D read

Cascades, Crises and Contagion:

Threshold Dynamics on Sparse Interaction Networks.

Duncan J. Watts

MIT Sloan School of Management

And

Santa Fe Institute

Threshold dynamics are extremely widespread in both the social and natural sciences, arising in the study of integrate-and-fire oscillators[1], self-organised criticality[2], cellular automata[3], the evolution of conventions[4] and co-operation[5], crowd behaviour[6], crime[7], competing technologies[9], the diffusion of innovations[10] and fads[11]. Here we present a simple threshold model and show that when interactions between elements are defined by a sparse network, large cascades can be generated by small initial perturbations. We determine a measure of a network's susceptibility to such cascades that is a function of its structure and the heterogeneity of the population, and find conditions under which a network can be generically susceptible to rare, but large, cascades. Finally, we use the insights gained from this model to examine the power transmission network of the western united states[12].

Binary decision problems have been studied in the social sciences under a number of guises, and vary widely in their micro-dynamical structure. The detailed decision mechanism underlying a choice between two competing technologies[9] is quite different to that governing the decision to join a protest[13] or contribute to a common good[6] Nevertheless, a useful rule of thumb in a broad class of binary decision problems is a *threshold rule*: If an individual infers that more than a certain fraction (f_c) of the population is in state 1, then it will switch to state 1, else it will switch to state 0 . Dynamics arises when a population of sites is initialised in state 0 (off) and some fraction F_0 is switched to state 1 (on) at time $t = 0$. Each site then updates its own state at each time step according the threshold rule, where the heterogeneity of the population can be captured by drawing individual thresholds randomly from a distribution $g(f)$ defined on the unit interval and normalised such that $\int_0^1 g(f)df = 1$. Threshold models similar to this one (but often exhibiting a characteristic *relaxation* time) also arise more generally in the natural sciences even when sites are not making decisions, but merely changing states; for instance in models of integrate-and-fire oscillators, self-organised criticality, and cascading failures in engineering systems.

The mean-field behaviour of this model -- where $N \to \infty$ and each site is connected equally to all others -- is particularly simple: $F_t = G(F_{t-1}) = \int_0^{F_{t-1}} g(f)df$. Hence the equilibria F_∞ of the dynamics are determined by the intersections of the 1D map $G(f)$ with the diagonal, and the stable equilibria are those for which $dG/df\,|_{f_*} = g(f_*) < 1$. Hence for a unimodal distribution with $g(0) = g(1) = 0$, the system will always converge to $F_\infty = \begin{cases} 0 & F_0 < f_* \\ 1 & F_0 > f_* \end{cases}$, with the intermediate, unstable equilibrium f_* acting as a switch.

One obvious consequence of complete connectivity is that, as long as $f_* > 0$, sufficiently small perturbations $(F_0 \to 0)$ will never grow to influence a macroscopic fraction of the population.

Large cascades--defined here as $F_\infty >> F_0$-- and even *global cascades* $(F_\infty = O(1))$, however, do occur in real social[13], economic[14] and engineering[15] systems, most strikingly when they are generated by small perturbations $(F_0 = o(1))$ or *shocks*. Furthermore, the system undergoing the cascade may have appeared stable with respect to many previous shocks of the same size--a phenomenon embodied in the concept of "normal accidents"[16]. Here we suggest that the vulnerability of large interconnected systems to such cascades can be explained in terms of the structure of a *sparse interaction network*, where individuals (vertices) still exhibit threshold dynamics, but measure only the states of their

immediate neighbours on the network, and where the average number of neighbours (K) is now much smaller than the population size $(K << N)$.

In the absence of any known geometry for the problem, a natural first choice of model for a local interaction network is a random graph[17] with average vertex degree $0 \leq K << N$. A useful feature of the threshold model is that for a homogeneous population $\left(g(f) = \delta(f_c)\right)$, the probability of a global cascade being generated by a single vertex switching on at time $t = 0$ is essentially a graph problem: Any vertex (v) adjacent to at least one *on* site will switch on if it has a threshold $f_c \leq 1/K_v$; otherwise it will not. Hence we may expect a large cascade to occur with finite probability when the random graph contains a *giant component*[17] (i.e. $O(N)$ vertices) consisting solely of vertices with $K_v \leq \left\lfloor 1/f_c \right\rfloor = M$. The general relation between M, K, and the size of the largest *M-delimited component* of a random graph is unknown[18]; however, it is known that when there are no restrictions on degree (i.e. $M = \infty$), the largest component in a random becomes a giant component when $K = 1$.[17] This result suggests[19] that an *M-delimited giant component* will appear when the mean of the *truncated* degree distribution $K_M = K \cdot Q(M, K)$ exceeds one (where the degree distribution $P(x)$ is truncated at $x = M$, $P(x)$ is the Poisson distribution and $Q(M, K)$ is the incomplete gamma function).

This condition defines a critical window in the (f_c, K) plane within which we expect cascades to occur. Figure 1 shows the window (solid line) along with boundaries (symbols) delineating the onset of cascades. For large f_c, there is a clear region inside the window in which cascades do occur $(F_\infty >> F_0)$, but are not global $(F_\infty = O(1))$; however this region diminishes as f_c decreases. Where the two regions overlap, the probability of a cascade is equal to the fractional size of the largest M-delimited component, but the size of any cascade that occurs is that of the entire connected component, regardless of M (Figure 1, Inset). This distinction is important, as along the upper boundary of the critical window, it results in cascades that are very infrequent but --because of the high connectedness of the graph--extremely large; hence the corresponding systems can appear stable through many small shocks and then suddenly exhibit a global cascade.

The same analysis can be applied to the more general case of a heterogeneous distribution of thresholds, where now K_M^* is calculated as a weighted average, and the weighting function is $G\left(1/f\right)$. For a normal distribution of thresholds, with mean f_c, Figure 1 (dashed lines) shows how the critical window opens as the width (σ) of the distribution increases. For random graphs then, the critical window appears to be a useful construct,

requiring only very general features of the network's threshold distribution and distribution of edges, to derive a condition for its stability.

Recent work[20-24] has suggested, however, that many real networks are not random, as they exhibit a much higher *local* density of connections than do random graphs, whilst remaining globally sparse $(K << N)$. The above methods can be generalised to incorporate local structure as follows: partition the population into $\tilde{N} = \lfloor N/n \rfloor$ disjoint groups of size n (where we assume here that N is evenly divisible by n). Each group is first allowed to *overlap* uniformly at random to an average of $\tilde{K}$ other groups and then edges are created, also at random, but only between vertices in the same and overlapping groups, such that each vertex still has an average of K neighbours. For fixed N and K, n and $\tilde{K}$ define a two-parameter family of graphs that interpolate between a set of isolated cliques (with maximal local density $\rho = 1$) when $\tilde{K} = 0$ and $n = K + 1$, and a random graph (with $\rho \approx K/N \sim 0$) when either $\tilde{K} = \tilde{N} - 1$ or $n = 1$ (see Figure 2). By fixing $n = K + 1$ and varying $0 \leq \tilde{K} \leq \tilde{N} - 1$, we can explore the effects of varying the local structure of a network through its maximum possible range without assuming anything about the underlying geometry of the network--an advance beyond previous attempts[25].

If a small seed $(s \approx n)$ is switched on initially, then the probability of any resulting cascade can be approximated by the random graph analysis above, but where K is now replaced by $\tilde{K}$ (see Figure 3) The dynamics is largely insensitive to the overall density of the graph (K)--and can yield cascades at much higher densities than for random graphs $(K >> 1)$--but highly sensitive to the density of inter-group connections $\left(\tilde{K}\right)$. In terms of social networks, if individuals in a population belong to too few groups (on average), then the local connections are dense enough to support the propagation of a cascade, but the population is insufficiently well connected for it to be grow beyond its seed group; conversely, if they belong to too many, the population is highly connected, but local connections are too sparse for the cascade to propagate along them.

It is possible to gain some intuition about real networks from the model -- a point we illustrate with the power transmission network of the western states coordinating council (WSCC). The WSCC network consists of 4941 power stations and sub-stations distributed throughout the US and Canada west of the Rocky Mountains, connected by 13,193 high voltage transmission cables. Although it has been shown to exhibit non-random structural features[20,26], as a first approximation we treat it as a random graph with $N = 4941$ and $K \approx 2.67$. Threshold dynamics are present because individual units are protected by relays that open under conditions of local stress or overloading[27], where

tripping of one relay typically places greater strain on its neighbours, and hence increases the likelihood of subsequent outages. Loading conditions and relay characteristics vary across the network, yielding a distribution of thresholds, where we recognise that this is an extremely simplistic--but illustrative--interpretation of the real dynamics.

Figure 4 shows a comparison between the probability and maximal size of cascades in the WSCC network with a homogeneous threshold f_c, and a cross-section of the critical window of a random graph at K=2.67 (See Figure 1). Two distinct regions--large, but not global cascades, and global cascades--are again apparent, where the maximum size of cascades in the outer region for the WSCC (roughly 5% of the network) is distinctly larger than that for a random graph, but only slightly larger than the size of the largest-ever cascading outage in the WSCC[27] (over 196 units or roughly 4% of the network[28]). The agreement with the random network model suggests that the density of the WSCC network $(K \approx 2.67)$ makes it nearly maximally vulnerable to threshold cascades; and further, that adding more connections to the network represents a trade-off between reducing the probability of a cascade but increasing the expected size of any that do occur.

In summary, we have presented a framework for assessing the robustness and stability of large, interconnected systems that focuses on the structure of the corresponding interaction network and demonstrates that, under certain conditions, apparently stable systems can exhibit occasional, very large cascades. Given the inherent difficult of testing, or even simulating, the response of networked systems to all possible disturbances under all possible conditions, we hope that this framework can potentially illuminate structural vulnerabilities of large networks, and that it is sufficiently generic to be adapted to multiple applications in the social, biological and engineering sciences.

1. Corral, Á., Pérez, C. J., Díaz-Guilera, A. and Arenas, A. Self-organized criticality and syncnronization in a lattice model of integrate-and-fire oscillators. *Phys. Rev. Lett.* **74(1),** 118-121 (1995).
2. Bak, P., Tang, C. and Wiesenfeld, K. Self-organized criticality: an explanation of $1/f$ noise. *Phys. Rev. Lett.* **59(4),** 381-384 (1987).
3. Moran, G. The r-majority vote action on 0-1 sequences. *Discrete Math.* **132,** 145-174 (1994).
4. Schelling, T. C. A study of binary choices with externalities. *J. Conflict Resolution.* **17(3),** 381-428 (1973).
5. Glance, N. S. and Huberman, B. A. The outbreak of cooperation. *J. Math. Soc.* **17(4),** 281-302 (1993).
6. Granovetter M. Threshold models of collective behavior. *Am. J. Soc.* **83(6),** 1420-1443 (1978).

7. Glaeser E. L., Sacerdote B. and Schheinkman J. A.. Crime and Social Interactions. *Quarterly J. Econ.* **111,** 507-548 (May 1996).
8. Arthur W. B. and Lane D. A. Information contagion. *Structural Change and Economic Dynamics.* **4(1),** 81-103 (1993).
9. Valente T. W. Social network thresholds in the diffusion of innovations. *Social Networks,* **18,** 69-89 (1996).
10. Bikhchandani, S., Hirshleifer, D. and Welch. I. A theory of fads, fashion, custom and cultural change as informational cascades. *J. Pol. Econ.* **100(5),** 992-1026 (1992).
11. Kosterev, D. N., Taylor, C. W. and Mittelstadt, W. A. Model validation for the August 10, 1996 WSCC System Outage. *IEEE Trans. on Power Systems.* **14(3),** 967-979 (1999).
12. Lohmann S. The dynamics of informational cascades: the Monday demonstrations in Leipzig, East Germany, 1989-91. *World Politics.* **47,** 42-101 (1994).
13. Aguirre, B. E., Quarantelli, E. L. and Mendoza, J. L. The collective behavior of fads: the characteristics, effects, and career of streaking. *Am. Soc. Rev.* **53,** 569-584 (1988).
14. De Vany, A. S. and Walls, W. D. Bose-Einstein dynamics and adaptive contracting in the motion picture industry. *The Economic Journal* **106,** 1493-1514 (1996).
15. Carlson, J. M. and Doyle, J. Highly optimized tolerance: a mechanism for power laws in designed systems. *Phys. Rev. E.* **60(2),** 1412-1427 (1999).
16. Perrow, C. *Normal accidents : living with high-risk technologies.* (Basic Books, New York, 1984).
17. Bollobas B. *Random Graphs.* (Academic, London, 1985).

18. Bollobas, B. Private communication.

19. S. Rasmussen, T. Luczak. Private communication

20. Watts D. J. and Strogatz S. H. Collective dynamics of 'small-world' networks. *Nature* **393,** 440-442 (1998).

21. Adamic L. A. The small world web. *Proceedings of the ECDL'99 Conference*, pp.443-452 (Springer, 1999).

22. Wagner A. and Fell. D. The small world inside large metabolic networks. (In review).

23. Sporns O., Tononi G. and Edelman G. M. Theoretical neuroanatomy: relating anatomical and functional connectivity in graphs and cortical connection matrices. *Cerebral Cortex* (in press).

24. Kogut B. and G. Walker. The small world of firm ownership in Germany: social capital and structural holes in large firm acquisitions--1993-1997. *Working paper,* (Reginald H. Jones Center, Wharton School 1999).

25. Watts D. J. *Small Worlds: The Dynamics of Networks between Order and Randomness*. (Princeton Univ. Press, Princeton, 1999).

26. Barabasi, A. and Albert, R. Emergence of scaling in random networks. *Science* **286,** 509-512 (1999).

27. Hauer, J. F. and Dagel. J. E. Consortium for electric reliability technology solutions, grid of the future: White paper on review of recent reliability issues and system events. U.S. Dept. of Energy. (1999). [get correct reference]

28. WSCC Operations Committee. Western Systems Coordinating Council Disturbance Report, August 10 1996. (Oct 18, 1996) [get correct reference]

Acknowledgements:

The author wishes to acknowledge helpful conversations with D. Callaway, P. Dodds, J. D. Farmer, J. Lynch, J. Spencer, and S. Strogatz.

Captions.

Figure 1. Phase diagram displaying critical window inside which cascades are predicted to occur on a random graph with N=1000 (and 100 realisations of the randomness). Solid lines represent the upper and lower values of K for which a giant component is expected in a random graph where no vertex in the component has degree $K_v > M = \left\lfloor 1/f_c \right\rfloor$. Symbols represent the onset of non-global (open circles) and global (solid circles) cascades that occur in the homogeneous dynamical model (i.e. every vertex has the same threshold f_c). Dotted lines show the transformation of the critical window when thresholds are now drawn randomly from a distribution $g(f)$ that is approximately normal and sharply peaked with mean f_c and $\sigma = 0.05$ (--), $\sigma = 0.1$ (-). Inset: cross section of the critical window for $f_c = 0.2$. The frequency of cascades (solid circles) is predicted by the size of the largest M-delimited component (solid line), where M=5, and the average size of the cascades (open circles), when they occur, is predicted by the size of the largest, standard connected component (dashed line). This indicates that rare, but catastrophic cascades are generically possible (at the right-hand extreme of the window).

Figure 2. Two-parameter family of networks for fixed N and K. The network is partitioned into $\tilde{N} = \left\lfloor \frac{N}{n} \right\rfloor$ distinct groups of size n (where, more generally, n could be drawn from a distribution); each group is associated with and average of $\tilde{K}$ other groups chosen uniformly at random, and vertices make random connections to an average of K others, where unlike random graphs, neighbours must be selected from within the same group or an immediately adjacent group. This constraint builds local structure into the network in a way that is reminiscent of social networks: individuals may have many friends, but they typically belong to a much smaller number of groups which overlap by virtue of joint memberships. When $\tilde{K} = 0$, all neighbours are in the same group, and the groups are isolated, resulting in a population of $\tilde{N}$ isolated cliques for $n = K + 1$. Hence, by setting $n = K + 1$ and varying $\tilde{K}$, the model can interpolate between maximal and minimal local structure (cliques and random graphs respectively), where the variation in n and $\tilde{K}$ is constrained by the conservation of total edges, $NK/2$.

Figure 3. Critical window for the network model in Figure 2, with N=10,000 and n=20. As in Figure 1, the solid lines represent the boundaries of the window as predicted by the existence of a giant M-delimited component, where now the nodes of the random graph are groups of size n that overlap by virtue of joint membership. Hence, in addition to the

threshold f_c, the governing parameter is $\tilde{K}$--the average number of groups to which a node belongs. The symbols again represent the onset of cascades that occur in the dynamical model.

Figure 4. Comparison between the frequency and maximal size of cascades in the WSCC network (circles and stars respectively) with those for random graphs with K=2.67 (solid and dashed lines). In both cases, a single node is switched on at t=0 and the results are averaged over 1000 random initial conditions.

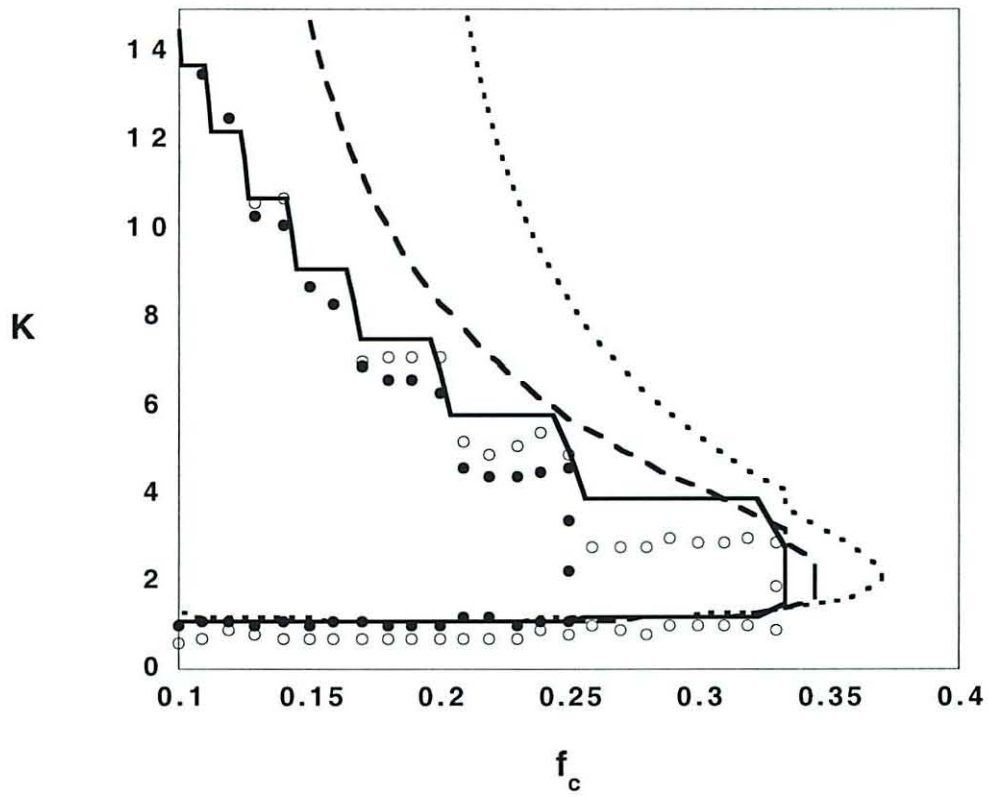


Figure 1

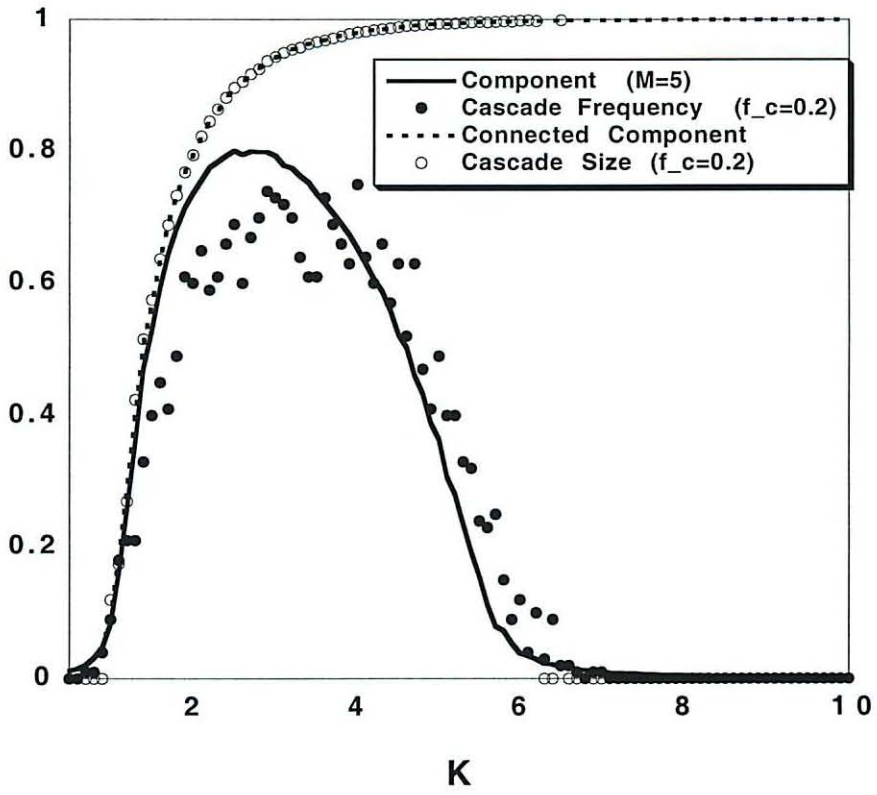


Figure 1 (insert)

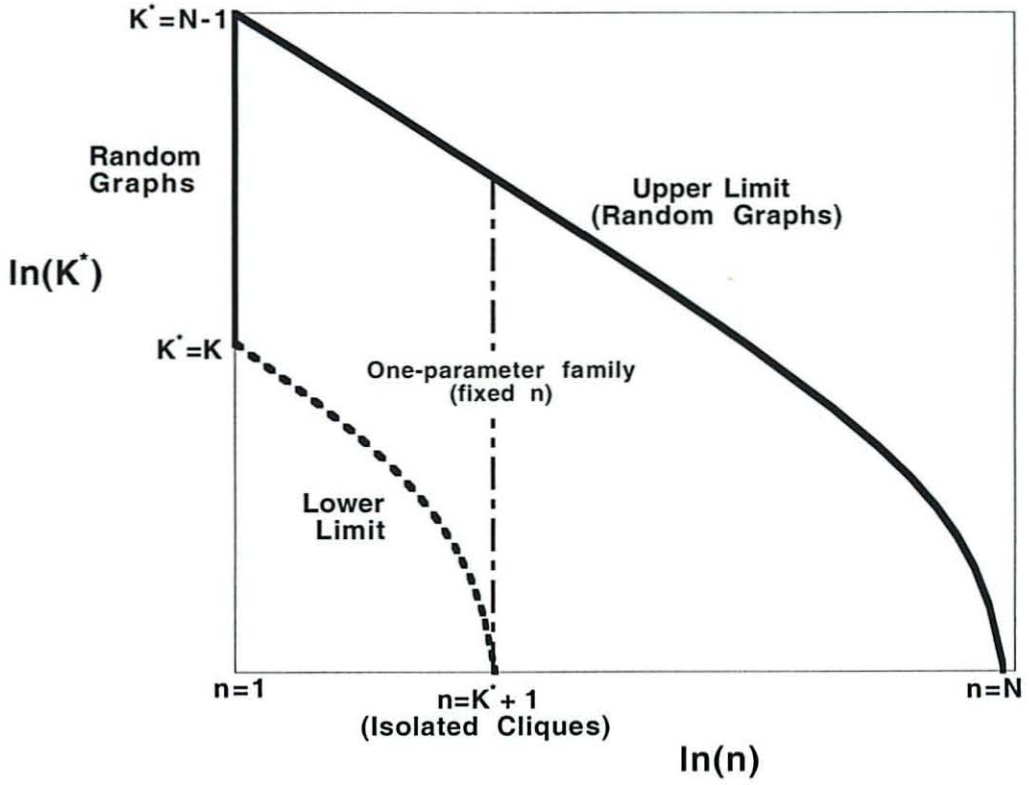

K*=N-1
Random
Graphs
ln(K*)
K*=K
Upper Limit
(Random Graphs)
One-parameter family
(fixed n)
Lower
Limit
n=1
n=K* + 1
(Isolated Cliques)
n=N
ln(n)

Figure 2

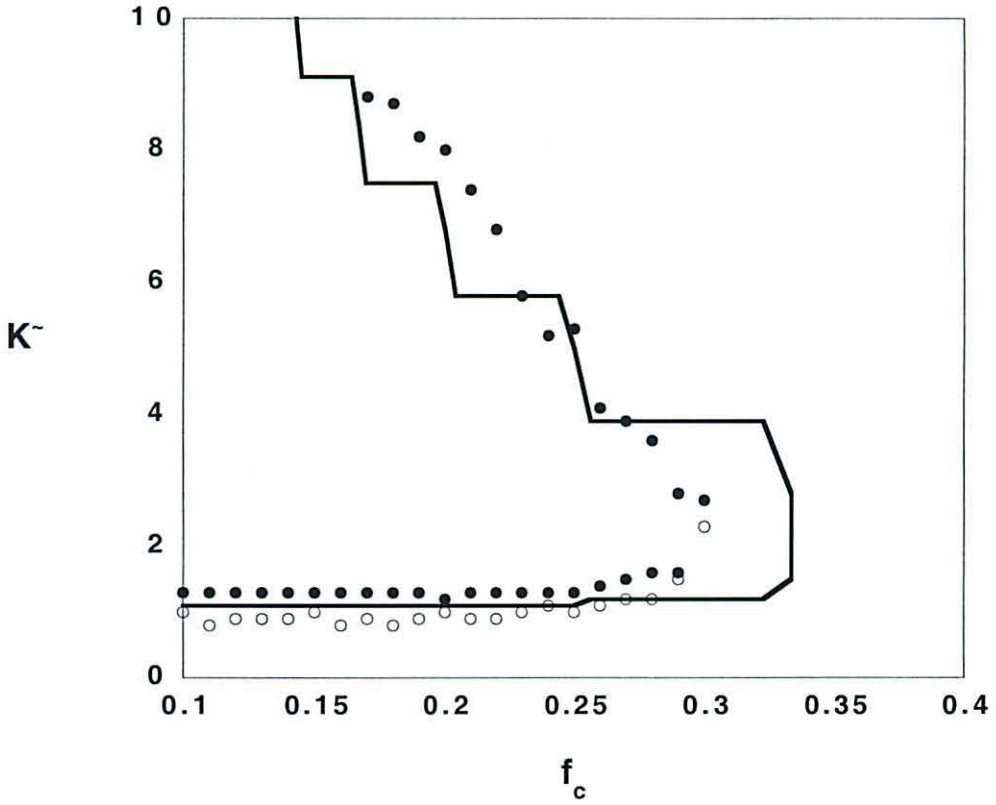


Figure 3

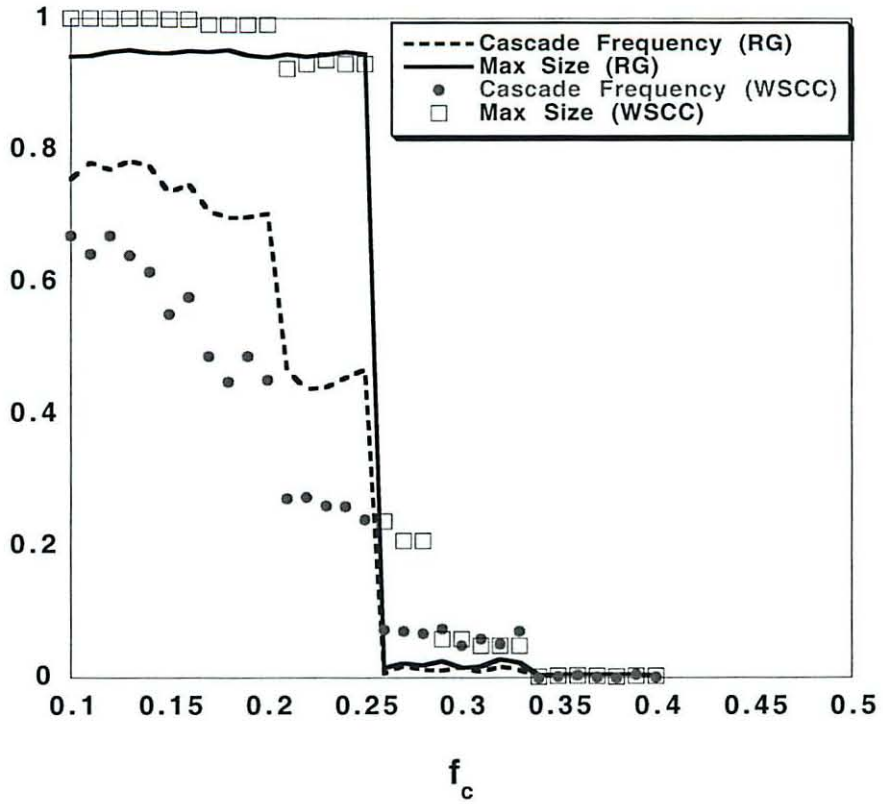

Cascade Frequency (RG)
Max Size (RG)
Cascade Frequency (WSCC)
Max Size (WSCC)
1
0.8
0.6
0.4
0.2
0
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
0.5
f_c

Figure 4