Analysis of Electric Power System Disturbance Data

Jie Chen
School of ECE
Cornell University
Ithaca, NY, 14853
jc227@cornell.edu

James S. Thorp
School of ECE
Cornell University
Ithaca, NY, 14853
jst6@cornell.edu

Manu Parashar
School of ECE
Cornell University
Ithaca, NY, 14853
mp28@cornell.edu

Abstract

In this paper, NERC (North American Reliability Council) records of power system disturbances for the year 1984 through 1999 are explored. The disturbance sizes show a power law distribution which confirms the early results [2]. Further, we find that the probability density of time intervals between disturbances can be given an exponential fit. Based on this observation, the suitability of applying SWV analysis to power system disturbances data is questioned. An artificial time series is constructed to support our idea. A fuse model using DC load flow and fuse protection is presented to simulating the cascading events in power transmission networks. Some initial simulation results are shown to be consistent with NERC data. Besides SOC and HOT, this model gives another way to investigate power law behavior in power system disturbances.

1. Introduction

Electric power transmission networks are complex systems which undergo non-periodic major cascading disruptions [1]. People tended to focus on individual causes of these disturbances and thought the probability of occurrence of blackouts decays exponentially with the event size. This is in contrast to the recently found power law "tail" of blackouts by Carreras et al. [2].

In this paper, after Carreras et al., we extend their analysis of NERC disturbance records for the years 1994 through 1998 to data from 1984 to 1999, which are the longest records we have found for disturbances in the North American power transmission system. We confirm their results of power law "tail" in blackout data, while challenge the correctness of applying SWV (scaled windowed variance) method to detect long term memory in blackout time series.

Many distributions of observed quantities in a wide variety of complex systems, such as earthquakes, sand-piles, and even biological evolution, exhibit power law form in their tails, sometimes called "heavy tail distribution". During the last decades, books, conference proceedings and papers have appeared concerning such power law scaling behavior. Several different power-law-producing mechanisms have been proposed, among which the theory of self-organized criticality (SOC) [4,8] is the most accepted and investigated. Recently, another idea, highly optimized tolerance (HOT) [9], was added to this toolkit helping explain power law behavior in designed systems. Here, we present a power-transmission-network oriented hidden failure model, DC fuse model, which can also produce power law distribution, and hope to capture the features of cascading events in power systems.

2. Analysis of NERC disturbances records

2.1. Description of NERC disturbances data

The disturbances data (NERC data) comes from the Disturbance Analysis Working Group (DAWG) Database which summarizes the disturbances that have occurred on the bulk electric systems of the electric utilities in North America [1]. It is the best-recorded source of blackouts in the North American power transmission system.

Here, we explore NERC disturbances data between the year 1984 and 1999, from which we construct blackout time series signals. The constructed time series is always zero except at those instances when blackouts occur, in which case the signal is equal to the size of the blackout happened at that time. The "event size", i.e., blackout size, is measured by three different quantities, the amount of power loss (MW), the number of customers affected and the restoration time (Minute). The time series is constructed with the resolution of a day according to these three measures. Also note that there are three electrically strongly interconnected areas: the entire eastern United States (Eastern US), Texas (ERCOT) and the western states (WSCC). A time series is constructed for each interconnection. Because ERCOT has too few records (only three in 16 years) to be analyzed, only six time series are constructed, i.e., three different measures within two interconnected areas (WSCC and Eastern US).

2.2. Power law "tail" in disturbance time series

Power law distribution has been found in many complex systems, so we are quite interested in whether or not it also exists in power systems. Since we only have limited data, using relative frequency distribution tends to have big fluctuations and hence it is difficult to identify the Probability Distribution Function (PDF) of the blackout data. Here, we turn to Cumulative Distribution Function (CDF), which is much nicer to deal with and tends to smooth out the fluctuations.

Suppose a random variable X has power law PDF f(x):

$$f(x) \sim x^{-\alpha} \quad (1)$$

Its CDF, then, will be

$$F(x) \sim \int_1^x t^{-\alpha} dt \sim 1 - x^{-(\alpha-1)} \sim 1 - x^{-\beta} \quad (2)$$

where • = •-1

For convenience, we consider following function P(x),

$$P(x) = P(X > x) = 1 - P(X \le X) = 1 - F(x) \sim x^{-\beta} \quad (3)$$

Taking logarithms to both sides, we get

$$\log P(x) \sim -\beta \log x \quad (4)$$

Therefore, if we draw the log-log plot of P(x) versus x (event size), we will find out if we can obtain a good linear fit, which means power law distribution.

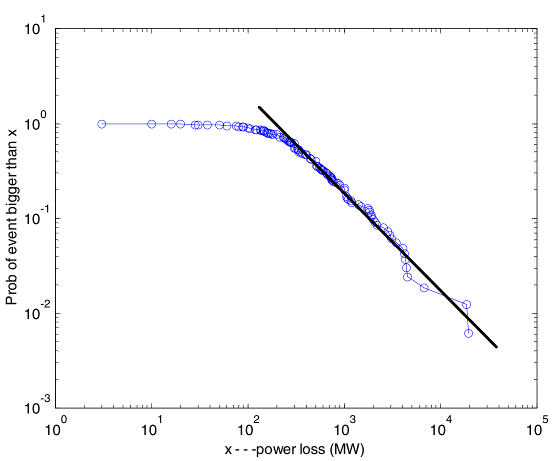

Figure 1. Power law "tail" of blackout power loss time series in Eastern US

Fig.1 shows such a plot of power loss data in the Eastern U. S. Obviously, its "tail" part can be fitted very well with a straight line. By drawing log-log plots of the remaining five time series data, we also can easily find a linear fit for each of them, respectively. Table 1 shows the fitted exponents • and • for each of the six data sets and the results confirm those found in [2].

As is known, if the exponent • of the power law distribution is smaller than 3, then no second moment will exist, and the standard deviation is infinite in the limit of infinite system size. Also, if • is smaller than 2, the mean will be unbounded. It is clear that based on the size measures we defined and disturbances data we used, the variance and mean of the blackout sizes are both unbounded for each of the two interconnected areas (WSCC and Eastern U.S), implying whole system size blackout could be possible.

2.3. Hurst exponent in disturbance time series

The Hurst exponent, often termed the self-similarity exponent or scaling exponent, is a measure developed to characterize the "dependence", or correlation, between distant samples in a time series. Carreras et al. [2] studied NERC data between the year 1994 and 1997 and found power system blackout time series had a Hurst exponent near 0.7. In this paper, we extend their analysis to data from 1984 to 1999.

First, we use the Scaled Windowed Variance (SWV) method, after Carreras et al. [2], to estimate the Hurst exponents for the six constructed time series. Three SWV methods, namely standard, linear detrended (LD) and bridge detrended (BD) [3], are applied. The results in Fig.2 and Fig.3 show that, within certain range of time lags, there is a clear linear fit, whose slope is the estimated Hurst exponent, for the averaged standard deviation versus window size for each time series.

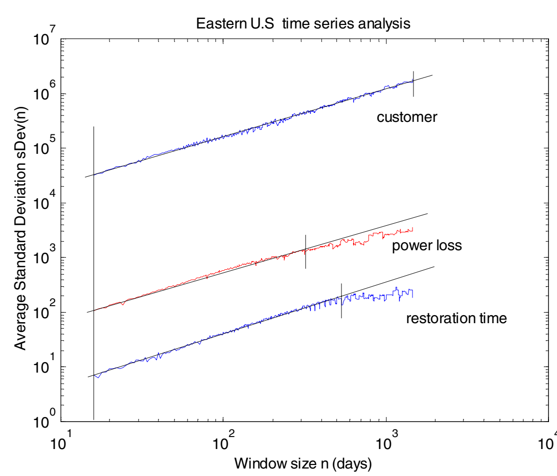

Figure 2. SWV analysis of time series for Eastern US blackouts.

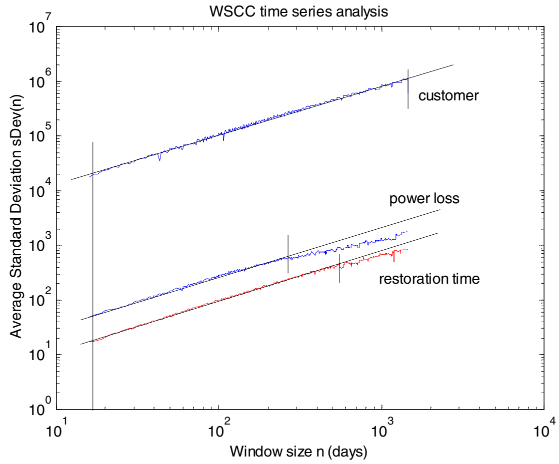

Figure 3. SWV analysis of time series for WSCC blackouts

Table 2 gives the Hurst exponents obtained for these six time series. Note that it was reported that standard SWV has negatively biased while LD SWV and BD SWV has positively biased estimates [3], the three Hurst exponents obtained by three SWV methods are averaged and we regard the averaged ones as the final estimates. Clearly, they are all around 0.8, greater than 0.5, which indicates a clear existence of long-range dependence. This result quite agrees with the one in [2]. However, based on our studies, which will be shown later in section 2.5, the results obtained above are suspect. The reason is that SWV methods probably are not appropriate for power system blackouts data and hence the conclusion from the analysis is not safe.

2.4. Distribution of time intervals between blackouts

In this section, we explore the distribution of time intervals between blackouts.

Assume there is no correlation between individual triggers of disturbances [2], then the disturbances will occur randomly at an average rate λ, as in a Poisson process. In this case, the PDF of the waiting times between events is exponential,

$$f(x) = \lambda e^{-\lambda x} \quad (5)$$

Further, as defined in section 2.2,

$$P(x) = e^{-\lambda x} \quad (6)$$

Taking logarithms to both sides, we get

$$\log P(x) = -\lambda x \quad (7)$$

So, if it is exponential distribution, by a semi-log plot of P(x) versus x (time intervals), we will give a linear fit for the data with a slope of λ.

To testify our assumption, we plot the semi-log diagrams of blackouts data. Fig.4 and Fig.5 show the result for WSCC data and Eastern US respectively. For both plots, if we ignore the biggest time interval, we can get a good straight line fit for each.

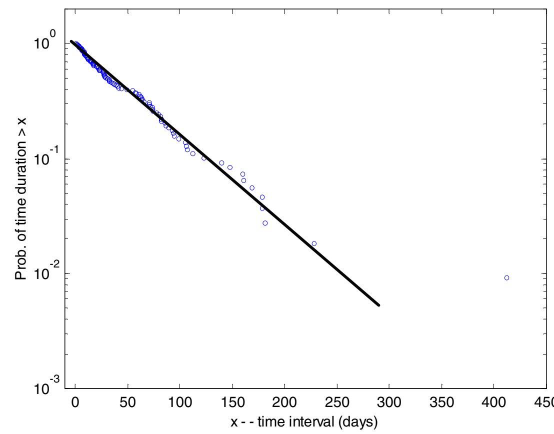

Figure 4. Distribution of time interval between events in WSCC region

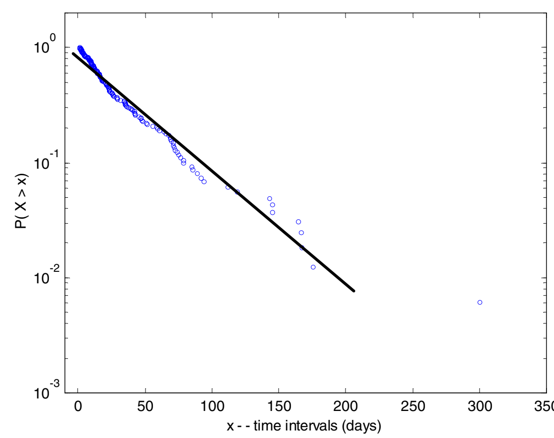

Figure 5. Distribution of time interval between events in Eastern US region

We also studied the distribution of time intervals, $f_{P0}(x)$, between events larger than a given size P_0. By setting different threshold sizes, we can get different sets of time-interval data. When applying semi-log plots mentioned above to these data, we get quite good linear fits. Fig.6 and Fig.7 give the time-interval data and their linear fits for selected threshold sizes of power loss time series. Table 3 shows the fitted slopes λ and their corresponding threshold P0. Note that we only give the fitted number up to threshold size 900MW, because for even larger P0, the number of data will become too small to lead to meaningful conclusions.

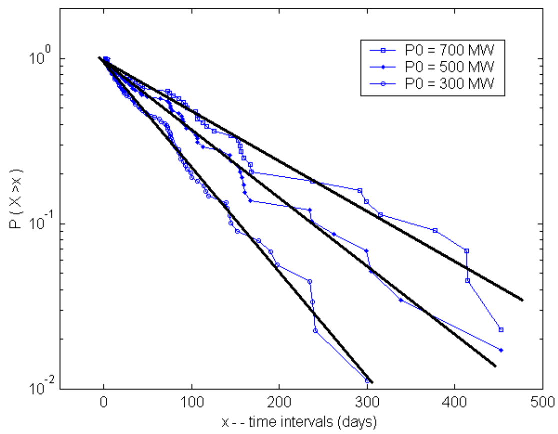

Figure 6. Distribution of time interval between events bigger than P0 in Eastern U.S region

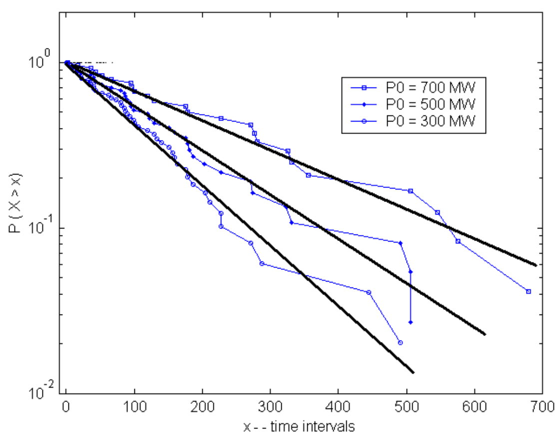

Figure 7. Distribution of time interval between events bigger than P0 in WSCC region

It should be noted that the small number of recorded blackouts (only several hundred), and correspondingly small number of time intervals are reflected in Fig.6 and Fig.7. Nonetheless, an exponential distribution fit for time intervals is not surprising, and could be understood from the variety of random causes for the disturbances.

2.5. Hurst exponent of artificial time series

Since exponential distribution of waiting time between events implies a "memoryless" process [11], what is the origin of the long-term dependence (Hurst exponent around 0.8) in the time series? In this section, we show some results that can be obtained by applying SWV methods to a constructed time series analysis.

In [3], it says "*the scaled windowed variance methods are only appropriate for use on signals whose differences form a stationary signal*" and "*if a signal does not exhibit scaling of* $\overline{SD}$ *versus window size n over more than two orders of magnitude, the conclusions that can be drawn from the estimate,* $\hat{H}$ *, will be limited*". In our case, first, it is quite difficult to detect the stationarity from the limited blackouts data, hence we can not make sure whether or not SWV analysis is appropriate. Besides, the results in Fig.2 and Fig.3 tell us that no one time series exhibits scaling more than two decades (only customer data exhibit scaling about two decades). According to the statement in [3], therefore, we can not safely conclude that long-range dependence is present in blackouts time series even if the estimated Hurst exponent is around 0.8.

To further support our idea, we applied SWV to estimate the Hurst exponent of a constructed artificial time series, which is "power-loss like" but obviously uncorrelated. This time series has similar features to the blackout time series. Its waiting time between events is exponentially distributed, ensuring that it has "no memory", and at each time when an event "occurs", we use a random power law distribution to generate the "event size". From how it is constructed, this time series should have no correlations.

The random exponential and power-law samples used are generated from uniform distribution by the inversion method. In brief, given that we wish to generate random variables with a PDF $f_X(\cdot)$, and CDF $F_X(\cdot)$, variable X generated from Equ.(8) has the required distribution [7]:

$$X = F_X^{-1}(R) \tag{8}$$

where R ~ U(0,1) is uniform distribution.

So, for exponential distribution:

$$f_X(x) = \lambda e^{-\lambda x}(x \ge 0) \tag{9}$$

The computer-generated random sample is

$$X = F_X^{-1}(R) = -\frac{1}{\lambda}\ln(1-R)) \tag{10}$$

Since R is identically distributed to (1-R),

$$X = -\frac{1}{\lambda}\ln R \tag{11}$$

Similarly, for power law distribution:

$$f_X(x) = kx^{-\alpha}(\alpha > 1) \tag{12}$$

The random sample will be

$$X = (\frac{\alpha-1}{k}R)^{-\frac{1}{\alpha-1}} = K \cdot R^{-\frac{1}{\alpha-1}} \tag{13}$$

In our experiment, we set λ = 0.02, K = 1, and α = 2 for the artificial time series. Fig.8 shows the constructed time series, and its estimated Hurst exponent is 0.72 by SWV analysis. To test the significance of the result, we run the experiment another ten times, and got the following Hurst exponents:

0.71, 0.87, 0.86, 0.72, 0.84, 0.67, 0.76, 0.82, 0.92, 0.65

Clearly, they are all well above 0.5, and should indicate long time dependence in each of these time series. However, this is in contrast to the way we construct the time series, which should be uncorrelated. Hence, although we can not, by doing this experiment, conclude that there is no correlation in the power system disturbances time series, we could doubt the suitability of applying the SWV method to such time series, which has a huge number of zeros and only a tiny fraction of nonzeros. Therefore, the value of 0.8 obtained in section 2.2 is not convincing, which renders the correlation-detection task inconclusive. In this case, longer, more detailed disturbances records, or more refined analysis methods are expected.

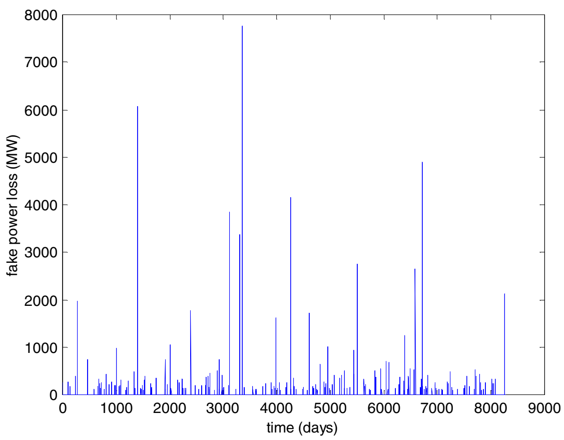

Figure 8. Artificial time series with exponentially distributed waiting time and power-law distributed event size

3. DC fuse model generates power law distribution

The power law "tail" of blackouts data, which was taken as evidence of SOC by Carreras et al. [2], is not necessarily an indication of dynamical self-organization into a critical stationary state. For example, Carlson and Doyle proposed another theory --- Highly Optimized Tolerance (HOT) [9] --- to explain power law distribution in power system blackouts. Here, we present another way of power law generation based on the detailed behavior of the power system protection system. Notice that either SOC or HOT tends to explain a variety of complex systems, while our model only focuses on power transmission networks.

3.1. Description of DC fuse model

Our presented model is a hidden failure model [5,10] with DC load flow and fuse protection. Fig.9 shows a diagram of the model.

The mesh represents the network of power transmission systems, and can be an artificial or real network structure. The nodes of the mesh represent the buses of the power network, and can be chosen as generations or loads. Branches act as transmission lines. For simplicity, they only have resistances and allow DC load flow. In every branch, there is a fuse acting as the relay system in a real system. Each line has a different load dependent probability of tripping incorrectly. A simple model is shown in Fig.10 where the probability of an exposed line tripping incorrectly is modeled as a function of line load flow seen by the line fuse.

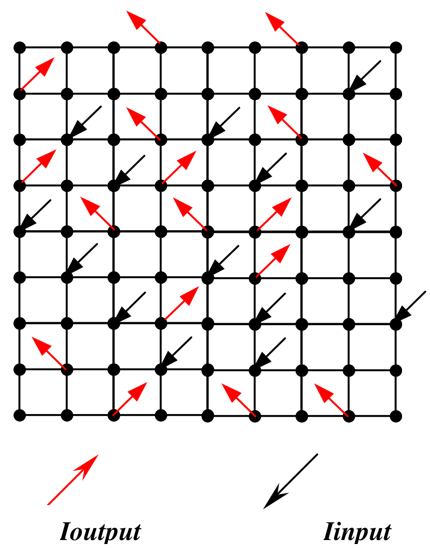

Figure 9. The fuse model of simplified power network

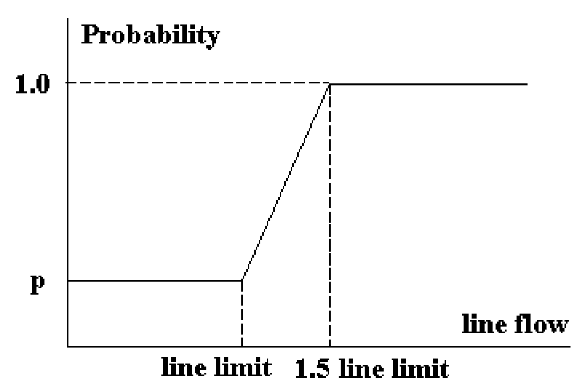

Figure 10. Probability of an exposed line tripping incorrectly

To calculate the "power loss" in cascading simulation events, a linear programming (LP) technique is used to find the DC load flow after each tripping. The LP problem formulation is:

$$\begin{aligned} \max \quad & f(I) = \sum I_{output} \\ \text{such that} \quad & I_{\min} <= I <= I_{\max} \\ & |J| <= J_{\max} \\ & I_{input} + I_{output} = 0 \end{aligned} \qquad (14)$$

where **I** is the input and output (generation and load), **J** is line load flow and $\mathbf{J_{max}}$ is line constraint, $\mathbf{I_{min}}$ and $\mathbf{I_{max}}$ are lower and upper bounds of **I**. (upper and lower limits of generator inputs and load outputs)

The brief analogy between the fuse model and the real network is shown in Table 4.

Table 4. Analogy between real power network and fuse model

Fuse Model	Real power network
Node	Bus
I_{input}	Generation
I_{output}	Load
Branch	Transmission Line
DC load flow	AC load flow
Fuse in branches	Relay system

There are two time scales in this model. The slow one has unit time intervals referred to as "days". In each "day", a blackout-triggering event will occur randomly. The other is the duration of individual cascading event, which is much faster than "day" and can be referred to as "minute". Hence, it is reasonable to assume no matter how long the cascading outage lasts, it will definitely end in a "day". Based on our analysis shown in section 2, exponentially distributed waiting time between outage-triggering events, which belongs to the slow time scale, has been applied. The growth in system load is also incorporated into the simulation model. For simplicity, linear growth is adopted.

3.2. Simulation algorithm

Before the simulation starts, we need to choose a network structure, determine generation and load pattern, initial DC load flow, and line constraints. Then the simulation is carried out on two time scales respectively. For the main loop, i.e., "day" time scale, below steps are followed:

1. Check if it is the "date" to increase the system load. Do as defined.
2. Decide whether or not a disturbance will occur, which is determined by exponential distribution of waiting time. If it will have a disturbance, enter sub loop; otherwise, record DC power loss as zero and start new main loop iteration.

For the sub loop, i.e., "minute" time scale, detailed simulation steps are listed as follows:

1. Randomly select a line as the initial tripping line leading to a possible cascading event.
2. Trip the selected line and compute the DC load flow using normal circuit equations.
3. Check for violations in line flow constraints and trip the line upon violation.
4. If there is no violation, determine exposed lines, which are all lines connected to the last tripped line and find the probability of tripping for each exposed line according to Fig.10. Note that the spread of disturbances is one-dimensional in power systems, hence, the case of more than one line trip at the same time rarely happens. In the simulation, we let one and only one line trip at one time. In specific, if more than one line might trip, the one with higher tripping probability is selected to be the next tripping line.
5. Shed load, if necessary, to keep the system stable, which requires all line flows less than their limits. LP technique defined as Equ.(14) is used to help out load shedding.
6. If no lines will be tripped, record the DC load lost and return to main loop starting a new iteration. Otherwise, go back to step 2.

3.3. Simulation result

In our initial simulation, the structure and parameters of transmission lines in WSCC 179-bus equivalent system [5] are used. We chose this particular system rather than a fake one for the sole purpose of testing if the simple model can generate results similar to those of a real system.

We run the simulation for 6000 "days", as if the system has been evolving during about 16 years. The system load is set to be increased 0.2% every other "month", which means 60 "days". Fig.11 and Fig.12 show the simulation results of DC power loss time series and relative cumulative frequency versus DC power loss. The power law "tail" of the simulated cascading events is obvious and the fitted exponent is 1.2, which is quite close to 1.07 for WSCC real blackout data.

With the evolving model, i.e., the load increases with time, we get quite similar result with real systems. Then what will happen if without the growth of load? It is a question of interest and can help us identify the role of load growth in producing power law distribution. The simulation result of the modified model with constant system load is shown in Fig.13, which indicates a power law distribution within the range of 3 ~ 20% loss of total system load. The power law exponent obtained in this case is 1.37, also a close result. This result could suggest that "load growth" may not play an important role in generating power law "tails".

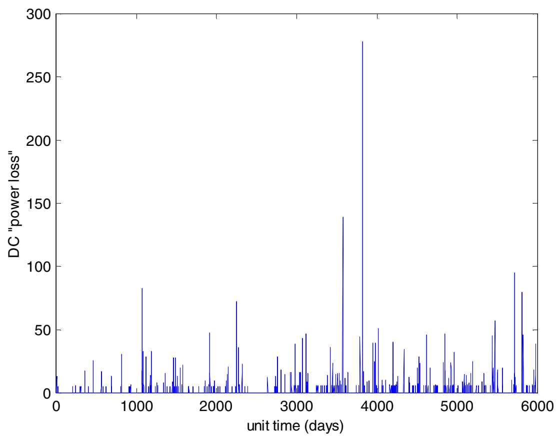

Figure 11. Simulated DC power loss time series in growing-load model

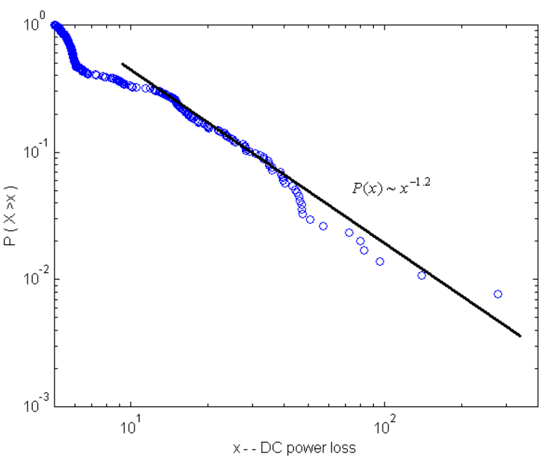

Figure 12. Power law "tail" of simulated DC power loss time series in growing-load model

Although our model includes basic components, such as load flow solving (circuit equations used), protective actions (line tripping and load shedding), we are still not quite sure about how much it can capture in simulating cascading events of real power systems. What we present here is a "power system like" model, which seems neither SOC nor HOT, but also can produce power law distribution of cascading events. That is, before further evidence is found to support SOC or HOT or whatever, we want to provide another way to look for the possible mechanism of generating power law behavior in power system disturbances.

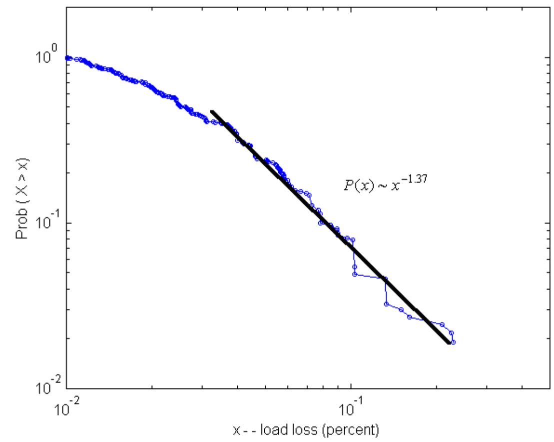

Figure 13. Power law "tail" of simulated DC power loss time series in constant-load model

4. Conclusion

Six time series of the NERC blackout data between the year 1984 and 1999 were constructed according to different measures of blackout size and different regions.

By analyzing these time series, we confirm the power law distribution of power system blackout size [2], and find the distribution of time intervals between events bigger than a given threshold is always exponential.

By the test of applying the SWV method to analyze an artificially constructed uncorrelated time series, we cast doubt on the suitability of detecting long time correlation among power system blackout data by using SWV analysis.

The DC fuse model is presented to explore the features of cascading events in power system. The initial simulation result shows power law distribution and is consistent with the result of real system. Different from either SOC or HOT, this model gives us another tool to investigate power law behavior in power system.

5. Acknowledgements

The authors acknowledge the research support by subcontract No.35352-6085 with Cornell University under WO 8333-04 from the Electric Power Institute and the U.S. Army Research Office.

6. Reference

[1] Information on electric systems disturbances in North America can be downloaded from the NERC website at http://www.nerc.com/dawg/database.html.

[2] B.A.Carreras, D.E.Newman, I.Dobson, and A.B.Poole, "Initial Evidence for Self-Organized Criticality in Electric Power System Blackouts", *Proceeding of Hawaii International Conference on System Sciences*, Jan. 4-7, 2000

[3] M.J.Cannon, D.B.Percival, D.C.Caccia, et al., "Evaluating Scaled Windowed Variance Methods for Estimating the Hurst Coefficient of Time Series", *Physica A,* vol. 241, pp.606-626, 1997.

[4] Per Bak, C.Tang, and K.Wiesenfeld, "Self-Organized Criticality: An Explanation of 1/f Noise", *Phys. Rev. Lett.*, vol. 59, pp. 381-4, 1987

[5] K.Bae, J.S.Throp, "A Stochastic Study of Hidden Failures in Power System Protection", *Decision Support Systems* 24, pp259-268, 1999

[6] Beran J. *Statistics for Long-Memory Processes*. New York: Chapman & Hall, 1994

[7] John Dagpunar, *Principles of Random Variate Generation*, Clarendon press, Oxford, 1988

[8] H.J. Jensen, Self-Organized Criticality – Emergent Complex Behavior in Physical and Biological Systems, Cambridge University Press, 1998

[9] J.M.Carlson, John Doyle, "Highly Optimized Tolerance: A Mechanism for Power Laws in Designed Systems", Phys. Rev. E 60, pp.1412-1427, 1999

[10] J.S.Thorp, A.G. Phadke, et al., "Anatomy of Power System Disturbances: Importance Sampling", PSCC, Dresden, Aug. 1996

[11] Leon-Garcia Alberto, *Probability and Random Processes for Electrical Engineering, 2nd ed.*, Addison-Wesley Publishing Company, 1994

Table 1. Power-law exponents of blackout time series

	Power loss		Customers affected		Restoration time	
	WSCC	Eastern U.S	WSCC	Eastern U.S	WSCC	Eastern U.S
β	1.07	0.97	0.75	0.71	0.94	0.83
α	2.07	1.97	1.75	1.71	1.94	1.83

Table 2. Hurst exponents of blackout time series

SWV method	Power loss		Customers affected		Restoration time	
	WSCC	Eastern U.S	WSCC	Eastern U.S	WSCC	Eastern U.S
Standard	0.728	0.724	0.761	0.754	0.754	0.786
LD	0.853	0.831	0.868	0.888	0.905	0.930
BD	0.850	0.830	0.874	0.880	0.910	0.939
Averaged H	0.810	0.795	0.834	0.841	0.856	0.885

Table 3. Fitted exponents λ versus threshold value P0

P0(MW)	0	100	200	300	400	500	600	700	800	900
Eastern U.S	0.0218	0.0189	0.0172	0.0145	0.0101	0.0095	0.0074	0.0073	0.0063	0.0068
WSCC	0.0175	0.0137	0.0100	0.0092	0.0076	0.0070	0.0063	0.0052	0.0043	0.0043