton Energy	1	1	Natural Gas	219200	2016	Trenton	NJ	0
Elk Mound	Z	90	Natural Gas	680516	2017	Elk Mound	WI	1
Tampa Electric Co NA 2	3	56	Natural Gas		2017	Tampa Bay	FL	0
Tampa Electric Co NA 2	4	56	Natural Gas		2018	Tampa Bay	FL	0
Polk	8	366	Natural Gas		2019	Mulberry	FL	0
Arvah B Hopkins	T5	46	Natural Gas	380218	2020	Tallahassee	FL	0

Appendix E – Planned Wind Power Plants in the EI

Plant Name	Unit ID	Capacity (MW)	Fuel Type	Bus Number	Year	City	State	In Model
Bishop Hill Energy LLC	1	200	Wind	636672	2012	Galva	IL	0
Blue Canyon Windpower VI LLC	1	100	Wind	521129	2012	Lawton	OK	1
Cimarron Windpower II	1	131	Wind	531469	2012	Cimarron	KS	0
Crossroads Wind Farm	98	227	Wind	515407	2012	Canton	OK	0
Ironwood Wind	2	167	Wind	531469	2012	Ford County	KS	0
Marble River Wind Farm	G1	200	Wind	137200	2012	Clinton	NY	0
Meadow Lake Wind Farm V LLC	N1	100	Wind	249524	2012	Brookston	IN	0
Post Rock Wind Power Project LLC	1	201	Wind	530592	2012	Ellsworth County	KS	0
Prairie Rose Wind Farm	R1	200	Wind	602039	2012	Jasper	MN	0
Bingham Wind	1	127	Wind		2013	Bingham	ME	0
Black Prairie Wind Farm LLC	N2	200	Wind	270673	2013	McLean County	IL	0
Blackstone Wind Farm IV	G2	100	Wind	270852	2013	Pontiac	IL	0
Lexington Chenoa Wind Farm II LLC	G1	100	Wind	270673	2013	Lexington	IL	0
Lexington Chenoa Wind Farm LLC	G2	200	Wind	270673	2013	Lexington	IL	0
Number Nine Wind Farm	G1	200	Wind		2013	Bridgewater	ME	0
Oakfield Wind Project	2	149	Wind		2013	Oakfield	ME	0
Waverly Wind Farm LLC	G1	200	Wind	532797	2013	Waverly	KS	0
Black Prairie Wind Farm LLC	N1	200	Wind	270673	2014	McLean County	IL	0
Simpson Ridge Wind Farm LLC	N1	100	Wind		2014	Hanna	WY	0

Appendix F – Planned Nuclear Power Plants in the EI

The following is a list of planned nuclear power plants due to open between 2012 and 2017.

Plant Name	Unit ID	Capacity (MW)	Fuel Type	Bus Number	Year	City	State	In Model
Watts Bar, Unit 2	2	1270	Nuclear	4022	2015	Spring City	TN	1
Virgil C. Summer, Unit 2	2	1100	Nuclear	370835	2016	Jenkinsville	SC	1
Vogtle Electric Generating Plant, Unit 3	3	1100	Nuclear	383753	2016	Waynesboro	GA	1
Vogtle Electric Generating Plant, Unit 4	4	1100	Nuclear	380115	2017	Waynesboro	GA	0

Appendix G – Case Index

Case ID	Generators Removed	Synchronous Condensers	Wind	Gas	Nuclear Removed	Add Nuclear	Year	Base Case
1	0	0	0	0	0	0	2012	S2012.sav
2	1	0	0	0	0	0	2012	S2012.sav
3	1	0	0	1	0	0	2012	case-2.sav
7	0	0	0	0	0	0	2015	S15.sav
10	1	0	0	0	0	0	2015	S15.sav
12	1	0	0	1	0	0	2015	case-10.sav
13	1	0	1	1	0	0	2015	case-12.sav
14	0	0	0	0	0	0	2017	S2017.sav
15	1	0	0	0	0	0	2017	S2017.sav
16	1	0	0	0	0	0	2016	S2016.sav
17	0	0	0	0	0	0	2016	S2016.sav
18	1	0	0	1	0	0	2016	S2016.sav
19	1	0	1	1	0	0	2016	case-18.sav
20	1	0	0	1	0	0	2017	case-15.sav
21	1	0	1	1	0	0	2017	case-20.sav
22	1	0	1	1	0	0	2012	case-3.sav
23	0	0	0	0	0	0	2013	S2013.sav
24	1	0	0	0	0	0	2013	S2013.sav
25	1	0	0	1	0	0	2013	case-24.sav
26	1	0	1	1	0	0	2013	case-25.sav
27	0	0	0	0	0	0	2014	S2017.sav
28	1	0	0	0	0	0	2014	S2014.sav
29	1	0	0	1	0	0	2014	case-28.sav
30	1	0	1	1	0	0	2014	case-29.sav

31	1	0	0	1	0	1	2012	case-3.sav
32	1	0	0	1	0	1	2013	case-25.sav
33	1	0	0	1	0	1	2014	case-29.sav
34	1	0	0	1	0	1	2015	case-12.sav
35	1	0	0	1	0	1	2016	case-18.sav
36	1	0	0	1	0	1	2017	case-20.sav
37	1	0	1	1	0	1	2012	case-31.sav
38	1	0	1	1	0	1	2013	case-32.sav
39	1	0	1	1	0	1	2014	case-33.sav
40	1	0	1	1	0	1	2015	case-34.sav
41	1	0	1	1	0	1	2016	case-35.sav
42	1	0	1	1	0	1	2017	case-36.sav
43	1	0	1	1	1	0	2012	case-37.sav
44	1	0	1	1	1	0	2013	case-38.sav
45	1	0	1	1	1	0	2014	case-39.sav
46	1	0	1	1	1	0	2015	case-40.sav
47	1	0	1	1	1	0	2016	case-41.sav
48	1	0	1	1	1	0	2017	case-42.sav
49	1	1	0	0	0	0	2015	case-10.sav
50	1	1	0	0	0	0	2012	case-2.sav
51	1	1	0	0	0	0	2013	case-24.sav
52	1	1	0	0	0	0	2014	case-28.sav
53	1	1	0	0	0	0	2016	case-16.sav
54	1	1	0	0	0	0	2017	case-15.sav
55	1	1	1	1	0	0	2012	case-22.sav
56	1	1	1	1	0	0	2013	case-26.sav
57	1	1	1	1	0	0	2014	case-30.sav
58	1	1	1	1	0	0	2015	case-30.sav
59	1	1	1	1	0	0	2016	case-13.sav
60	1	1	1	1	0	0	2017	case-19.sav
61	1	1	1	1	0	1	2012	case-37.sav
62	1	1	1	1	0	1	2013	case-38.sav
63	1	1	1	1	0	1	2014	case-39.sav
64	1	1	1	1	0	1	2015	case-40.sav
65	1	1	1	1	0	1	2016	case-41.sav
66	1	1	1	1	0	1	2017	case-42.sav

Appendix H – Maps of New Generators by Startup Year

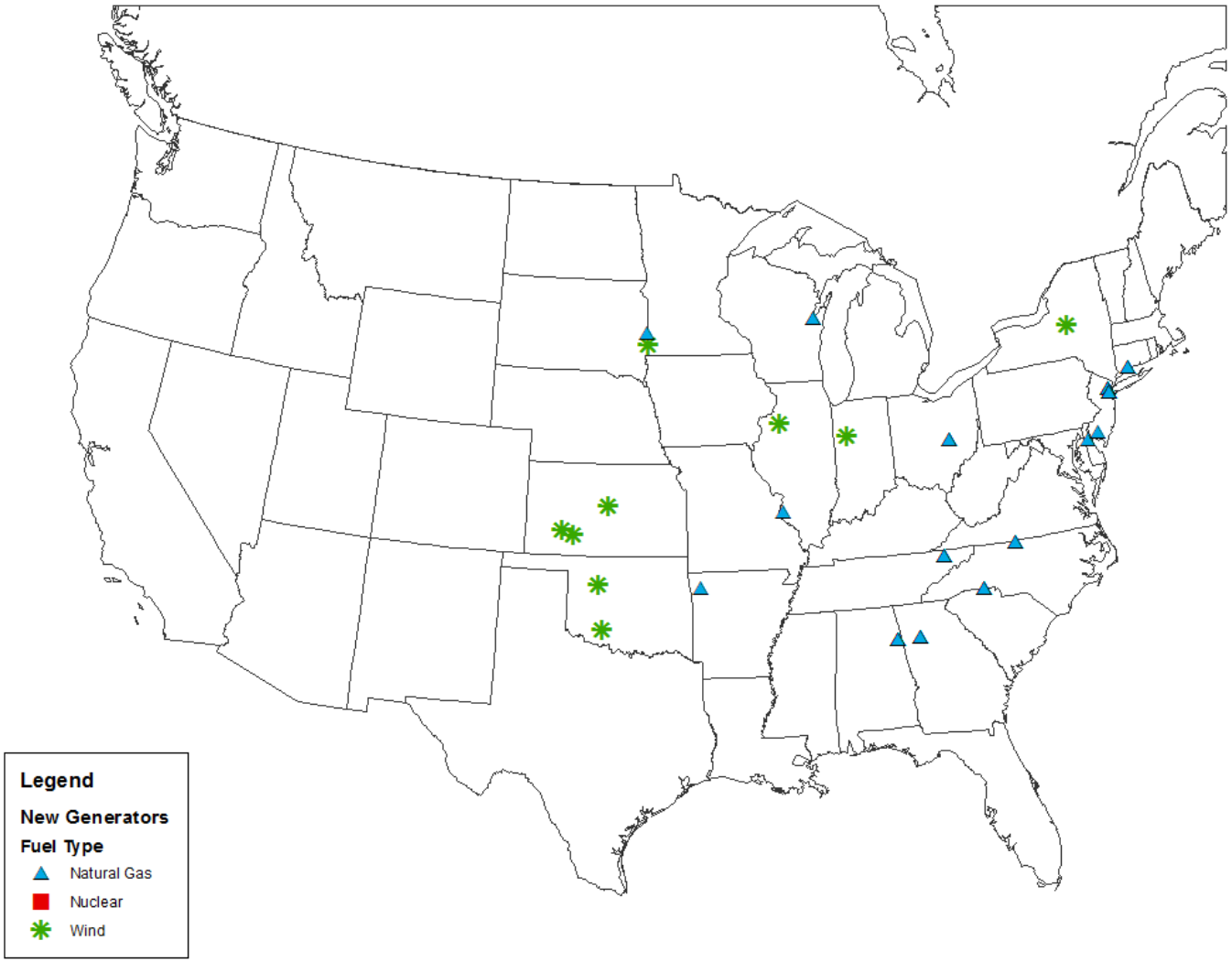


Fig. H.1: Map of new generators in 2012

Fig. H.2: Map of new generators in 2013

Fig. H.3: Map of new generators in 2014

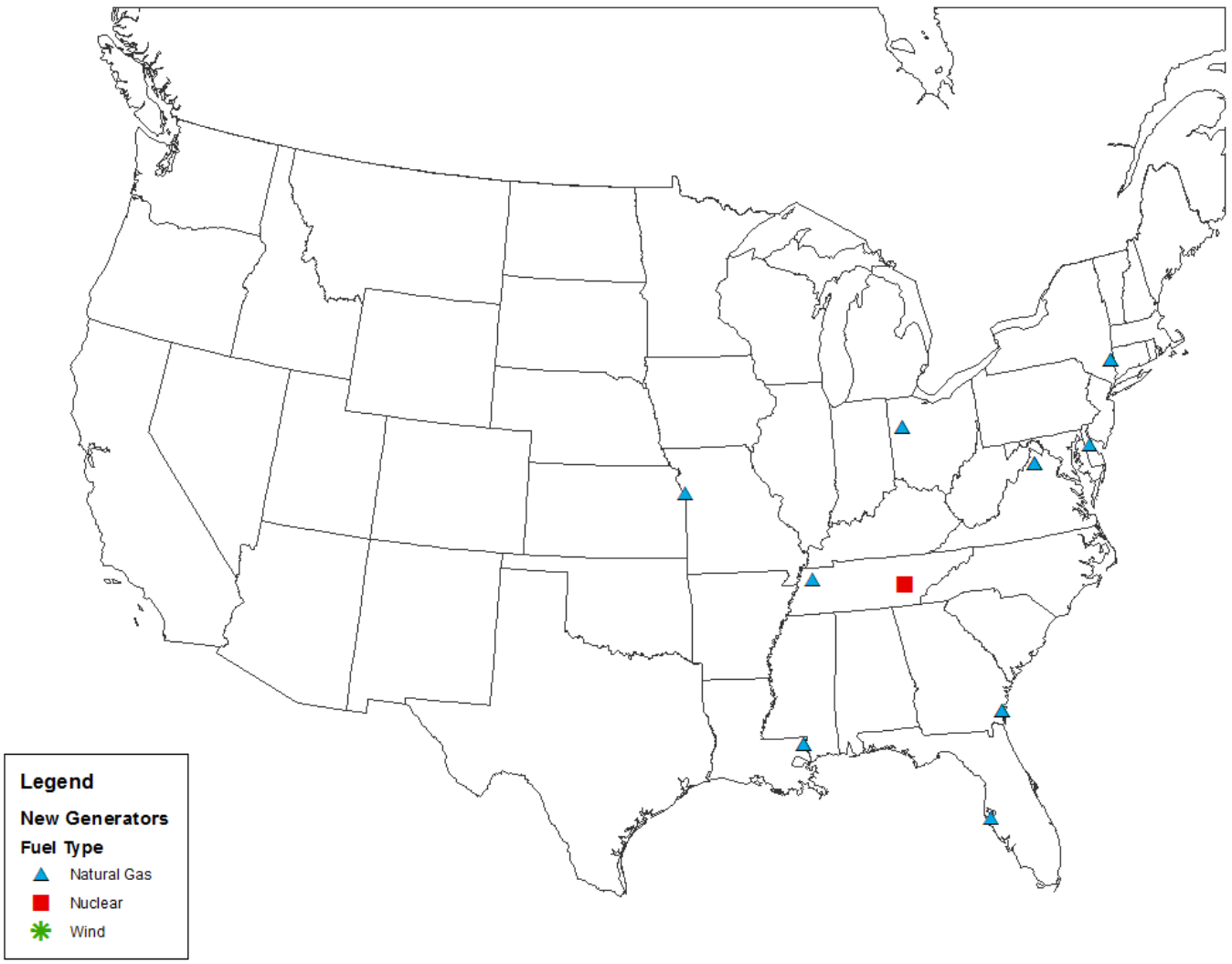


Fig. H.4: Map of new generators in 2015

Fig. H.5: Map of new generators in 2016

Fig. H.6: Map of new generators in 2017

Publications

P. Markham, Y. Zhang, *et al*., "Wide-Area Power System Frequency Measurement Applications," *Future of Instrumentation International Workshop*, Gatlinburg, TN, 2012.

Y. Liu, Z. Yuan, P. Markham, R.W. Conners, and Y. Liu, "Application of Power System Frequency for Digital Audio Authentication," *Power Delivery, IEEE Trans. on*, vol. 27, no. 4, pp. 1820-1828, Oct. 2012.

M. Young, P. Markham, Y. Liu, Y. Ye, L. Chen, Y.C. Zhang, and Y. Liu, "Advanced Power System Monitoring, Modeling, and Control Applications using FNET," *16th IFAC Symposium on System Identification*, 11-13 July 2012.

P. Markham, Y. Zhang, Y. Liu, J. Guo, T. Bilke, and D. Bertagnolli, "Analysis of Frequency Extrema in the Eastern and Western Interconnections, 2010-2011," *Power and Energy Society General Meeting, 2012 IEEE*, 22-26 July 2012.

Y. Liu, P. Markham, and Y. Liu, "Power Grid Frequency Data Conditioning Using Robust Statistics and B-spline Functions," *Power and Energy Society General Meeting, 2012 IEEE*, 22-26 July 2012.

P. Markham, Y. Liu, T. Bilke, and D. Bertagnolli, "Analysis of Frequency Extrema in the Eastern and Western Interconnections," *Power and Energy Society General Meeting, 2011 IEEE*, 24-28 July 2011.

Y. Liu, Z. Yuan, P. Markham, R. Conners, and Y. Liu, "Wide-area frequency as a criterion for digital audio recording authentication," *Power and Energy Society General Meeting, 2011 IEEE*, 24-28 July 2011.

Y. Zhang, L. Chen, Y. Ye, P. Markham et al., "Visualization of Wide Area Measurement Information from the FNET System," *Power and Energy Society General Meeting, 2011 IEEE*, 24-28 July 2011.

L. Chen, P. Markham, C. Chen, and Y. Liu, "Analysis of Societal Event Impacts on the Power System Frequency using FNET Measurements," *Power and Energy Society General Meeting, 2011 IEEE*, 24-28 July 2011.

L. Chen, M. Gandhi, J. Hurtt, P. Markham, Y. Zhang, and Y. Liu, "Island Grid Modeling and Validation Using Single Phase Phasor Measurement Data," *Power and Energy Society General Meeting, 2011 IEEE*, 24-28 July 2011.

Y.C. Zhang, P. Markham et al., "Wide-Area Frequency Monitoring Network (FNET) Architecture and Applications," *Smart Grid, IEEE Transactions on*, vol. 1, no. 2, pp. 159-167, Sept. 2010.

Zhiyong Yuan; T. Xia; Y. Zhang; L. Chen; P.N. Markham; R.M. Gardner; Y. Liu; , "Inter-area oscillation analysis using wide area voltage angle measurements from FNET," *Power and Energy Society General Meeting, 2010 IEEE*, pp.1-7, 25-29 July 2010.

Vita

Penn Markham is a native of Charleston, West Virginia. He received his B.S. in Electrical Engineering from Virginia Polytechnic Institute and State University in 2006. His research interests include applications of wide-area measurements and information technology in power systems. He is a Min Kao Fellow in the Department of Electrical Engineering and Computer Science at the University of Tennessee, and a member of both the IEEE and the Power & Energy Society.