

Vulnerability Analysis Of Interdependent Infrastructure Systems

Giulio Galvan

Doctoral Researcher, Dept. of Civil Engineering, University of Bristol, Bristol, United Kingdom

Jitendra Agarwal

Senior Lecturer, Dept. of Civil Engineering, University of Bristol, Bristol, United Kingdom

ABSTRACT: As resilience of infrastructure systems gains importance to deal with the uncertainty related to extreme natural events, there is increasing emphasis on the design of systems that do not fail catastrophically. The consequences of a perturbation on a system depend both on the magnitude of the perturbation and the vulnerability of the system. The assessment of the vulnerability of infrastructure systems presents the challenge of dealing with their complexity. This paper presents a method to identify the elements of a system which have the potential to trigger cascading failures thus making the system vulnerable. A new predictive metric (X_I) is introduced and variations in the system parameters that could affect its predictive capabilities are explored. Networks which have properties comparable to real-world infrastructures such as transportation and utility supply systems are simulated. It is found that the correlation between the new metric and the behaviour of the system holds across all the spectrum of the simulations performed.

1. INTRODUCTION

Increasing interconnectivity between the elements of infrastructure systems is necessary to efficiently satisfy the needs of modern societies. Interdependent systems, however, allow damage to spread more widely thus increasing the severity of failure consequences. The existence of such risks has been demonstrated by events such as the 2003 blackout in the Northeast US, where the damage has been estimated to be 6.4 billion USD (Anderson & Geckil 2003) and the 2011 eruption of the Eyjafjallajökull volcano, that caused a systemic failure in the European air transportation network (Wilkinson et al. 2011). It is also widely acknowledged that the risk landscape that needs to be considered when managing interdependent infrastructure is constantly changing. Urbanization is leading to concentration of exposure in relatively small geographical areas. Climate change has the potential to alter the distribution and the intensity of adverse weather events. Further, as existing infrastructure ages, inherent safety margins are progressively being eroded.

Cost-benefit analyses and probabilistic risk assessments usually do not take into account systemic risks that arise due to the propagation of damage in one system to the interconnected systems. The exact magnitude of these effects is difficult to evaluate because of the complexity of the systems involved. In order to deal with this scenario of emerging systemic risks and uncertainty of extreme natural events, resilience has recently become the new safety paradigm.

Resilience requires robustness and vulnerability is sufficient for a lack of robustness (Blockley et al 2012). A first step towards resilience is to assess which hazards, internal or external, have the potential to affect large portions of the system and which disturbances, on the other hand, can be absorbed locally. This paper provides a methodology to identify the inherent vulnerabilities of an infrastructure system so that the resilience plans can be devised accordingly. The concept of resilience and its usefulness in practice is reviewed in Section 2. Section 3 summarizes the state-of-the-art on vulnerability research and illustrates its relation with resilience.

A methodology to assess the vulnerability of infrastructure networks towards cascading failure is presented in Section 4. Its performance under a wide range of conditions is evaluated through the electric power transmission system described in Section 5. The results are discussed in Section 6.

2. RESILIENCE

Resilience is often advocated as a desirable feature of complex systems, yet there are multiple and contrasting interpretations of this concept. For the purpose of this paper, it is taken as “the ability of a system, community or society exposed to hazards to resist, absorb, accommodate to and recover from the effects of a hazard in a timely and efficient manner, including through the preservation and restoration of its essential basic structures and functions”, as given by the UNISDR (2009). This definition encourages a holistic view of resilience where resources are allocated to all the phases of the disaster management process.

The first advantage with this definition of resilience is in the design methodology for protective infrastructure. Recent severe events such as the Tohōku Earthquake or Superstorm Sandy have highlighted that probabilistic approaches to design cannot provide all the answers to natural hazards. The tails of the statistical distributions of adverse natural events are affected by severe uncertainty and dealing with low-probability high-impact events is difficult. Planning for safety requires a severity threshold under which event probabilities can be modelled and protective infrastructure designed accordingly. Scenarios exceeding this severity level should be managed differently.

The second advantage lies in the costs of resilience measures. State-of-the-art resilience frameworks (Arup & Rockefeller Foundation 2014) include non-structural measures such as contingency planning, land-use regulations and supply chain diversification. Most of these interventions costs are negligible as compared to the upfront investments necessary to build the protective infrastructure required to withstand tail events.

There are, however, a number of issues with this broad definition of resilience. The dynamics of the post-disaster phase is highly nonlinear and poorly understood. Assessing a-priori the recovery trajectory of a disaster-struck community system is a challenging task. It depends (a) on the behaviour of community members, businesses and decision-makers, (b) on how the system is related to the rest of the environment and (c) on the extent of the initial damage compared to the size of the system. With these challenges in mind, it is argued that the first step towards the creation of a resilient system is to understand where it is vulnerable.

3. VULNERABILITY

“Vulnerability is susceptibility to damage – especially where small damage leads to disproportionate consequences” (Blockley et al. 2012). It derives from the internal organization of the system. The superposition of vulnerability with external hazards can give rise to negative consequences and a natural hazard can turn into a disaster.

This definition of vulnerability echoes with that of systemic risk given in Lorenz et al. (2009) i.e. a vulnerable system presents greater risks due to its internal structure. Until a perturbation appears in the right location and with the appropriate intensity, such vulnerabilities may be latent, but the risks associated with them are particularly high and must be taken care of. The recovery from adverse events has been shown to strongly depend on the initial damage it causes (Quarantelli 1999); therefore, a vulnerable system is also less likely to show resilience.

This distinction between local and systemic risks has also been pointed out by Taleb et al. (2014) when they argue that different risk management policies are needed to address different types of risks. Interconnectivity of infrastructures may generate systemic risks because failure in one part of the system has the potential to affect large portions of it.

3.1. Identifying vulnerabilities

Spatially distributed infrastructures can be modelled as network of elements providing the support for the flow of services. Research on complex networks has produced a number of metrics that provide diverse information about a graph and its components. In large scale-free networks nodal degree (i.e. the number of links at a node) has been to shown be an effective way to identify the elements that affect the vulnerability of the whole system (Albert et al. 2000), but on smaller networks with different topologies the correlation between nodal degree and system vulnerability is much lower (Dunn & Wilkinson 2013). The betweenness centrality of a node (i.e. the number of shortest paths through it) simulates well the flow of commodities through a network but it cannot reliably predict which element, if removed, will trigger the most severe consequences (Zio & Sansavini 2010). State-of-the-art research on spatial infrastructure networks uses a range of approaches and an extensive review is available in Ouyang (2014). Many researchers use full simulation of cascading failure processes in order to identify the elements that are the root cause of the system vulnerability. Such simulations, however, are computationally intensive and a simplified approach is presented in the next section.

4. METHODOLOGY

The vulnerability assessment approach has three steps (i) creating a model of the infrastructure, (ii) initiating a local damage and (iii) identifying the vulnerable elements.

4.1. Infrastructure modelling

A well-established modelling approach for spatially distributed infrastructures is to integrate complex network theory with engineering models (Johansson et al. 2013). Infrastructure systems are modelled using graphs and a flow model of physical quantity is defined. A graph is a set of n nodes (or vertices) joined by m links (or edges). Nodes may represent sources (origins), sinks (destinations) or junctions. Links are the flow channels.

Flow models of different complexity are possible. In this paper, the Motter and Lai (ML) model (Motter & Lai 2002) was used to simulate the distribution of flows in the system. It is a generic model that can be used to distribute the flow of a commodity or a service to the elements of different types of network. Two subsets of nodes need to be identified, generation V_G and distribution V_D , with cardinalities n_G and n_D respectively. The model assigns to each node i a load L_i proportional to the number of shortest paths between each generation-to-distribution pair that go through it. Mathematically,

$$L_i = \frac{1}{n_G n_D} \sum_{j \in V_G, k \in V_D} \frac{n_{jk(i)}}{n_{jk}} \quad (1)$$

where n_{jk} is the number of shortest paths from node j to node k and $n_{jk(i)}$ is the number of those paths that include i .

Once the flows through the network have been identified, the capacity of every node needs to be established. In engineering practice, the capacity is greater than the operating load by a safety margin α :

$$C_i = (1 + \alpha) * L_i \quad (2)$$

The baseline for the capacity considered in this work is 120% of the original load, which governs the state of the electric power transmission system used as a case study (Section 5). The ratio between load and capacity, however, varies with the fluctuations in the demand as well as with the design standards of the system. The effects of variations in the capacity distribution are shown in Section 6.

4.2. Disruption modelling

Disruptions to infrastructure systems are modelled either by considering the physical nature of the threat (Dueñas-Osorio et al. 2007), or by considering abstract scenarios such as the removal of elements in the network (Agarwal et al. 2001). The first approach has the advantage of producing hazard-informed risk assessments, while the second investigates the vulnerability of systems irrespective of the nature of the external threat. In this paper, the latter approach is followed and perturbation to the system has been modelled by the removal of nodes.

When a node is removed, the flow through the network changes. Every node for which the new load exceeds the original capacity is then considered as failed and the process is repeated until there no subsequent failure. This process develops in discrete time steps starting with the removal of the test node at time $t=0$. For every time step the fraction of failed nodes is identified as:

$$X_t = \frac{n_{failed}(t)}{n} \quad (3)$$

where n is the total number of nodes. The final fraction of failed nodes is represented in this work as X_∞ . This has been used as a measure of robustness in the literature on complex networks (Havlin et al. 2010). If the probability of the removal of a network node is assumed to be the same for every node, then X_∞ is a measure of the risk associated with the scenario involving the removal of that node.

4.3. Vulnerability metrics

In this paper, a new vulnerability metric, identified here as X_1 , is defined. It is the fraction of nodes failed after the first step ($t=1$) of the cascading process, i.e.

$$X_1 = \frac{n_{failed}(1)}{n} \quad (4)$$

The computation of this metric requires to run the flow model once for the undisturbed network and once for every disruption. By avoiding the full cascading failure simulation, it provides a mean to balance result accuracy with computational effort. Network analysis itself is regarded as the first step of a more complete risk assessment process (Eusgeld et al. 2009), and therefore it is important to achieve this balance.

Vulnerability metric X_1 was used to rank every node in the case study and its performance was compared with other commonly used metrics in infrastructure networks research. Nodal degree (D), original flow through the node (F) and change in the efficiency of the network (E), computed as the change in the mean of the means of the shortest path lengths when the node is removed (Crucitti et al. 2004), are amongst the less computationally intensive metrics frequently used in the research literature.

In this paper a complete assessment on the predictive capabilities of each of these metrics is presented.

The cascading failure model described above (Equations 1 to 3) was run on the two different systems described in Section 5 and rankings of the nodes according to X_∞ were obtained. This ranking was taken as the reference. Subsequently, each of the rankings according to D , F , E and X_1 was compared to the reference by using the Spearman's rank correlation coefficient (ρ):

$$\rho = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2-1)} \quad (5)$$

where n is the number of elements in the two rankings and d_i the difference between the rankings of each element according to the two criteria. The value of ρ can vary between -1 and 1.

4.4. Sensitivity analysis

In order to show the robustness of X_1 for the identification of important nodes in the network, a sensitivity analysis was performed. The effects on the predictive capability of D , F , E and X_1 were investigated after changing network topology, average link density, capacity distribution, flow model and network size. One thousand networks with different parameters were generated during each step of the analysis. The full cascading failure process was run on every sample network. Its nodes were ranked according to X_∞ and the four predictive metrics, and finally Spearman's ρ coefficients were calculated for each of them.

5. CASE STUDY

The IEEE-RTS96 power transmission system was used as the starting point for this study. The One Area network (24 nodes) as well as the Two Area interdependent system (48 nodes) (Grigg et al. 1999) were considered. The Two Area system is obtained by joining two copies of the One Area system with three connecting links. This interdependency alters the flow through the single network, changing the cascading failure dynamics. The defining parameters of these systems were then perturbed during the sensitivity analysis to explore the performance of the methodology based on X_1 .

5.1. Network topology

The first sensitivity analysis that was performed consisted in systematically rewiring the network edges in order to create new topologies. The average link density $c = 2m/n$ was fixed to the value (2.833) of the IEEE system taken as the reference. This procedure generated networks with a fixed number of nodes and a constantly changing edge distribution.

5.2. Average link density

The second analysis consisted in sampling the average link density c from a uniform distribution $[c_{min}, c_{max}]$ where c_{min} characterizes a treelike network and c_{max} a fully connected graph. Networks with the original number of nodes and varying number of edges were generated. The line properties for the edges were sampled from a uniform distribution $[x_{min}, x_{max}]$ where x_{min} and x_{max} are the minimum and maximum line impedances in the original network.

5.3. Element capacity distribution

During the previous analyses, capacity was fixed for every node at 120% of the initial load. The third sensitivity analysis sampled the capacity of each node from a uniform distribution. The extremes of the distributions are 1.05 and 2 times the original load, reflecting the heterogeneity typical of real networks.

5.4. Flow model

The ML model described above can produce cascading failure results which are similar to higher-fidelity engineering model. It belongs, however, to the class of topology-based model and makes only minimal use of the engineering information on the system. The effects on X_I of using a Direct Current (DC) Power Flow model were investigated by solving the DC power flow equations (Pepyne 2007) on one thousand different network topologies:

$$F = CA(B)^{-1} P \quad (6)$$

where F is the vector of power flow in the lines, C is the line properties matrix, A is the edge-node incidence matrix, B is the bus susceptance matrix and P is the vector of power injections at nodes.

For the purpose of determining the power injections, every generator was assumed to contribute to satisfying the demand on the system with the same proportion of its maximum generating capacity.

5.5. Size of the network

Finally, a sensitivity analysis was performed on the size of the network. In this case the topology of the Two Area network was used as the starting point. On this double size system the rewiring described in Section 5.1 was performed, with c equal to 2.958.

6. RESULTS AND DISCUSSION

6.1. Analysis of the IEEE Systems

The correlation (Equation 5) between each of the four rankings (D , F , E and X_I) and the reference ranking (X_∞) is given in Table 1 for the two original networks.

Table 1: Spearman's ρ , ML flow model

Case	D	F	E	X_I
24 nodes	0.526	0.674	0.733	0.937
48 nodes	0.374	0.634	0.660	0.818

A drop in the predictive capabilities of the degree of the node is observed here, suggesting that the dynamics introduced by the interdependency links cannot be captured by such a simple metric.

In both cases X_I proves to be the metric that is most successful in identifying the nodes that cause the largest cascading failures. While it may seem natural that the correlation is the highest among X_I and X_∞ , research papers on complex systems such as critical infrastructures often stress their inherent unpredictability (Zio 2014). In theory, the bulk of the damage to the network may happen at any stage of the process, and could evolve with drastically different dynamics. Here, however, it is shown that such unpredictability is bounded, and that X_I was able to reliably identify the system vulnerabilities without a full dynamic simulation during a wide array of simulations.

6.2. Sensitivity analysis results

The value of Spearman's ρ was calculated for each metric at the end of every simulation in each of the five sensitivity analyses. The results of the first 1000 simulations (random topologies with 24 nodes and fixed link density) are plotted in Figure 1 and Figure 2. Similar diagrams can be plotted for every sensitivity analysis. Instead, the mean values of ρ and the respective coefficient of variation $\Delta\rho/\bar{\rho}$ are given in Table 2 and Table 3 for all the cases described above. The results show that X_I performs much better than the other metrics, with the sample mean of the Spearman's rank correlation coefficient ($\bar{\rho}$) being the highest among the four metrics.

Table 2: $\bar{\rho}$, sensitivity analyses

Case	D	F	E	X_I
Topology(24)	0.653	0.697	0.623	0.953
Link density	0.476	0.650	0.613	0.962
Capacity	0.449	0.701	0.635	0.953
Flow model	0.401	0.603	0.238	0.747
Topology(48)	0.670	0.746	0.674	0.940

The coefficient of variation of each sample is the smallest in the case of X_I , indicating a more robust performance. The distribution of the results is also skewed to the right.

Table 3: $\Delta\rho/\bar{\rho}$, sensitivity analyses

Case	D	F	E	X_I
Topology(24)	0.166	0.193	0.221	0.047
Link density	0.416	0.217	0.221	0.040
Capacity	0.215	0.181	0.094	0.043
Flow model	0.482	0.290	0.875	0.231
Topology(48)	0.119	0.105	0.131	0.043

Values of ρ that allow to reject the null hypothesis of no correlation with a 5% level of significance depend on the number of elements of the ranking. These are: $\rho_{s,5} = 0.344$ for $n = 24$ and $\rho_{s,5} = 0.240$ for $n = 48$. This suggests that each metric, to some extent, is able to identify criticalities. The mean values of their correlations, however, show considerable variation between different analyses.

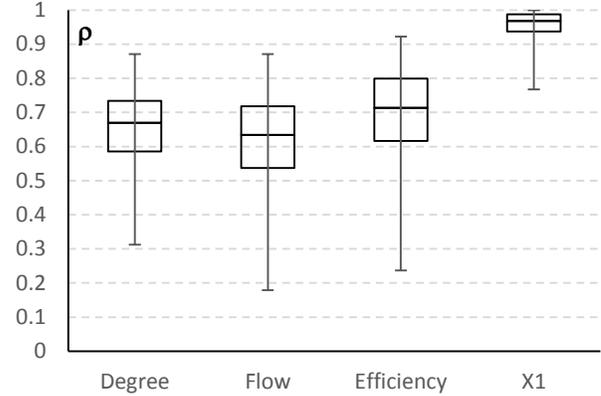


Figure 1 - Boxplot of the distributions of the correlation coefficient between X_∞ and D, F E and X_I (random topologies, 24 nodes $c=2.583$).

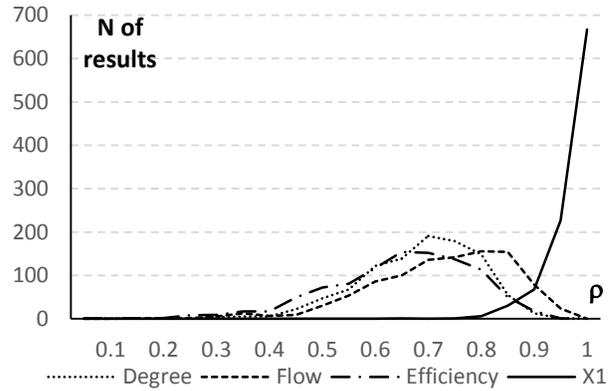


Figure 2 - Distributions of ρ between X_∞ and D, F E and X_I (random topologies, 24 nodes $c=2.583$).

The mean correlation between X_∞ and node degree (D) or change in efficiency (E) drops respectively by 38% and 62% when using the DC flow model, suggesting that these two metrics are the less robust in their predictive capabilities. The original flow (F) is comparatively stable in the samples considered in the analysis, with the mean value of ρ changing by at most 15% when using the DC flow model. This proves once again the effectiveness of the ML model in simulating the flows through the network.

X_I consistently outperforms the traditional metrics: when the ML flow model is used, the ranking based on X_I has a mean Spearman correlation of over 94% on 4000 different networks with varying topology, size, average link density and element capacity distribution.

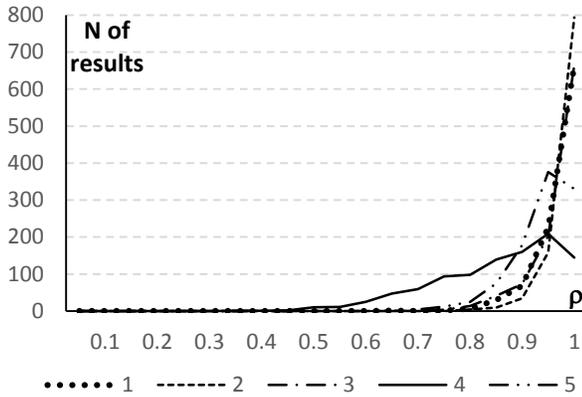


Figure 3 - Distributions of ρ between X_1 and X_∞ for the five sensitivity analyses performed: 1.topology (24 nodes), 2.link density, 3.capacity, 4.flow model, 5.topology (48 nodes).

The distributions of ρ between X_1 and X_∞ obtained in the five sensitivity analyses are shown in Figure 3. The test that puts the highest amount of strain on the predictive capabilities of X_1 is the change in the flow model from topological (the ML model) to functional (the DC model). Even in this case, however, X_1 performs better than all the other indicators with a sample mean of 0.747 and coefficient of variation equal to 0.231.

6.3. High risk scenarios

The sum of the X_∞ values obtained by removing the nodes in the top 20% of the ranking (according to each metric) was taken as an aggregate measure of the risk in those scenarios. The ratio of this number to the sum of X_∞ in the same number of critical nodes identified by the full cascading failure model is indicated as $R20$. While Spearman's ρ correlates pair of metrics across the whole ranking, $R20$ explores how the most severe scenarios identified with the simplified metrics correlate with the result of the full simulation.

In other words, if the decision maker was to act on the top 20% elements identified with the simplified assessments, how much of the risk would be mitigated? The 20% mark was chosen because those few elements represent over 70% of the aggregated cascading failure risk. Table 4 and Table 5 present the results of such assessment for the two different flow models.

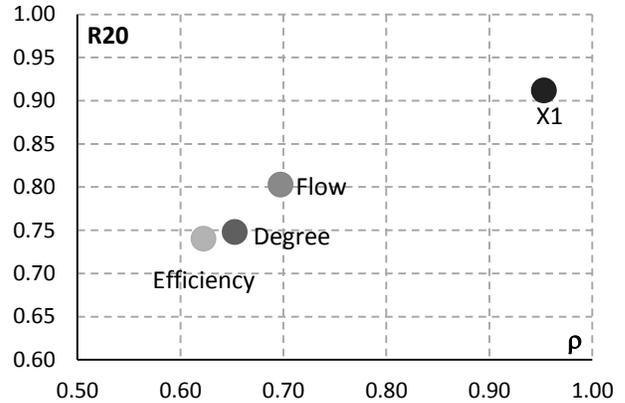


Figure 4 - Evaluation of each metric on ρ and $R5$ (ML flow model, random topologies, 24 nodes)

Table 4: $\overline{R20}$, different flow models

Case	D	F	E	X_1
ML model	0.749	0.803	0.740	0.911
DC model	0.710	0.790	0.620	0.854

Table 5:

Case	D	F	E	X_1
ML model	0.166	0.193	0.221	0.047
DC model	0.198	0.146	0.268	0.099

Figure 4 shows, for the ML model, how the four indicators perform when evaluated simultaneously against ρ and $R20$. Metric X_1 is again found to outperform the other metrics.

7. CONCLUSION

(i) Managing infrastructure systems requires to explicitly take into account their interconnectivity and the risks it carries. (ii) A new metric for vulnerability assessment, X_1 , was introduced. It represents the fraction of nodes failed at the first step of the cascading failure process. The method contrasts with previous research, where computationally intensive dynamical models were used. Identification of system vulnerabilities performed with X_1 are very similar to the results of a full cascading failure simulation. (iii) A systematic evaluation of the metric X_1 against different metrics including degree, original flow and change in network efficiency, clearly demonstrates the merits of X_1 . (iv) For a system with a large number of nodes, this metric has the

advantage of having a fixed computational time, as it does not depend on the dynamics of the cascading failure process. Since the number of such computations may be high in infrastructure systems, the advantage is significant. (v) X_I is a good predictor of X_∞ under a wide range of conditions. It can be used in the formulation of new analytical models of cascading failures of complex networks.

8. REFERENCES

- Agarwal, J., Blockley, D.I. & Woodman, N.J., 2001. Vulnerability of Systems. *Civil Engineering and Environmental Systems*, 18(2), pp.141–165.
- Albert, R., Jeong, H. & Barabasi, A., 2000. Error and attack tolerance of complex networks. *Nature*, 406(6794), pp.378–82.
- Anderson, P.L. & Geckil, I.K., 2003. Northeast blackout likely to reduce US earnings by \$6.4 billion. *Anderson Economic Group*
- Arup & Rockefeller Foundation, 2014. City Resilience Framework. Available at: <http://www.rockefellerfoundation.org/blog/framework-articulating-city-resilience>
- Blockley, D., Godfrey, P. & Agarwal, J., 2012. Infrastructure resilience for high-impact low-chance risks. *Proceedings of the Institution of Civil Engineers*, 165(Civil Engineering Special Issue), pp.13–19.
- Crucitti, P., Latora, V. & Marchiori, M., 2004. Model for cascading failures in complex networks. *Physical Review E*, 69(4): 045104.
- Dueñas-Osorio, L., Craig, J.I. & Goodno, B.J., 2007. Seismic response of critical interdependent networks. *Earthquake Engineering and Structural Dynamics*, 36, pp.285–306.
- Dunn, S. & Wilkinson, S.M., 2013. Identifying Critical Components in Infrastructure Networks Using Network Topology. *Journal of Infrastructure Systems*, 19, pp.157–165.
- Eusgeld, I., Kröger, W., Sansavini, G. & Schläpfer, M., 2009. The role of network theory and object-oriented modeling within a framework for the vulnerability analysis of critical infrastructures. *Reliability Engineering & System Safety*, 94(5), pp.954–963.
- Grigg, C. et al., 1999. The IEEE reliability test system - 1996. *IEEE Transactions on Power Systems*, 14(3), pp.1010–1020.
- Havlin, S. et al., 2010. Catastrophic Cascade of Failures in Interdependent Networks. *Nature*, 464(7291), pp.1025–1028.
- Johansson, J., Hassel, H. & Zio, E., 2013. Reliability and vulnerability analyses of critical infrastructures: Comparing two approaches in the context of power systems. *Reliability Engineering & System Safety*, 120, pp.27–38.
- Lorenz, J., Battiston, S. & Schweitzer, F., 2009. Systemic risk in a unifying framework for cascading processes on networks. *The European Physical Journal B*, 71(4), pp.441–460.
- Motter, A. & Lai, Y.-C., 2002. Cascade-based attacks on complex networks. *Physical Review E*, 66(6): 065102.
- Ouyang, M., 2014. Review on modeling and simulation of interdependent critical infrastructure systems. *Reliability Engineering and systems safety*, 121, pp.43–60.
- Pepyne, D.L., 2007. Topology and cascading line outages in power grids. *Journal of Systems Science and Systems Engineering*, 16(2), pp.202–221.
- Quarantelli, E.L., 1999. The disaster recovery process: what we know and do not know from research, *University of Delaware Disaster Research Center Preliminary Papers*. Available at: <http://dspace.udel.edu/handle/19716/309>
- Taleb, N.N., Bar-Yam, Y., Douady, R., Norman, J., Read, R., 2014. The Precautionary Principle: Fragility and Black Swans from Policy Actions, *NYU School of Engineering Working Paper Series* (2014). Available at: <http://arxiv.org/pdf/1410.5787v1.pdf>
- UNISDR, 2009. *UNISDR terminology on Disaster Risk Reduction*. Available at: <http://www.unisdr.org/we/inform/terminology>.
- Wilkinson, S.M., Dunn, S. & Ma, S., 2011. The vulnerability of the European air traffic network to spatial hazards. *Natural Hazards*, 60(3), pp.1027–1036.
- Zio, E., 2014. Vulnerability and Risk Analysis of Critical Infrastructures. *Vulnerability, Uncertainty, and Risk*, ASCE, pp.23–30.
- Zio, E. & Sansavini, G., 2010. Modeling failure cascades in critical infrastructures with physically-characterized components and interdependencies. *ESREL 2010 Annual Conference*, pp.651–652.