Estimating Failure Propagation in Models of Cascading Blackouts

Ian Dobson   Benjamin A. Carreras   Vickie E. Lynch   Bertrand Nkei   David E. Newman

Abstract— **We compare and test statistical estimates of failure propagation in data from versions of a probabilistic model of loading-dependent cascading failure and a power systems blackout model of cascading transmission line overloads. The comparisons suggest mechanisms affecting failure propagation and are an initial step towards monitoring failure propagation from practical system data. Approximations to the probabilistic model describe the forms of probability distributions of cascade sizes.**

I. Introduction

Large blackouts of electric power transmission systems are typically caused by cascading failure of loaded system components. For example, long, intricate cascades of events caused the Western North American blackout of 30,390 MW in August 1996 [18] and the Eastern North America blackout of 61,800 MW in August 2003 [21]. The vital importance of the electrical infrastructure to society motivates the analysis and monitoring of the risks of cascading failure. In particular, in addition to limiting the start of outages that cascade, it is useful to be able to monitor the tendency of cascading failures to propagate after they are started [14], [5].

CASCADE is a probabilistic model of loading-dependent cascading failure that is simple enough to be analytically tractable [13], [16], [15]. CASCADE contains no power system modeling, but does seem to approximately capture some of the salient features of cascading failure in large blackouts. The CASCADE model has many identical components randomly loaded. An initial disturbance adds load to each component and causes some components to fail by exceeding their loading limit. Failure of a component causes a fixed load increase for other components. As components fail, the system becomes more loaded and cascading failure of further components becomes likely.

The CASCADE model can be well approximated by a Galton-Watson branching process in which failures occur in

I. Dobson is with the ECE Dept., University of Wisconsin, Madison WI 53706 USA (email: dobson@engr.wisc.edu). B.A. Carreras, V. Lynch, and B. Nkei are with Oak Ridge National Laboratory, Oak Ridge TN 37831 USA (email: carrerasba@ornl.gov). D.E. Newman is with the Physics Dept., University of Alaska, Fairbanks AK 99775 USA (email: ffden@uaf.edu).

stages and each failure in each stage causes further failures in the next stage according to a Poisson distribution [14]. The average number of failures in the initial disturbance is θ and the subsequent stochastic propagation of the failures is controlled by the parameter λ, which is the average number of failures caused by each failure in the previous stage.

OPA is a power system blackout model that represents probabilistic cascading line outages and overloads [3]. The network is conventionally modeled using DC load flow and LP dispatch of the generation. The initial disturbance is generated by random line outages and load variations. Overloaded lines outage with a given probability and the subsequent power flow redistribution and generator redispatch can overload further lines, which can then probabilistically outage in a cascading fashion. There is no attempt to represent all the diverse interactions that can occur during a blackout. However, the modeling does represent a feasible cascading blackout consistent with some basic network and operational constraints. OPA can also model the slow evolution of the network as load grows and the network is upgraded in response to blackouts [1], [12], [2], [4], but in this paper the network is assumed to be fixed and these complex systems dynamics are neglected.

Other authors have constructed power system blackout models involving cascading failure emphasizing different aspects of the problem. Chen and Thorp [6], [7] model hidden failures and compute vulnerability of key lines using importance sampling and examine criticality and blackout mitigation. Ni, McCalley, Vittal, and Tayyib [19] show how to monitor the risk of a variety of system limits being exceeded; minimizing this risk would have the effect of limiting the risk of cascading events starting. Chen, Zhu, and McCalley [8] show how to evaluate the risk of the first few likely cascading failures. Rios, Kirschen, Jawayeera, Nedic, and Allan [20] use Monte Carlo simulation to estimate the cost of security taking account of hidden failures, cascading outages and transient instability.

Our ultimate goal is to understand cascading failure in large blackouts from a global systems point of view, identify the main parameters governing the cascading process, and suggest ways to estimate these parameters from real or simulated outage data. These metrics will allow monitoring of the risk of cascading failure and quantifying of the tradeoffs involved in blackout mitigation. In this paper, we take a step towards this goal by comparing the abstract cascading failure model CASCADE with the power system blackout model OPA. The comparison reveals which features of the OPA blackouts are captured by the CASCADE model. In particular, we seek to characterize in OPA and measure from OPA results the parameter λ governing the propagation of failures after the

start of the cascade. Resolving problems in measuring λ from OPA results is a first step towards measuring the degree to which failures propagate in power systems. If the overall system stress is such that failures propagate minimally, then any failures that occur are likely to be a single failure or a short sequences of failures that cause small blackouts or no blackout. However, if the overall system stress is such that failures propagate readily, then there is a substantial risk of cascading failure leading to large blackouts and it is in the national interest to quantify this risk and examine the economics and engineering of mitigating this risk.

II. CASCADE MODEL AND BRANCHING PROCESS PARAMETERS

This section summarizes the CASCADE model of probabilistic load-dependent cascading failure and its branching process approximation [13], [16], [15], [14]. (Here the normalized version of CASCADE is summarized; for many purposes, the unnormalized version is more useful and flexible [13], [16].)

The CASCADE model has n identical components with random initial loads. For each component the minimum initial load is 0 and the maximum initial load is 1. For $j=1,2,...,n$, component j has initial load ℓ_j that is a random variable uniformly distributed in $[0,1]$. $\ell_1, \ell_2, \cdots, \ell_n$ are independent.

Components fail when their load exceeds 1. When a component fails, a fixed amount of load $p \geq 0$ is transferred to each of k components. The k components to which load is transferred are chosen randomly each time a component fails [15].

To start the cascade, we assume an initial disturbance that loads each component by an additional amount d. Other components may then fail depending on their initial loads ℓ_j and the failure of any of these components will distribute the additional load p that can cause further failures in a cascade. The cascade proceeds in stages with M_1 failures due to the initial disturbance, M_2 failures due to load increments from the M_1 failures, M_3 failures due to load increments from the M_2 failures, and so on. The size of the cascading failure is measured by the total number of components failed S.

For the case $k = n$ in which load is transferred to all the system components when each failure occurs, the distribution of S is a saturating quasibinomial distribution [16], [13], [9]:

$$
P[S=r] = \begin{cases} \binom{n}{r} \phi(d)(d+rp)^{r-1}(\phi(1-d-rp))^{n-r}, \\ \qquad\qquad\qquad\qquad r = 0,1,...,n-1 \\ 1 - \sum_{s=0}^{n-1} P[S=s], \qquad r = n, \end{cases} \quad (1)
$$

where the saturation function ϕ is

$$
\phi(x) = \begin{cases} 0 & ; x < 0, \\ x & ; 0 \leq x \leq 1, \\ 1 & ; x > 1. \end{cases} \quad (2)
$$

Note that (1) uses $0^0 \equiv 1$ and $0/0 \equiv 1$ when needed.

In the case $k < n$, no analytic formula such as (1) is currently available, but it can be shown that the following approximation (4) remains valid [15].

Define

$$
\lambda = kp \quad \text{and} \quad \theta = nd \quad (3)
$$

λ may be interpreted as the total amount of load increment associated with any failure and is a measure of how much the components interact. θ may be interpreted as the average number of failures due to the initial disturbance.

Now we approximate the CASCADE model [14], [15]. Let $n \to \infty, k \to \infty$ and $p \to 0, d \to 0$ in such a way that $\lambda = kp$ and $\theta = nd$ are fixed. For $\theta \geq 0$,

$$
P[S=r] \approx
$$
$$
\begin{cases} \theta(r\lambda+\theta)^{r-1}\dfrac{e^{-r\lambda-\theta}}{r!} & ; 0 \leq r \leq (n-\theta)/\lambda, \ r < n \\ 0 & ; (n-\theta)/\lambda < r < n, \ r \geq 0 \\ 1 - \displaystyle\sum_{s=0}^{n-1} P[S=s] & ; r = n \end{cases} \quad (4)
$$

The approximate distribution (4) is a saturating form of the generalized Poisson distribution [11], [10]. Moreover, under the same approximation, the stages of the CASCADE model become stages of a Galton-Watson branching process [14], [17]. In particular, the initial failures are produced by a Poisson distribution with parameter θ. Each initial failure independently produces more failures according to a Poisson distribution with parameter λ, and each of those failures independently produces more failures according to a Poisson distribution with parameter λ, and so on. This branching process leads to another interpretation of λ as the average number of failures per failure in the previous stage. λ is a measure of the average propagation of the failures [14].

The expected number of failures in stage j of the branching process is given by

$$
EM_j = \theta\lambda^{j-1} \quad (5)
$$

until saturation due to the system size occurs. Formula (5) is exact for the branching process before saturation and an approximation for the expected number of failures in each stage of CASCADE.

Further approximation is useful. Using Stirling's formula and a limiting expression for an exponential for $r \gg 1$, (4) becomes

$$
P[S=r] \approx \frac{\theta}{\lambda\sqrt{2\pi}} \exp[(1-\lambda)\tfrac{\theta}{\lambda}]\frac{\exp[-r(\lambda-1-\ln\lambda)]}{(r+\frac{\theta}{\lambda})\sqrt{r}}
$$
$$
; 1 \ll r < r_1 = \min\{n/\lambda, n\} \quad (6)
$$

and if $\theta/\lambda \sim 1$ so that also $r \gg \theta/\lambda$,

$$
P[S=r] \approx \frac{\theta\exp[(1-\lambda)\tfrac{\theta}{\lambda}]}{\lambda\sqrt{2\pi}} r^{-\frac{3}{2}} \exp[-r(\lambda-1-\ln\lambda)]
$$
$$
; 1 \ll r < r_1 = \min\{n/\lambda, n\} \quad (7)
$$

Let

$$
r_0 = (\lambda - 1 - \ln\lambda)^{-1} \quad (8)
$$

In the approximation (7), the term $r^{-\frac{3}{2}}$ dominates for $r \lesssim r_0$ and the exponential term dominates for $r \gtrsim r_0$. Thus (7) reveals that the distribution of the number of failures has an approximate power law region of exponent -1.5 for $1 \ll r \lesssim r_0$ and an exponential tail for $r_0 \lesssim r < r_1$. The approximation

(7) implies that r_0 is only a function of λ and does not depend on θ or the system size n.

We discuss some of the implications of the approximation for the form of the distribution of S.

1) For $\lambda = 1$, (6) becomes $P[S = r] \approx \frac{\theta}{\sqrt{2\pi}} r^{-\frac{3}{2}}$, r_0 becomes infinite and the power law region extends up to the system size.

2) $r_0^{-1} = \lambda - 1 - \ln\lambda$ is a nonnegative function of λ with a quadratic minimum of zero at $\lambda = 1$. Therefore, for a range of λ near 1, $r_0 = (\lambda - 1 - \ln\lambda)^{-1}$ is large and the power law region extends to large r.

3) The risk of large cascades as compared with the risk of small cascades is approximately determined by λ. Small λ gives a distribution with an exponential tail past a small number of failures and a negligible probability of large cascades. λ near 1 gives a power tail with a significant probability of large cascades and $\lambda > 1$ gives a significant probability of all components failing.

4) If $\lambda < 1$, then $P[S = r]$ increases with θ. The increase is linear for small θ and exponential with large θ. Thus reducing the probability of failures by decreasing the size of the initial disturbance is most effective when λ is less than and bounded away from 1 and θ is large.

III. OPA BLACKOUT MODEL

In the OPA model [3], there is a fast time scale of the order of minutes to hours, over which cascading overloads or outages may lead to a blackout. Cascading blackouts are modeled by overloads and outages of lines determined in the context of LP dispatch of a DC load flow model. To start the cascade, random line outages are triggered with a probability p_0. A cascading overload may also start if one or more lines are overloaded in the solution of the LP optimization. In this situation, we assume that there is a probability p_1 that an overloaded line will outage. When a solution is found, the overloaded lines of the solution are tested for possible outages. Outaged lines are in effect removed from the network and then a new solution is calculated. This process can lead to multiple iterations, and the process continues until a solution with no more line outages is found. We regard each iteration as one stage of the cascading blackout process. The overall effect of the process is to generate a possible cascade of line outages that is consistent with the network constraints and the LP dispatch optimization.

The parameters p_0 and p_1 determine the initial disturbance. The level of stress on the system is determined by a multiplier on the loads in the power system.

IV. ESTIMATING θ AND λ FROM DATA

This section proposes methods of estimating θ and λ from the data produced by CASCADE or OPA. Both CASCADE and OPA produce a stochastic sequence of failures in stages with M_1 failures due to the initial disturbance and subsequent numbers of failures $M_2, M_3,...$. In the case of OPA the failures are transmission line outages. If at any stage (including the first stage), there are zero failures, then the cascade of failures ends.

For $d > 0$, the probability of a nontrivial cascade in the CASCADE model is easily obtained from (1) as

$$P[S > 0] = 1 - P[S = 0]$$
$$= 1 - (1 - d)^n = 1 - (1 - \theta/n)^n \qquad (9)$$

Let the observed frequency of nontrivial cascades be

$$f = \frac{\text{\# cascades with } S > 0}{\text{\# samples}} \qquad (10)$$

Then (9) suggests the following estimator for θ:

$$\hat{\theta} = n - n(1 - f)^{\frac{1}{n}} \qquad (11)$$

Let the sample mean of the number of failures in stage j of the cascade be

$$m_j = \frac{1}{\text{\# samples}} \sum_{\text{samples}} \text{\# failures in stage } j \qquad (12)$$

Then (5) suggests the following estimator for λ based on the data from stage j of the cascade:

$$\hat{\lambda}_j = (m_j/\hat{\theta})^{\frac{1}{j-1}} \qquad (13)$$

The naive estimators in (11) and (13) have been tested on data produced by CASCADE and they appear empirically to be useful statistics. For example, for $\lambda < 1.3$, Fig. 1 shows the estimated $\hat{\lambda}_j$ as constant with respect to the stage j as expected. (For $\lambda > 1.3$, the estimated $\hat{\lambda}_j$ decreases with the stage j because at higher λ and at higher j there are more cascades with all 190 components failed and this saturation effect reduces $\hat{\lambda}_j$. Recall that the formula (5) used to derive (13) assumes no saturation.)

Fig. 1. $\hat{\lambda}_j$ as a function of stage number j from CASCADE model for $\lambda = 0.0017, 0.0114, 0.038, 0.106, 0.408, 0.7125, 1.13, 1.467,$ and 1.56. There are $n = 190$ components and $\theta = 0.095$.

V. RESULTS

A. Comparing CASCADE and OPA

The OPA model on a 190 node tree-like network [2] was used to produce line outage data. The load multiplier parameter was varied to vary the system stress. The $\hat{\lambda}_j$ computed from the OPA results is plotted in Fig. 2. We can see that at high load $\hat{\lambda}_j$ is a decreasing function of the stage j while for low loads $\hat{\lambda}_j$ is an increasing function of the stage j. This functional form is not seen in the CASCADE model results in Fig. 1.

The probability distributions for number of lines outaged in OPA corresponding to Fig. 2 are shown in Fig. 3. We can attempt to match these probability distributions with CASCADE by using $\hat{\theta}$ from the OPA results as an estimate of θ and using $\hat{\lambda}_1$ from the OPA results as an estimate of λ. The resulting CASCADE probability distributions are shown in Fig. 4. Although there is reasonable qualitative agreement between the probability distributions from OPA and CASCADE for smaller λ, the OPA probability distributions for larger λ contain a peak not present in the CASCADE probability distributions. We consider a modification to CASCADE to explain this peak in subsection V-B.

Fig. 3. Probability distributions of number of line outages from OPA model for various values of loading multiplier.

Fig. 2. $\hat{\lambda}_j$ as a function of stage number j from OPA model for various values of loading multiplier.

B. Blackout inhibition modification to CASCADE

In a blackout, there is not only an effect by which line outages further load the system and tend to cause further outages. There is also an effect by which sufficient line outages will cause load to be shed and this load shedding reduces the load on the system. (It is also possible, but perhaps less common, for load shedding to introduce large disturbances and imbalances that further stress portions of the system.) Moreover sufficient line outages will tend to island the system

and this can have the effect of limiting further outages. That is, sufficiently many line outages can have an inhibitory effect on further cascading outages.

We attribute the peak in the OPA probability distributions for larger λ to this inhibitory effect. One can argue that for small λ, it is not likely that the cascade will include enough line outages to encounter the inhibitory effect. Moreover, the inhibitory effect could result in the decrease in $\hat{\lambda}_j$ as the stage j increases observed for larger λ in Fig. 2.

CASCADE does not model the inhibitory effect and one way to test these explanations is to modify CASCADE to model the inhibitory effect. A crude modeling of the inhibitory effect in CASCADE is to halt the cascading process after a fixed number of components $r_{\max}$ have failed. That is, when $r_{\max}$ components have failed, the current stage of the cascade is completed, thus allowing more than $r_{\max}$ components to fail, but the next stage of the cascade is suppressed.

The results of the modified CASCADE model with $r_{\max} = 10$ are shown in Figs. 5 and 6. The decrease in $\hat{\lambda}_j$ with j for larger λ is evident in Fig. 5 and the peak in the probability distribution for larger λ is evident in Fig. 6. These qualitative dependencies in the modified CASCADE results are similar to the OPA results in Figs. 3 and 4. However, Fig. 5 does not show the increase in $\hat{\lambda}_j$ with j for smaller λ observed in Fig. 2 and a further modification to CASCADE to examine this is considered in subsection V-C.

We comment further on the modified CASCADE results in Fig. 5. The value $\hat{\lambda}_1$ in the first stage agrees with the input λ. That is, the inhibition does not seem to affect the initial propagation of the cascade. Also $\hat{\lambda}_j$ appears to decrease to a limiting value $\hat{\lambda}_*$ for values of $\lambda > \hat{\lambda}_*$. For $\lambda < \hat{\lambda}_*$, $\hat{\lambda}_j$ is independent of the stage j.
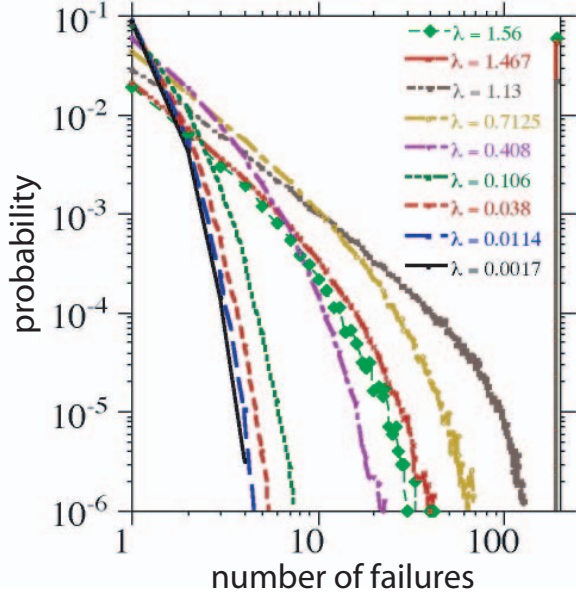
Fig. 4. Probability distributions of number of failures from CASCADE model using the values of $\hat{\lambda}_1$ from Fig. 2 and $\hat{\theta} = 0.095$. There are $n = 190$ components. Results for $\lambda > 1$ show a significant probability of all 190 components failing.

C. Random line failure modification to CASCADE

One effect present in OPA but not present in CASCADE is that overloaded lines do not always fail, but rather fail with probability p_1. Implementing this additional modification in CASCADE for various values of p_1 gives $\hat{\lambda}_j$ values as shown in Fig. 7. Some similar results for OPA are shown in Fig. 8 and there is now some qualitative similarity between OPA and the further modified version of CASCADE. In particular, for lower values of p_1, $\hat{\lambda}_j$ increases with stage j.

VI. CONCLUSION

We have used the CASCADE probabilistic model of cascading failure and its approximations to define an estimator $\hat{\lambda}_j$ of the propagation of failures at stage j of the cascade. The approximations to CASCADE also describe the extent of the region of power law behavior in probability distributions of cascade size. Testing the estimator $\hat{\lambda}_j$ on data produced by the cascading blackout model OPA suggests that, while $\hat{\lambda}_1$ appears to reflect the initial propagation of line outages, $\hat{\lambda}_j$ may decrease or increase with j. Modifications to the CASCADE model that also produce the decrease or increase of $\hat{\lambda}_j$ with j suggest explanations of these effects. For example, the decrease in $\hat{\lambda}_j$ for larger λ may be attributed to the inhibition of line outages by load shedding after a sufficient number of lines are outaged.

These initial results show that the interplay between the CASCADE and OPA models is useful for understanding the propagation of failures in cascading blackouts and in particular will be helpful in devising and testing statistical estimators to quantify this propagation.
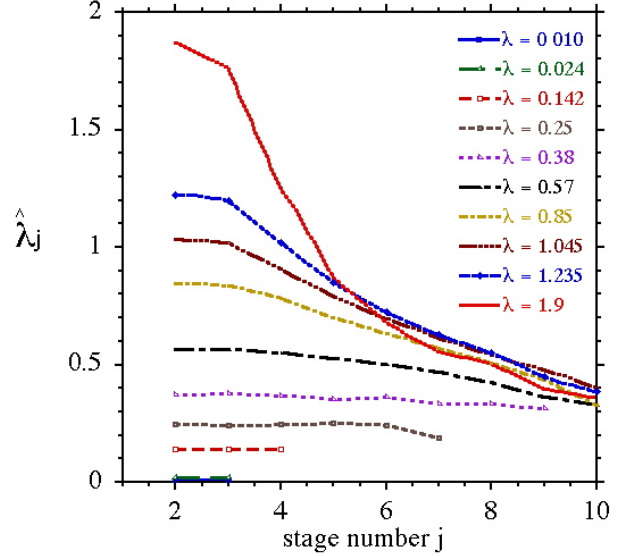

Fig. 5. $\hat{\lambda}_j$ as a function of stage j from CASCADE model with inhibition of line outages for various values of λ.
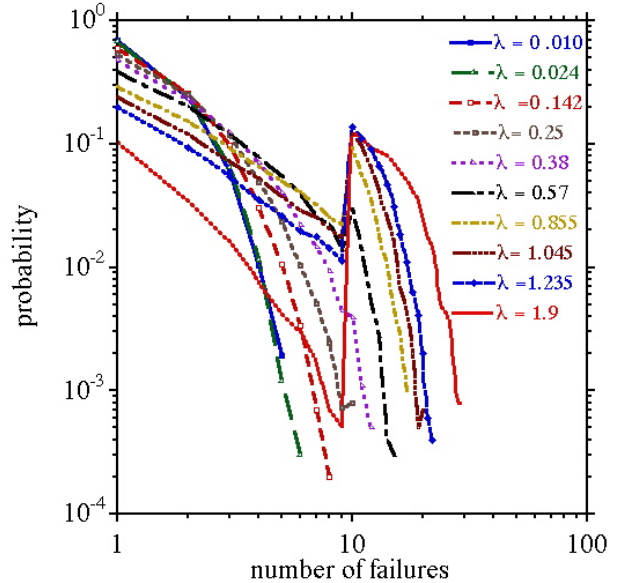

Fig. 6. Probability distributions of number of failures from CASCADE model with inhibition of line outages for various values of λ.

REFERENCES

[1] B.A. Carreras, V.E. Lynch, M. L. Sachtjen, I. Dobson, D.E. Newman, "Modeling blackout dynamics in power transmission networks with simple structure," *34th Hawaii Intl. Conf. System Sciences*, Maui, Hawaii, Jan. 2001.
   Note: authors' papers are available at
   http://eceserv0.ece.wisc.edu/~dobson/home.html
[2] B.A. Carreras, V.E. Lynch, I. Dobson, D.E. Newman, "Dynamics, criticality and self-organization in a model for blackouts in power transmission systems," *35th Hawaii Intl. Conf. System Sciences*, Hawaii, Jan. 2002.
[3] B.A. Carreras, V.E. Lynch, I. Dobson, D.E. Newman, "Critical points and transitions in an electric power transmission model for cascading failure blackouts," *Chaos*, vol. 12, no. 4, December 2002, pp. 985-994.

[4] B.A. Carreras, V.E. Lynch, D.E. Newman, I. Dobson, "Blackout mitigation assessment in power transmission systems," *36th Hawaii Intl. Conf. System Sciences*, Hawaii, Jan. 2003.

[5] B.A. Carreras, V.E. Lynch, I. Dobson, D.E. Newman, "Dynamical and probabilistic approaches to the study of blackout vulnerability of the power transmission grid," *37th Hawaii Intl. Conf. System Sciences*, Hawaii, Jan. 2004.

[6] J. Chen, J.S. Thorp, "A reliability study of transmission system protection via a hidden failure DC load flow model," *IEE Fifth International Conference on Power System Management and Control*, 2002, pp. 384-389.

[7] J. Chen, J.S. Thorp, I. Dobson, "Cascading dynamics and mitigation assessment in power system disturbances via a hidden failure model," preprint, 2003.

[8] Q. Chen, K. Zhu, J.D. McCalley, "Dynamic decision-event trees for rapid response to unfolding events in bulk transmission systems," *IEEE Porto Power Tech Proceedings*, vol. 2, Portugal 2001.

[9] P.C. Consul, "A simple urn model dependent upon predetermined strategy, *Sankhyā: The Indian Journal of Statistics, Series B*, vol. 36, part 4, 1974, pp. 391-399.

[10] P.C. Consul, "On some models leading to the generalized Poisson distribution," *Communications in Statistics - Theory and Methods*, vol. 17, no. 2, pp. 423-442, 1988.

[11] P.C. Consul, *Generalized Poisson distributions*, New York:Dekker, 1989.

[12] I. Dobson, B.A. Carreras, V. Lynch, D.E. Newman, "An initial model for complex dynamics in electric power system blackouts," *34th Hawaii Intl. Conf. System Sciences*, Maui, Hawaii, Jan. 2001.

[13] I. Dobson, B.A. Carreras, D.E. Newman, "A probabilistic loading-dependent model of cascading failure and possible implications for blackouts," *36th Hawaii Intl. Conf. System Sciences*, Hawaii, January 2003.

[14] I. Dobson, B.A. Carreras, D.E. Newman, "A branching process approximation to cascading load-dependent system failure," *37th Hawaii Intl. Conf. System Sciences*," Hawaii, January 2004.

[15] I. Dobson, B.A. Carreras, D.E. Newman, "Probabilistic load-dependent cascading failure with limited component interactions," *IEEE International Symposium on Circuits and Systems*, Vancouver Canada, May 2004.

[16] I. Dobson, B. A. Carreras, D. E. Newman, "A loading-dependent model of probabilistic cascading failure," to appear in *Probability in the Engineering and Informational Sciences*, 2005.

[17] T.E. Harris, *Theory of branching processes*, Dover, New York 1989.

[18] NERC, *1996 system disturbances*, North American Electric Reliability Council, Princeton Forrestal Village, 116-390 Village Boulevard, Princeton, New Jersey 08540-5731, 2002.

[19] M. Ni, J.D. McCalley, V. Vittal, T. Tayyib, "Online risk-based security assessment," *IEEE Transactions on Power Systems*, vol. 18, no. 1, pp. 258 -265, Feb. 2003.

[20] M.A. Rios, D.S. Kirschen, D. Jawayeera, D.P. Nedic, R.N. Allan, "Value of security: modeling time-dependent phenomena and weather conditions," *IEEE Transactions on Power Systems*, vol. 17, no. 3, 2002, pp. 543-548.

[21] U.S.-Canada Power System Outage Task Force, *Final Report on the August 14th blackout in the United States and Canada*, United States Department of Energy and National Resources Canada, April 2004.
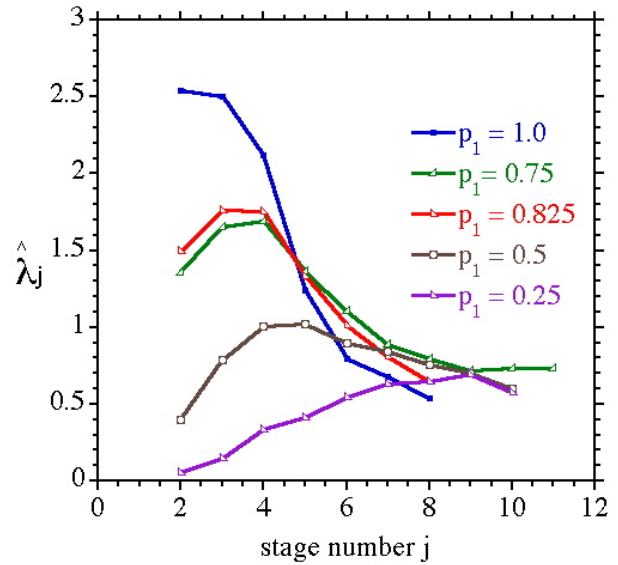
Fig. 7. $\hat{\lambda}_j$ as a function of stage j from CASCADE model with inhibition of line outages and overloaded lines outaged with probability p_1 for various values of p_1. The parameters are $p = 0.25$, $k = 10$, $r_{\max} = 15$, $n = 200$.
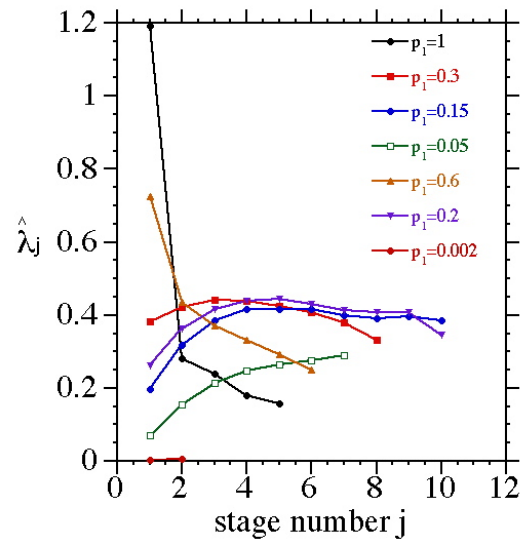

Fig. 8. $\hat{\lambda}_j$ as a function of stage j from OPA model for various values of p_1 on a 94 node tree-like network.