

## Validating the OPA cascading blackout model on a 19402 bus transmission network with both mesh and tree structures

Benjamin A. Carreras  
BACV Solutions  
Mohawk Drive  
Oak Ridge TN USA  
bacarreras@gmail.com

José M. Reynolds-Barredo  
Departamento de Física  
Universidad Carlos III  
Madrid, Spain  
jmr2002@gmail.com

Ian Dobson  
ECpE Department  
Iowa State University  
Ames Iowa USA  
dobson@iastate.edu

David E. Newman  
Physics Department  
University of Alaska  
Fairbanks AK USA  
denewman@alaska.edu

### Abstract

*The OPA model calculates the long-term risk of cascading blackouts by simulating cascading outages and the slow process of network upgrade in response to blackouts. We validate OPA on a detailed 19402 bus network model of the Western Electricity Coordinating Council (WECC) interconnection with publicly available data. To do this, we examine scalings on a series of WECC interconnection models with increasing detail. The most detailed, 19402 bus network has more tree structures at the edges of the main mesh structure, and we extend the OPA model to account for this. The higher-risk cascading outages are the large cascades that extend across interconnections, so validating cascading models on large networks is crucial to understanding how the real grid behaves. Finally, exploring networks with mixed mesh and tree like structure has implications for the risk analysis for both the transmission grid and other network infrastructures.*

### 1. Introduction

Cascading outages are the main way that electric transmission grid blackouts become widespread. While there are many different models and simulations of cascading outages producing plausible sequences of outages [1, 2], a necessary next step is to validate these models. Indeed the IEEE PES cascading failure working group is vigorously pursuing the benchmarking and validation of cascading outage simulations [3,4].

Cascading outages and blackouts occur at all scales, but the larger blackouts that cascade across interconnections, although infrequent, have higher risk than the smaller blackouts because the probability distribution of blackout size has a heavy tail, making large blackouts more likely than might be expected [5]. Therefore it is

important to validate cascading models and simulations for large interconnections.

The OPA model<sup>1</sup> [6–9] simulates cascading blackouts in power grid transmission systems as a complex system in which the reliability is shaped over time by the grid evolution and the engineering responses to blackouts. This paper validates OPA on a series of models of the Western Electricity Coordinating Council (WECC) interconnection in North America with publicly available data, including the observed historical distribution of blackout size. Here we concentrate on the probability distribution of blackout size and other measures of the cascading; the duration and cost components of blackout risk are also challenging and are discussed and approximately quantified in [10].

Our network models of the WECC range from 1553 buses to 19402 buses, allowing the examination of scalings and topology as the model detail increases. Importantly, the smaller network models have mainly a meshed character, whereas the 19402 bus network model has a meshed central portion that has many appendages of tree-like form. Therefore we extend the OPA model to properly account for this change in topological structure in the 19402 bus network model.

Previous work [11] validated the OPA model using data from the WECC on the 1553 bus network. This paper describes the following further advances:

1. Validation on 9402 and 19402 bus network models of the WECC interconnection. This is enabled by rewriting OPA to be much faster [12].
2. Description of network size scaling and change in structure as the network model of the interconnection is made more detailed. While familiar to many engineers, the fact that differently detailed models of the same transmission grid can have a different topological

<sup>1</sup>OPA stands for “Oak Ridge National Lab, Power Systems Engineering Research Center, and University of Alaska”.

structure, or regions with different structure, seems to have been overlooked in the complex network literature.

3. Enhancement of OPA, originally designed for networks with mainly a meshed structure, to handle networks that also have significant tree structures.
4. Enhancement of OPA to better model load variation.

The rest of the paper is organized as follows. Section 2 summarizes the OPA model. The model networks of the WECC system used in this paper are described in Section 3 and summary review of the available WECC data is given in Section 4. Section 5 presents the initial results of the OPA model, showing some problems when applied to the 19402 bus model. Section 6 investigates the reason for the problems and Section 7 presents a new modification of OPA addressing those issues. Finally, discussion and conclusions are given in Section 8.

## 2. OPA Model

The OPA model, which is defined in detail in [6–8], describes the complex dynamical evolution of a power transmission system. There are two timescales: a fast timescale describing the cascading process of the blackouts and a slow time scale describing the evolution of the grid. In the fast timescale, OPA models the transmission lines, loads and generators using the DC load flow approximation. The evolution of the system starts from a good dispatch solution for a given network, then for each day, independent random line outages are triggered with probability  $p_0$ . Whenever a line is outaged, the generation and load are redispatched using standard linear programming methods. The cost function is weighted to ensure that load shedding is avoided where possible. If any lines were overloaded during the optimization, then these lines are outaged with a fixed probability  $p_1$ . This leads to a process of testing for outages and redispatch that is iterated until there are no more line outages. The total load shed is then the total power lost during the blackout. The way the cascading is modeled neglects many of the details of the cascading processes in a real blackout and the timing of events, but it does represent in a simplified way a process of cascading overloads and outages that is consistent with some of its basic features. In the slow timescale, OPA carries out the complex system dynamics of the transmission grid that is evolving in response to a slowly increasing power demand and also whose reliability is shaped via engineering that upgrades the system in response to blackouts. The slow complex system dynamics is carried out first by multiplying each day all loads by a fixed parameter  $g$  that represents the daily rate of increase in electricity demand, which we take to be about 2% a year. To maintain coordination

between generation capacity and transmission capacity, the generation maximum power increases after a period of time  $T$  when the capacity margin decreases below a given critical level  $\Delta P/P$ . A second process in the slow time scale is that when a blackout occurs, the lines involved in the blackout have their line flow limits increased slightly by multiplying by a parameter  $\mu$ . That is, the parts of the system involved in the last blackout are upgraded. For model simplicity, the grid topology remains fixed in the upgrade of the lines. A list of the main OPA parameters is given in Table 1.

**Table 1. Main OPA parameters**

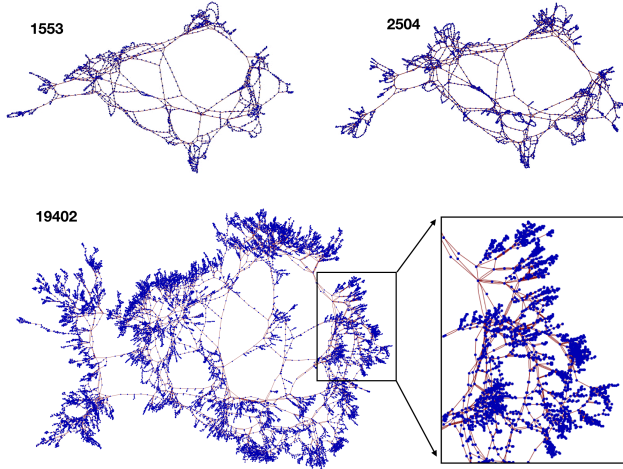
$p_0$	$0.106/N_L$	probability initial line outage
$p_1$	$63.0/N_L$	probability an overloaded line outages
$\gamma$	1.07	width of load variation
$\Delta P/P$	0.2	critical margin
$\mu$	1.07	line upgrade rate
$T$	365	time to build new generators
<hr/>		
$N_L$ is number of lines in network		
$N_L = 2114$ (1553 bus network)		
$N_L = 3345$ (2504 bus network)		
$N_L = 22113$ (19402 bus network)		

On top of the increase of the averaged load, there is also a daily variation on the demand to represent the varying conditions of the power grid. In this paper, the modeling of the variation of the demand has been changed and is discussed in the Appendix.

Given assumptions about blackout cost [5], OPA can calculate the distribution of blackout size over the long-term and hence the blackout risk, accounting for the way that the power grid evolves as blackouts shape this risk. Therefore these calculations can be used to evaluate the long-term impact of transmission system upgrades on the risk of small blackouts and of large blackouts. These calculations for an evolving grid do not address the short term risk of blackout of a fixed power grid.

## 3. WECC network models

We have three model networks of the WECC transmission system with increasing detail; they have 1553, 2504, and 19402 buses respectively, and are shown in Fig. 1. As expected, the overall global structure of the three networks is very similar. However, as the modeling detail increases, the detailed network structure changes. The 19402 bus network has many more tree-type structures attached to the main meshed part of the network. The 1553 bus network has 11% tree-type lines, the 2504 bus network has 19% tree-type



**Figure 1. Three model networks with 1553, 2504, and 19402 buses respectively. The bottom right frame blows up a small portion of an edge of the 19402 bus network to show tree structures.**

lines, while the 19402 bus network has 56% tree-type lines. The ratio of lines to busses for the 1553 and 2504 bus networks is approximately 1.35, whereas the 19404 bus network has the lower (more tree-like) ratio of 1.14. This is also indicated by the network clustering, which reduces by more than a factor of 4 from 0.078 for the 1553 bus network to 0.019 for the 19402 bus network, showing that there are less possible paths between any two buses chosen randomly for the 19402 bus network.

We use the smaller 1553 bus network explore parameters, because the calculations are faster and we can get very good statistical information by performing the calculations for 200 000 simulated days. The objective of this parameter exploration is to find a set of parameters which lead to results that are consistent with the WECC data. To have a consistent set of parameters for the other two networks, we apply the following scaling conditions:

$$\begin{aligned} p_0 N_L &= \text{constant} \\ p_1 N_L &= \text{constant} \end{aligned} \quad (1)$$

where  $N_L$  is the number of lines of the network. All the other parameters are the same for the three networks. We found in the past using artificial networks that this way of scaling the parameters gives consistent results.

#### 4. Historical data for WECC outages

There are a variety of data available (or potentially available) on blackouts and outages of the WECC transmission grid. They are all important for validation of

the modeling of the blackout complex system dynamics. The main source of data on blackouts in the North American grid is the North American Electrical Reliability Council (NERC). This data includes power lost, number of customers with service, duration of the blackout, etc. Analysis of this data [5, 10, 13] shows the existence of power law regions in the probability distribution function and in the rank function of the blackout size as measured by the power lost.

Another source of information we have used is the detailed transmission line outage data for 8864 outages recorded by a WECC utility over a period of ten years [14], which is similar to that collected by the Transmission Availability Data System [15]. The value of this detailed outage data for validation is noted and the authors are very grateful that this data has been made available. Given the data available to us, we must assume that this data, which is only for one WECC utility, is representative of data for the whole WECC. The data has been processed to extract information on cascading events, including the probability distribution of outages in the first generation, the probability distribution of the total number of outages after cascading, and the probability distribution of the number of generations (iterations) in the cascades. Details of the data processing are in [16].

We also extracted from the detailed outage data the average propagation  $\lambda$ . This is the standard Harris estimator for the cascade propagation [16, 17].  $\lambda$  is calculated by taking the average of the ratio of child failures (generation  $i$ ) to parent failures (generation  $i-1$ ) over all the cascading events. This parameter provides important information on how cascades propagate that needs to be matched by the model. Another issue in comparing with data is the definition of a blackout. The thresholds for a reportable blackout include uncontrolled loss of 300 MW or more of firm system load for more than 15 minutes from a single incident, load shedding of 100 MW or more implemented under emergency operational policy, loss of electric service to more than 50 000 customers for one hour or more, and other criteria detailed in the United States Department of Energy form EIA-417. Since we are carrying out the calculations over many decades in order to have good statistics and we have a constant increase in the demand during this period of time, it does not make sense to use a fixed threshold for the power loss. Instead we use a threshold for the power shed normalized to the power demand. This threshold is 0.0001 for a blackout with cascading and for a blackout with no cascading, which implies a localized event, we use a threshold of 0.0004.

## 5. Solution for 1553 and 2504 bus networks and scaling to the 19402 bus network

Exploring the parameters listed in Table 1, one can find multiple solutions for the 1553 bus network that match the WECC data reasonably well. Maintaining the parameters in reasonable ranges we can find a few solutions relatively close to each other. Here, we only discuss one typical solution. The scaling laws in (1) are then used to obtain the parameters  $p_0$  and  $p_1$  for the 2504 bus network. This scaling works well. The parameters obtained for these two networks are listed in Table 1.

As examples of how the numerical results from OPA match the WECC data, first we show in Fig. 2 the rank function for the load shed normalized to the power demand for the 1553 and 2504 bus model networks compared to the data. The agreement is very good for all range of values.

Another important data characteristic to match with the model is the probability distribution of the number of line outages per blackout. The results for the 1553 and 2504 bus model networks are shown in Fig. 3 and the agreement is good. Both sets of results show that the model can reproduce very well the tails of the distributions and so we can expect to reproduce the average cascade propagation in the blackouts. To test that we compare in Fig. 4 the  $\lambda$  function for the two model networks and the data. We limit the comparison to the first 20 iterations because the size of the model networks limits the number of iterations. Thus, using a larger sized network we should be able to extend this comparison to a larger number of iterations. In this limited case, the comparison shown in Fig. 4 is satisfactory. Other data characteristics are also matched by the model, but due to limited space we will present only these three measures in this paper.

The next step in this validation is to scale these results to the 19402 bus network using the appropriate scaling factors (1). The result of the scaling for the 19402 bus network compared with the data for the rank function for the load shed normalized to the power demand and the probability distribution function of the number of line outages per blackout is shown in Fig. 5. The comparison fails fairly badly in this case, and is caused by the mixture of topological structures in the 19402 bus network as discussed in the next section.

## 6. Reducing the tree-like parts of the 19402 bus network

As mentioned earlier, the 19402 bus network has substructures with different network topological charac-

ter. In particular, there are many filamentary and tree-like subnetworks attached to the main meshed structure. This can be seen in Fig. 1, which enlarges a small portion of the boundary of the 19402 bus network.

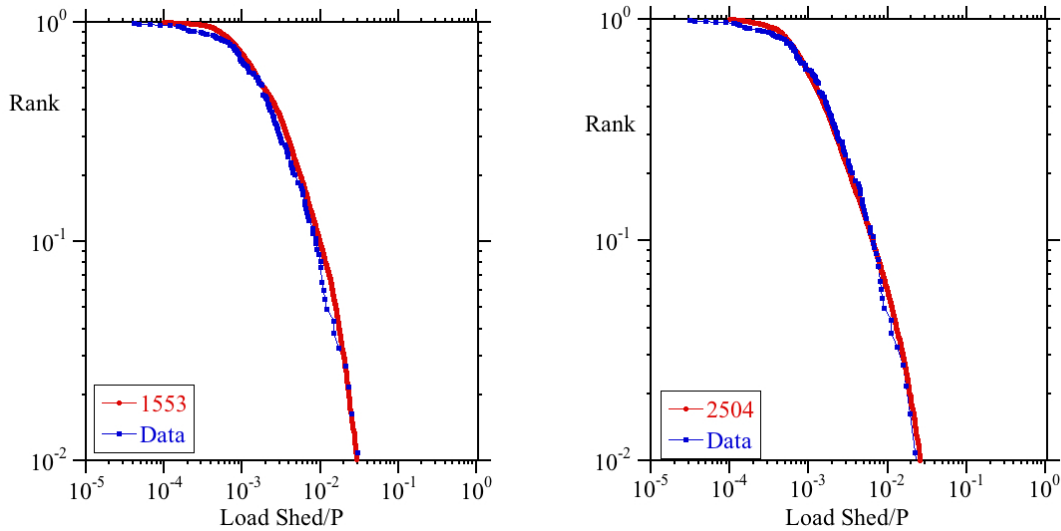
We built three new networks by peeling off or compacting some of these substructures:

1. An 8590 bus network is obtained from the 19402 bus network by substituting for the tree-like structures attached to the core network. Each tree is substituted with a single bus with the total load and generation of the tree.
2. A 7427 bus network is obtained from the 19402 bus network by substituting for the appendix structures attached to the core network. An appendix is defined as a subnetwork connected to the core network by a single bus. Note that an appendix can contain a loop. Each appendix is substituted by a single bus with the total load and generation of the appendix.
3. A 4866 bus network is obtained from the 7427 bus network by substituting chains of buses (that is, consecutive buses with degree 2) with a single bus with the chain total load and generation.

The three networks are shown in Fig. 7, and are not intended to be good models for the WECC; they are built to study the effects of suppressing tree-like parts of 19402 bus network. With these reduced networks we can test the scaling given by (1) and compare the results with the data. Figs. 6 and 8 show the rank function for the load shed normalized to the power demand for the three networks together with the WECC data. (For estimating the propagation  $\lambda$  in the 4866 bus network in Fig. 8, there is more noise for larger  $\lambda$  due to limited data). We clearly do not encounter the problem that we had with the 19402 bus network in Fig. 5. The numerical results describe the data very well. The same happens with the probability distribution function of the number of outages and iterations per blackout and the  $\lambda$  propagation function shown in Figs. 6 and 8. The study in this section of the hypothetical removal of the tree-like substructures identifies the problem for the realistic 19402 bus network.

## 7. Enhanced OPA model

As described in Section 6, the problem with the scaling with size comes from the regions of the 19402 bus network with a tree-like structure. The two parameters that define the probability of outage,  $p_0$  and  $p_1$ , scale following (1) for a fixed topology, but their values can have a different impact on the complex system dynamics of the cascading depending on the topology of the network. Thus to model the electric grid adequately, we must consider different values of the parameters for the regions with different topology.



**Figure 2. Rank function for the load shed normalized to the power demand for the historical data and the 1553 and 2504 bus networks**

We first found that varying the value of the initial failure probability  $p_0$  for the tree-like region did not significantly change the results, so we retain the same value of  $p_0$  throughout the network. As mentioned earlier, the values used for both  $p_0$  and  $\Delta P/P$  are determined, within a range, by the real grid data. The important parameter here is  $p_1$ , the probability that an overloaded line outages. Determining  $p_1$  is more complicated because its real-world counterpart combines a number of factors and is therefore difficult to determine directly from the real grid data. We therefore determine  $p_1$  phenomenologically by adjusting it within a physically reasonable range to make the OPA output fit the real world data. It should be noted that  $p_1$  is not an arbitrary factor, rather a poorly determined physical parameter, and one of the factors which affects  $p_1$  is the topology. We modified  $p_1$  by dividing the value of  $p_1$  for the tree-like lines by 10. Requiring a decrease in  $p_1$  for the tree-like lines makes qualitative sense because of the different upgrading that follows from the different ways that power is redistributed and load is shed in tree-like versus a meshed topologies. In particular, in a tree-like part of the network supplying load, outage of a line disconnects all downstream buses and always causes some blackout and hence in OPA always a line upgrade, while in a meshed network there are several paths to provide load power, and a line outage more rarely causes a blackout. Therefore for a balanced upgrade of all the network, overloads in the tree-like portion must outage less frequently.

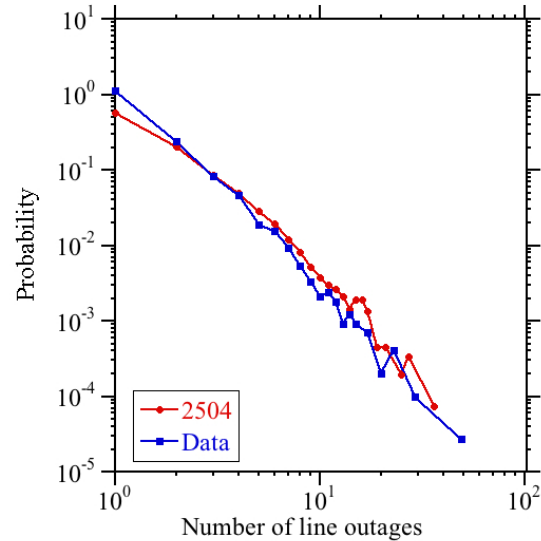
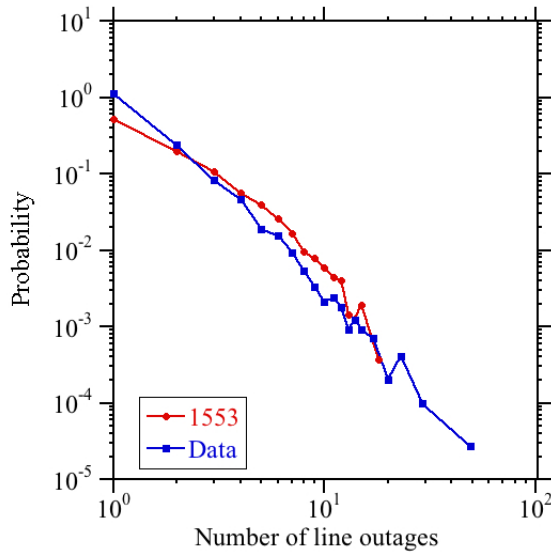
The tenfold decrease in  $p_1$  for the tree-like lines was confirmed empirically by running OPA to show that it is adequate to describe the WECC data. Figs. 9 and 10

show the rank function for the load shed normalized to the power demand and probability distribution function of the number of line outages per blackout for the 19402 bus network. There is very good agreement with data and the mismatch shown in Fig. 5 is no longer present. Obtaining sufficient statistical accuracy for these results on the 19402 bus network required 50 days of computation. The parameters for the computations are in Table 1.

## 8. Conclusion

Proper risk analysis requires a well validated model. For this, it is important to simulate cascading on large network models in order to properly represent the interconnection-scale cascades that are the highest risk cascading blackouts. Accordingly we extend and validate the OPA model on a 19402 bus network model of the WECC.

OPA was previously designed and validated on smaller WECC network models with primarily a meshed structure [11]. However, in this paper we show that the tree-like portions of large transmission network models behave differently than the meshed networks. This highlights a general conclusion that researchers should be mindful of the heterogeneous structure of large power grid networks. We therefore extend OPA to model the tree-like portions of a larger network by reducing the probability of overloaded line outaging in the tree-like portions of the network. This bulk probability can be phenomenologically determined for the different parts of the network and combined with the other OPA parameters obtained from data in [11] to satisfactorily re-



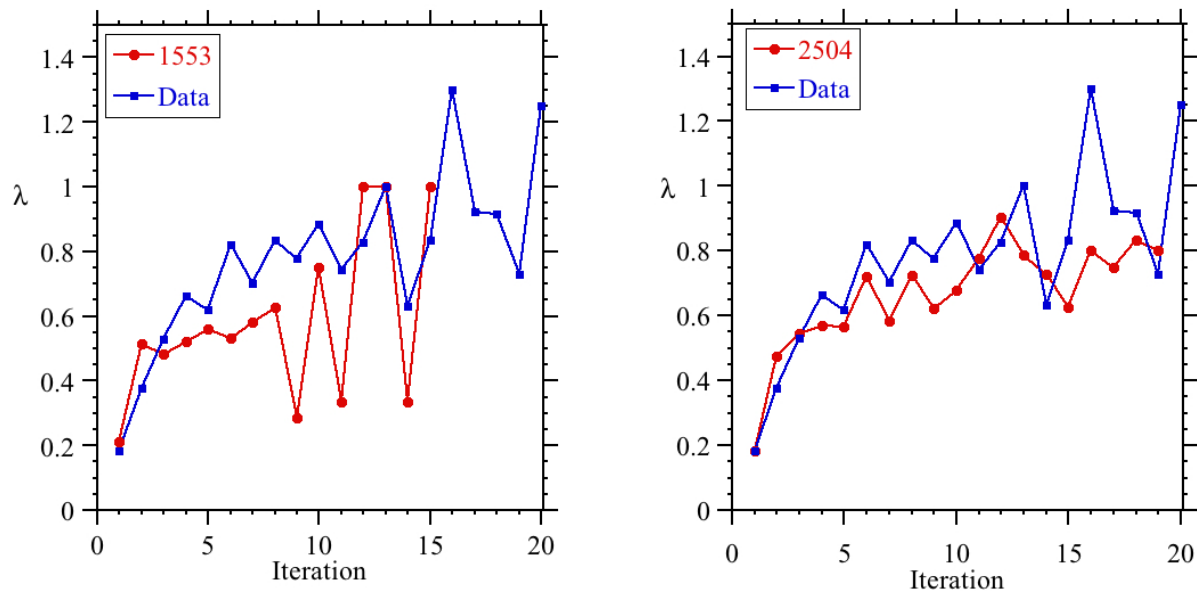
**Figure 3. Probability distribution of the number of line outages per blackout for the historical data and the 1553 and 2504 bus networks**

produce observed bulk statistical features of the WECC obtained from observed data, including the distribution of blackout sizes, the average propagation, and the distribution of the number of line outages. This constitutes validation of the more complete heterogeneous grid. We also improve the modeling of load variation in OPA.

While the OPA model represents cascading processes in a fairly simple manner, it has the unique feature of computing the long-term reliability of the power system as the network slowly upgrades in response to blackouts. The inclusion of the continual engineering efforts to maintain reliability in the modeling may well be responsible for its successful reproduction of observed WECC statistics. The validation of OPA on a large network model increases confidence in this approach to modeling the long-term reliability with respect to cascading blackouts.

## References

- [1] IEEE PES CAMS Task Force on Cascading Failure, Initial review of methods for cascading failure analysis in electric power transmission systems, IEEE PES General Meeting, Pittsburgh PA USA, July 2008.
- [2] —, Survey of tools for risk assessment of cascading outages, IEEE PES General Meeting, Detroit, MI 2011.
- [3] —, Benchmarking and validation of cascading failure analysis tools, IEEE Trans. Power Systems, vol. 31, no. 6, November 2016, pp. 4887-4900.
- [4] E. Ciapessoni et al., Benchmarking quasi-steady state cascading outage analysis methodologies, Probability Methods Applied to Power Systems, Boise ID 2018.
- [5] B.A. Carreras, D.E. Newman, I. Dobson, North American blackout time series statistics and implications for blackout risk, IEEE Trans. Power Systems, vol. 31, no. 6, November 2016, pp. 4406-4414.
- [6] B.A. Carreras, V.E. Lynch, I. Dobson, D.E. Newman, Complex dynamics of blackouts in power transmission systems, Chaos, vol. 14, no. 3, Sept. 2004, pp. 643-652.
- [7] I. Dobson, B.A. Carreras, V.E. Lynch, D.E. Newman, Complex systems analysis of series of blackouts: cascading failure, critical points, and self-organization, Chaos, vol. 17, no. 2, 026103 June 2007.
- [8] S. Mei, X. Zhang, M. Cao, Power Grid Complexity, Tsinghua Univ. Press, Beijing and Springer, Berlin 2011.
- [9] H. Ren, I. Dobson, B.A. Carreras, Long-term effect of the n-1 criterion on cascading line outages in an evolving power transmission grid, IEEE Trans. Power Systems, vol. 23, no. 3, Aug. 2008, pp. 1217-1225.
- [10] D.E. Newman et al., Exploring complex systems aspects of blackout risk and mitigation, IEEE Trans. Reliability, vol. 60, no. 1, March 2011, pp. 134-143.
- [11] B.A. Carreras, D.E. Newman, I. Dobson, N.S. Degala, Validating OPA with WECC data, 46th Hawaii Intl. Conf. System Sciences, Maui, HI, January 2013.
- [12] J. M. Reynolds-Barredo et al., The interplay of network structure and dispatch solutions in power grid cascading failures, Chaos 26, 113111, 2016.
- [13] P. Hines, J. Apt, S. Talukdar, Large blackouts in North America: historical trends and policy implications, Energy Policy, vol. 37, 2009, pp. 5249-5259.
- [14] Bonneville Power Administration Transmission Services Operations & Reliability website <http://transmission.bpa.gov/Business/Operations/Outages>
- [15] North American Electric Reliability Corporation (NERC), Transmission Availability Data System (TADS) Data Reporting Instruction Manual, Aug. 2014.



**Figure 4. Average propagation  $\lambda$  for the historical data and the 1553 and 2504 bus networks**

- [16] I. Dobson, Estimating the propagation and extent of cascading line outages from utility data with a branching process, *IEEE Trans. Power Systems*, vol. 27, no. 4, November 2012, pp. 2146-2155.
- [17] T.E. Harris, *Theory of branching processes*, Dover NY 1989.

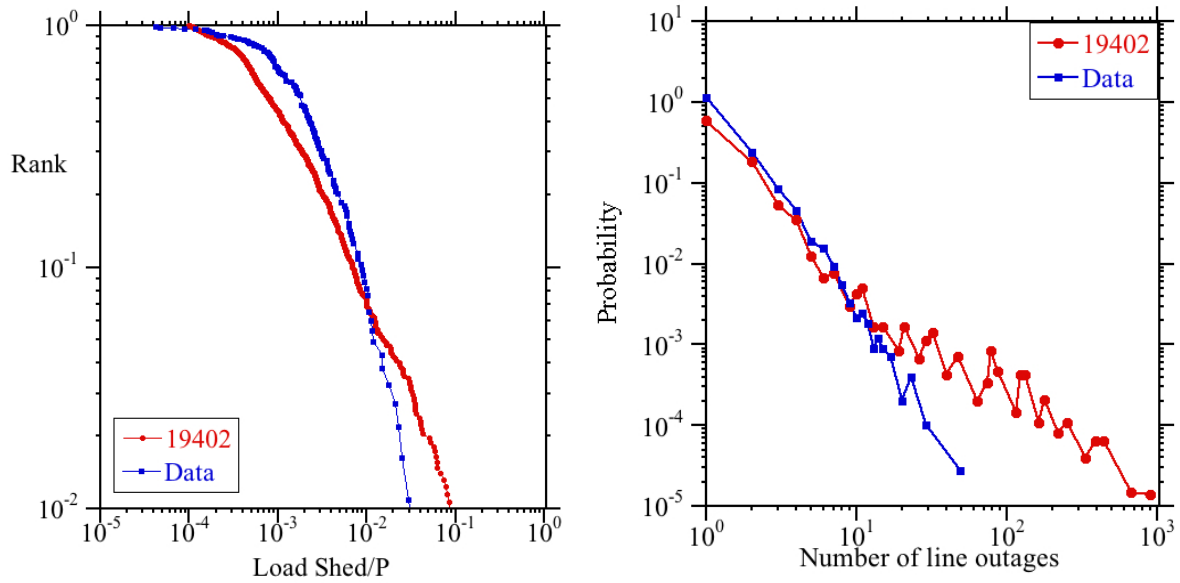
**Acknowledgements:** We gratefully thank BPA for making the historical outage data public. The analysis and any conclusions are strictly the author's and not BPA's. ID acknowledges support in part from NSF grant 1609080. Calculations were performed in Uranus, a supercomputer cluster funded jointly by EU-FEDER and the Spanish Government via the National Projects Nos. UNC313-4E-2361, ENE2009-12213-C03-03, ENE2012-33219, ENE2012-31753 and ENE2015-68265-P.

## Appendix: New model for load variation

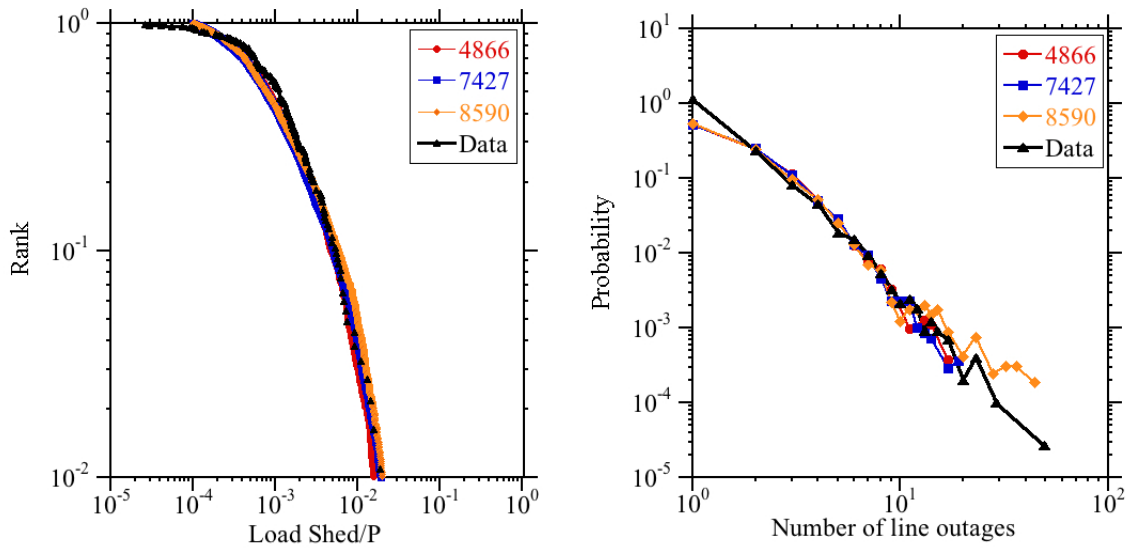
The OPA model previously used a constant distribution of the variation of the demand to determine its variability. This has some problems. If we change the width of the variation of the demand keeping constant the critical margin, we see an abrupt change of the complex system dynamics when the edges of this width are closer to the critical margin, causing an abrupt increase in the frequency and size of blackouts. Another consequence of having a flat distribution for the variation of the demand is that the probability of blackouts is relatively independent of the width of the demand variation and increases sharply when this reaches the critical margin. These results do not seem very reasonable, so we have built a demand variation model which is more consistent

with available data. The New England ISO has made available electricity demand data for every hour of 304 days, which is 7296 hours. These data allow an estimate of the variations of the peak demand. There are daily periodic changes and yearly periodic changes that we subtract from the data, because we do not want the average variation of the demand but rather the statistical fluctuations around these periodic daily and yearly variations. Since the OPA model uses the peak of the daily demand, for each day we can calculate the maximum load. Fig. 11 shows this maximum load as a function of the day during a year. Then, we fit the data in Fig. 11 with  $a + b \sin \left[ \frac{2\pi}{180} x + c \right]$ , where  $x$  is the day in the sequence. The result of this fit is shown in Fig. 11 as a continuous line. Therefore, this fit can be used to describe the yearly evolution of the daily peak value of the load demand. Subtracting this averaged value from the data, we obtain the daily variation of the peak demand. We normalize these values to the averaged value of the demand to get the relative variation of the demand and we can use this daily data to find its distribution. The distribution together with a Gaussian fit is shown in Fig. 11.

From these results, we incorporated in OPA a variation of the load demand based on a Gaussian distribution with a standard deviation of 7% as indicated by Fig. 11. To avoid events with very large demand variations which do not make sense, we limit the maximum variation to three standard deviations from the mean. This new method of modeling the demand variation is driven by observed data and avoids the problems discussed at the start of the Appendix.



**Figure 5. Rank function for the load shed normalized to the power demand and the probability distribution of the number of line outages per blackout for the 19402 bus network compared with historical data.**



**Figure 6. Rank functions of load shed normalized to the power demand, and the number of line outages for the 4866, 7427, and 8590 bus networks.**



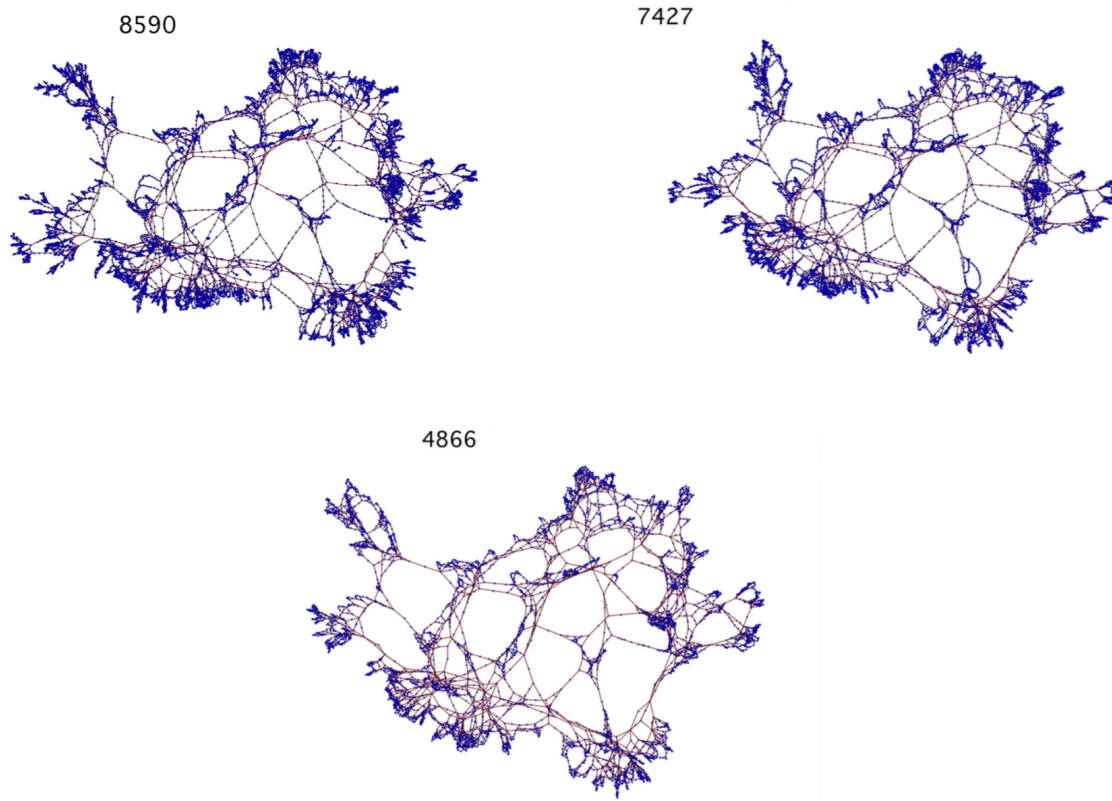


Figure 7. Three model networks of 8590, 7427, and 4866 buses reduced from 19402 bus network.

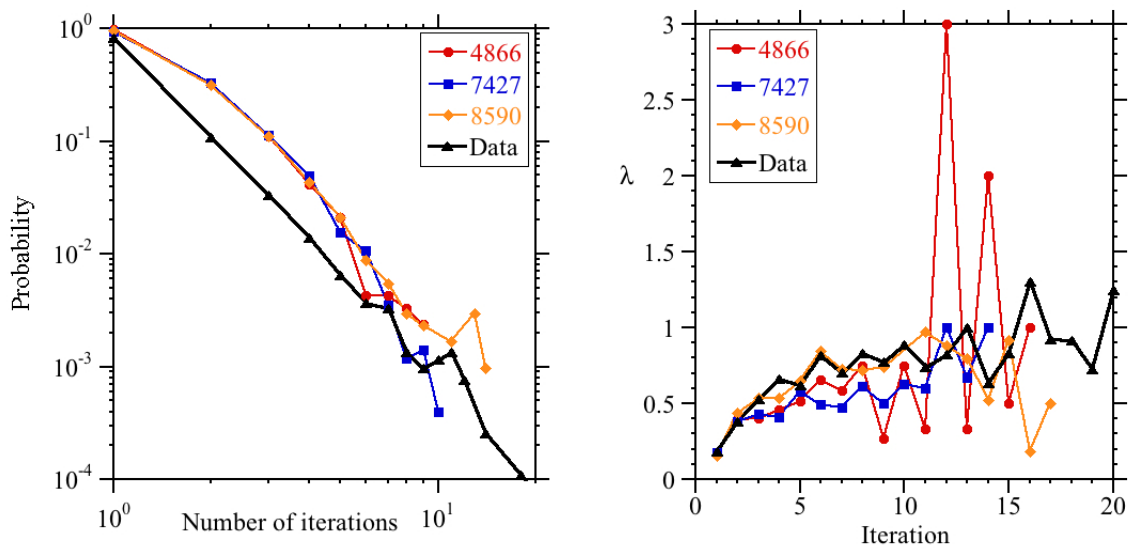
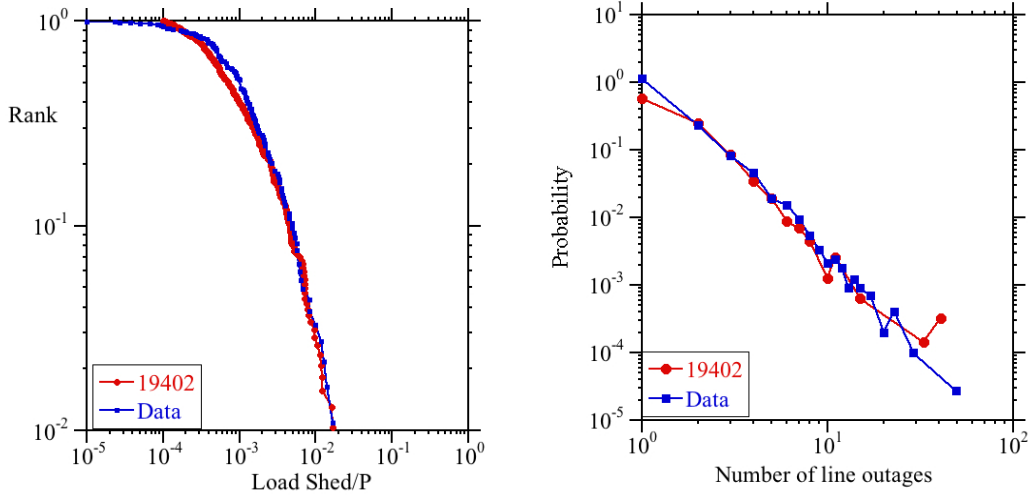
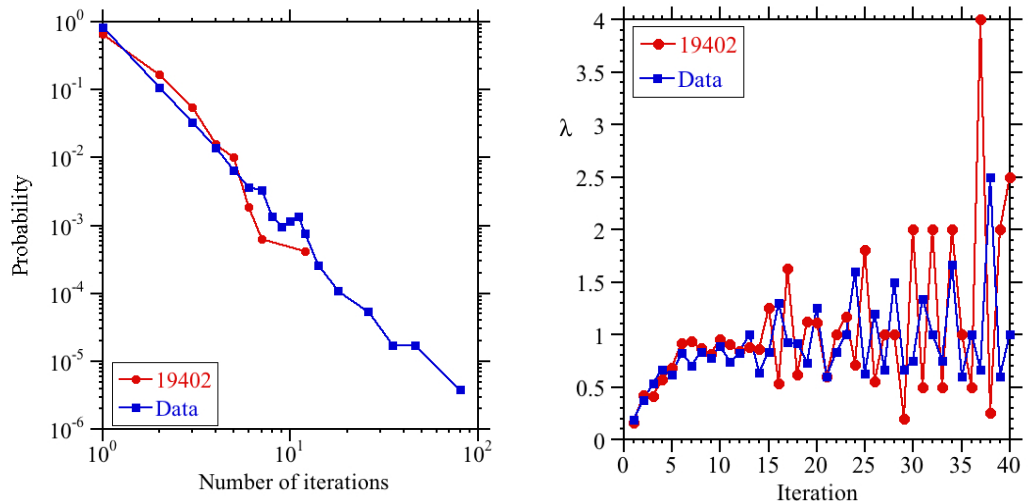


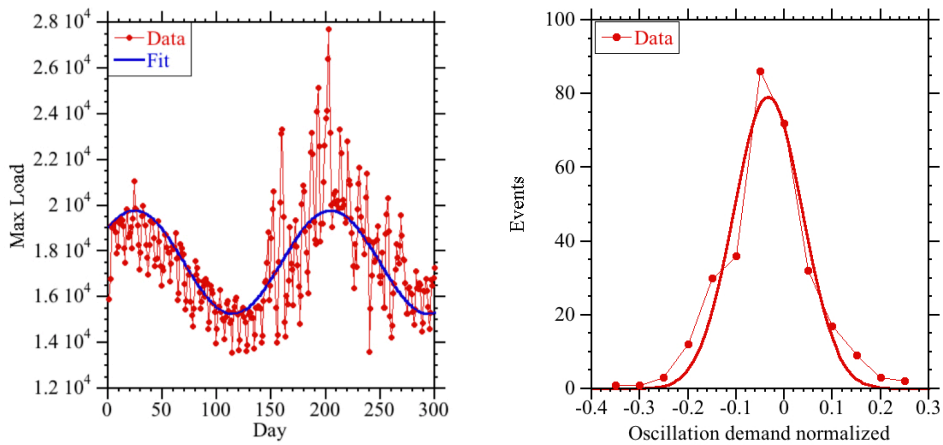
Figure 8. Rank functions of iterations per blackout and average propagation  $\lambda$  for the 4866, 7427, and 8590 bus networks.



**Figure 9. Match of 19402 bus network OPA results with data. Lefthand: Rank function for load shed normalized to power demand. Righthand: PDF of number of line outages per blackout.**



**Figure 10. Match of 19402 bus network OPA results with historical data. Lefthand: Distribution of the number of iterations per cascade. Righthand: Average propagation  $\lambda$  in each iteration.**



**Figure 11. Lefthand: Maximum load as a function of the day during a year and a sinusoidal fit. Righthand: Distribution of the variation of the load around the sinusoidal average.**