# Identifying Cascading Failures on Synthetic Power Transmission Systems

Ryan M Sparks • Ryan Hoff • Dr. Nathan Johnson • Dr. Mikhail Chester

## Abstract

Cascading failures across a network propagate localized issues to more broad and potentially unexpected failures in the network. In power networks, where load must be delivered in real-time by a generation source, network layout is an important part of cascading failure analysis. In lieu of real power network data protected for security reasons, we can use synthetic networks for academic purposes in developing a validating methodology. A contingency analysis technique is used to identify cascading failures, and this involves randomly selecting initial failure points in the network and observing how current violations propagate across the network. This process is repeated many times to understand the breadth of potential failures that may occur, and the observed trends in failure propagation are analyzed and compared to generate recommendations to prevent and adapt to failure. Emphasis is placed on power transmission networks where failures can be more catastrophic.

## Introduction

Power transmission systems are critical infrastructure that require significant planning, maintenance, and upkeep, and are highly vulnerable to environmental hazards. Failure at some point in a system's lifetime is inevitable, and power grids are designed with redundancy to protect the system and continue to deliver power even in cases of component failure. Even so, there are some scenarios wherein the loss of critical components results in a power outage.

While localized power outages on distribution systems are somewhat common and can often be repaired quickly, unexpected failures on the transmission system can be significant and involve the disturbance of large amounts of power. [1] Studying real transmission systems is often difficult because of the lack of available data, largely in part due to infrastructure security concerns [2]. To construct realistic networks and make meaningful conclusions about situations without using real data, fictional but realistic (i.e., synthetic) networks are used, possessing realistic power system attributes and topology, but which do not reveal sensitive information about their real infrastructure counterparts [3].

## Methodology

There are several approaches for identifying cascading failures which examine different causes and results of cascading failure. [1] For this study, we will examine the vulnerability of sub-transmission level substations to single and multiple thermal overloads. The identification of cascading failure scenarios can be accomplished by performing a contingency analysis on the system, and repeatedly removing any lines that are thermally overloaded by the contingency. When the resulting network has all lines under their thermal overload threshold, any sub-transmission load buses that have become isolated from the system are identified, and the outage is recorded along with the initial lines that failed in the contingency analysis.

**Synthetic Network Input**

The synthetic power network for the Phoenix city area shown in figure 1 is constructed using systematically placed substations followed by an iterative line synthesis process described in Birchfield, Overbye. [8] This can be effectively divided into three parts: substation siting and sizing, transmission line construction, and reactive power planning. The first step in the process is *substation siting and sizing*, involving the identification of geographic location, load values, generation capacity, and voltage levels. From there, *transmission line construction* identifies the best line connectivity for the network, prioritizing lines that best reduce overloads from critical contingencies with penalties for line length and poor network connectivity. Finally, *reactive power planning* adds capacitor banks to ensure power quality throughout the network. [8] Due to the small size of the studied networks and the assumption of general compliance to the power factor policy of the utility, we can assume that any power factor correction equipment resulting from reactive power planning will be negligible in terms of cascading failure analysis.

**Step 1: Substation Planning:** Substation planning begins with geographic zip code population data from the 2020 US Census. [4] The process starts by heuristically identifying a target number of load substations to be created (dividing or combining zip codes as necessary), then identifying substation geographic locations as the centroid of each zip code region. [14] To meet the target number of substations, zip codes may be divided or clustered together geographically. The load applied to each substation is determined by the population in the substation's service region. Gegner et. al. [14] finds a strong correlation between population and power delivered, with a typical value of 2.007kW and 0.573kVAR per person. Generation substations are added based on location from the public Department of Homeland Security power plants dataset [5], with generation capacity scaled to total system load. Voltages are assigned to each substation with all buses containing a sub-transmission (base) voltage bus, and higher voltages are assigned randomly with probabilities proportional to generation and load. [6]

**Step 2: Transmission Line Planning:** Transmission line planning is an iterative process that begins by identifying all possible lines that can be constructed in the network. Previous work by Birchfield et. al. [7] guides the construction of realistic synthetic transmission lines, which are contained in the Delaunay triangulation 1st, 2nd, and 3rd neighbors (for each voltage level). This same work finds that realistic power networks typically contain between 1.15 and 1.25 times as many lines as buses (for each voltage level). [7] A random subset of 1.2n lines (where n is the number of buses at that voltage level) are selected, and lines are iteratively removed at random and added back based on the line's sensitivity towards correcting the most severe contingency in the network, connectivity, redundancy, and line length. Through "random removal followed by targeted addition" [8], this process aims to "anneal" the network into a synthetic network suitable for our purposes.

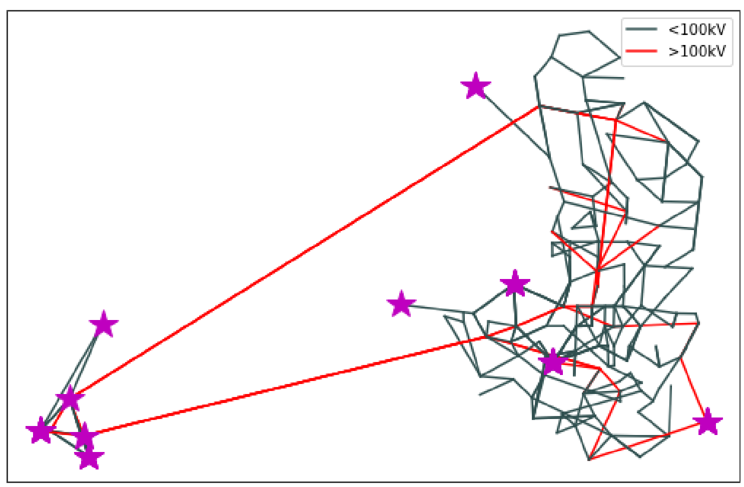

*Figure 1: Sample Input Network based on the Phoenix city area (stars represent generation)*

**Step 3: Reactive Power Planning:** Omitted from this study due to the small size of the studied networks and the assumption of general compliance to the power factor policy of the utility.

The resultant network includes static thermal line rating data depending on the needs of the system. The constructed network for the Phoenix city area is shown in figure 1. It should be noted that the thermal line rating data used here and in real transmission systems may neglect climate data, which could impact line ampacity. [15][17] The ampacity determined by the synthetic network planning process will be used as a threshold beyond which the line will be automatically opened to protect the system.

**Cascading Failure Identification**

Contingency analysis on the synthetic power network is done through the removal of one or more transmission lines, representing a failure and protection relay response on that line. Upon introducing the failure, a new power flow analysis is run; any resulting thermal line overloads are identified. These overloaded lines are removed, representing automatic thermal overload protection on the line. Once the power network is stable (or has completely collapsed), any substations which have become isolated from the network are marked as outaged. Notably, substations connected to overloaded lines may not become outaged if they are still connected to the network by other lines. This process is summarized in figure 2.

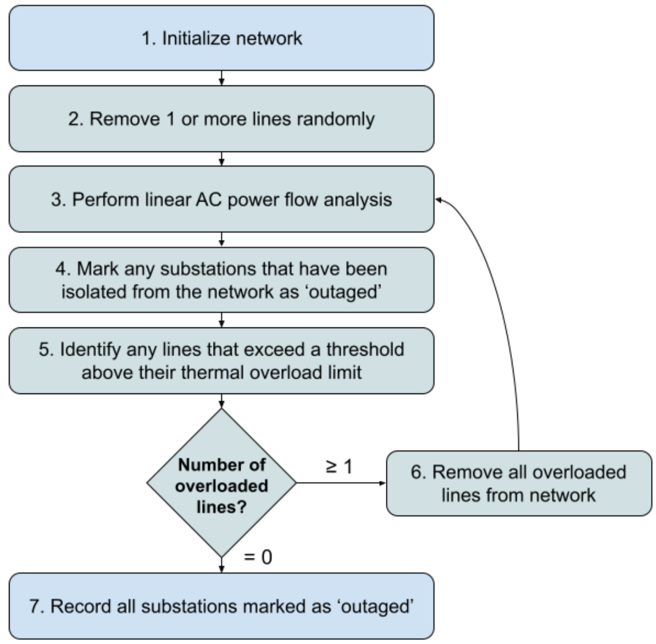


*Figure 2: Transmission network cascading failure methodology*

The outage starting point can consist of one or more of any of the lines randomly and arbitrarily selected for failure. To simulate a failure on the lines, we will assume that protection circuit breakers on the line are able to identify the failure and isolate it from the system. [16] An open-source framework in Python known as PyPSA has been used for solving power flow problems. [12] One example of a synthetic PyPSA model that has received extensive review and testing is a model of the European power system known as PyPSA-Eur. [13] After the initial failure(s) have been applied to the network, PyPSA will be used to run a power flow analysis, revealing any substations that have become completely isolated from the original network. These substations will be marked as 'outaged', with no pathway to deliver power to their customers.

To identify the possibility of any cascading failures, it can be assumed that each line is typical in that it possesses a static thermal line rating appropriately sized for its conditions. [9] The power flow through each line after the initial contingency analysis will be identified and compared with that line's thermal MVA limit, which is included in the imported synthetic network [6]. Because transmission lines may be safely overloaded for short periods of time, a permissible overload timeframe must be selected for the study. This permissible overload timeframe is correlated with how high above the limit the line is, and after this time-frame circuit breakers will open the line to prevent physical infrastructure damage. [10] In this study, we will look at three scenarios: 1-minute, 5-minute, and 10-minute overload scenarios, giving grid operators and automatic control systems some amount of time to correct the overload before automatic protection is presumably activated. These failure scenarios correspond with 245%, 135%, and 115% overload thresholds according to Carneiro and Ferrarini. [11].

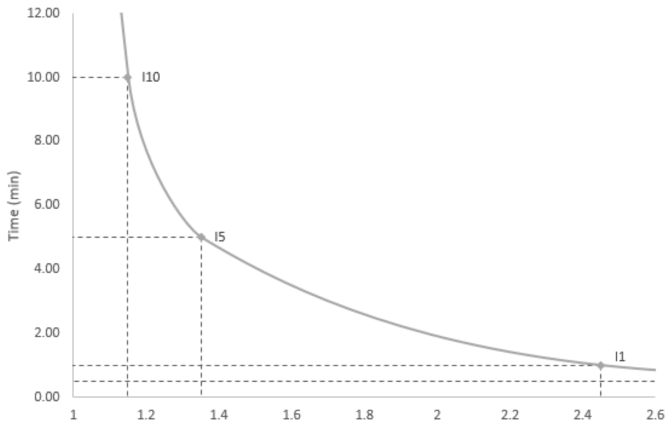

*Figure 3: Recreated from Figure 1 in [11], "Time-dependent characteristic of the classical overload protection."*

To identify cascading failures, we will systematically remove any lines that are overloaded past the specified threshold, under the assumption that they would be automatically isolated by thermal protective systems in the network. A subsequent linear power flow is run on the model using PyPSA. Any additional substations that have become completely isolated from the original network are again marked as 'outaged', and the process is repeated for any additional overloaded lines.

The process terminates when all lines in the network are under their specified overload threshold, and any number of substations may be disconnected from the network (outaged). The process returns outaged substations and which lines have been overloaded and removed from the network. To capture substations that may be more susceptible to outage, the process is repeated many times with a different 'starting point' selected each time.

**Repetition**

By repeating the process of randomly creating a failure scenario and then identifying subsequently outaged substations, information can be collected regarding outage frequency on each substation. Each scenario is identical other than the initially failed lines, so simulations can be run in parallel. For testing and validation, 10,000 iterations are used, yielding data on which substations fail and what the failure "start-point" was.

# Results

Using a synthetic transmission network in the Phoenix city area and corresponding thermal line limits on the network, the cascading failure identification process can be performed with 10,000 iterations in each scenario. These scenarios included all combinations of 1, 2, 3, and 4 simultaneous initial line failures, and 115% (10-minute), 135% (5-minute), and 245% (1-minute) overload thresholds.

The results are compiled in a list of scenarios including the initially failed lines and any subsequent failures that occurred. The process returns any buses that may have been isolated from the system, but we are primarily interested in the lower (sub-transmission) substation outages that occur (at the 69kV level in Phoenix) since distribution systems are connected to these buses.

After running the cascading failure identification process for each scenario, the results can be analyzed. In many contingency scenarios, no service regions were found to experience an outage and the system continued to operate as normal, which is to be expected from a realistic system. However, depending on the initial conditions, some service regions were found to experience outages more frequently than others.

It is immediately evident that one of the low-volume, solar generating stations close to the city fails in every scenario with at least one substation outage, however this substation does not have a load attached and is not represented by one of the service regions. Other outaged substations vary significantly depending on which initial lines were removed in the contingency scenario.

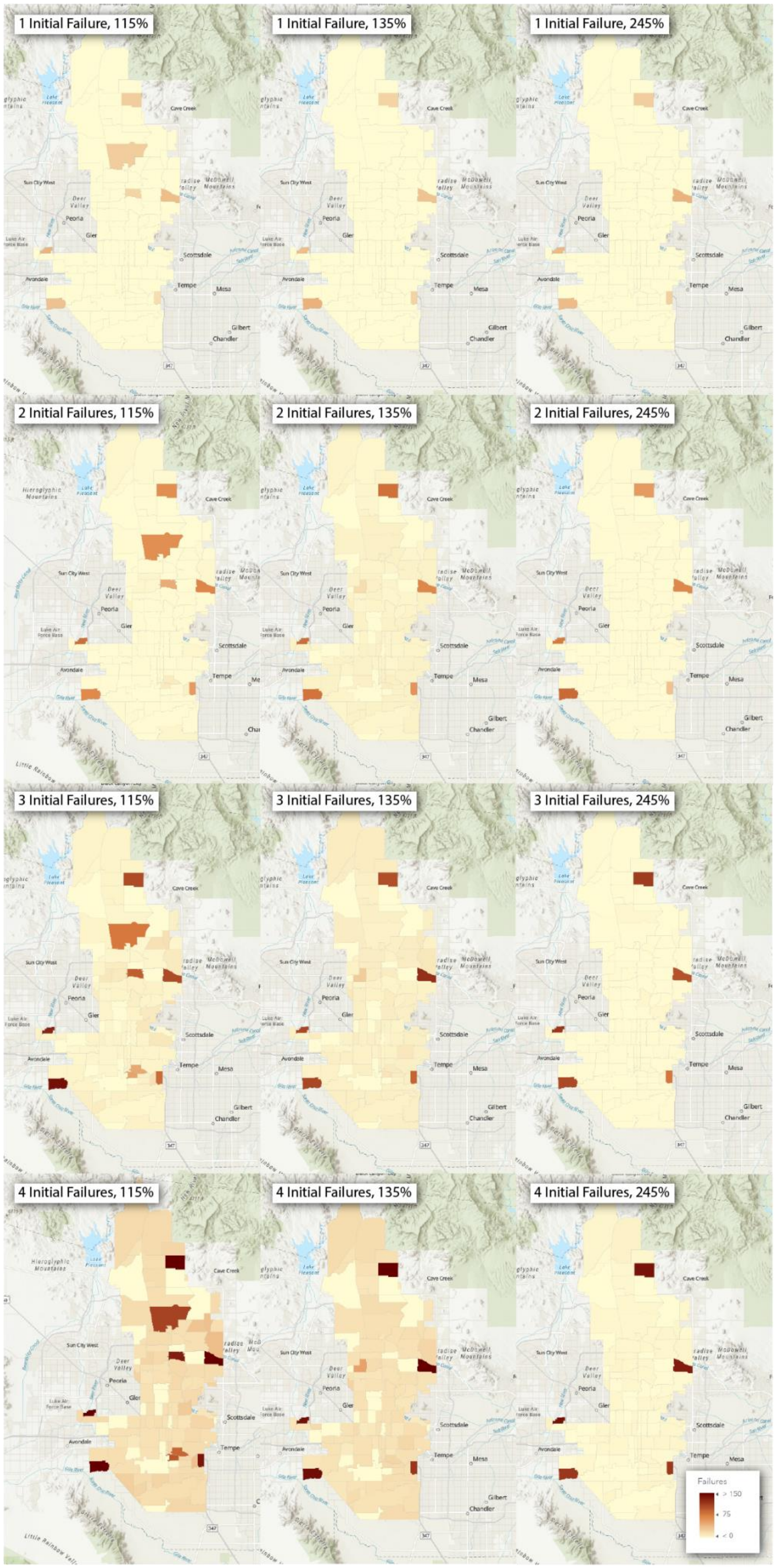


*Figure 4: Number of outage scenarios in each service region after 10,000 iterations for different combinations of overload thresholds and different numbers of initially failed lines*

It is intuitive that with more lines initially failing, there will be more outages experienced by both initial and cascading failures. This effect can be seen by examining the difference in scenarios when the overload threshold is held constant, as can be seen in figure 4. Similarly, a higher timeframe corresponds with a lower overload threshold, and it can be expected that more substations will fail when a lower overload threshold is specified. This can also be observed in figure 4. Service regions are shown on the map, with color representing the number of outages experienced in the 10,000 different scenarios.

These results are consistent with the expectations and limitations of the scenario. While this process can easily identify outaged substations that have been isolated from the network, it ignores any voltage and frequency problems that may be present on the network. In addition, the studied regions are small enough that reactive power issues on the grid are ignored. Understanding the limitations, this process can provide a reasonably useful model for identifying potential cascading failures as a result of thermal overloads on transmission lines.

# Discussion

The results of this process identify several substations in the synthetic transmission network which are more prone to outages than others. The process is contingent on a networked system and would not provide any meaningful results on radial systems such as common distribution feeders. It is also important to note that this study is performed on a synthetic network, and it cannot be assumed that there is any geographical significance to the results. However, this cascading failure methodology could be applied to a realistic system if such data was available, or it could be used to help validate a synthetic system against real scenarios. Even with the synthetic data, information is revealed about the situation leading up to the cascading failure, line connectivity, and outage frequency in response to the line failures.

Understanding the limitations of this model, it can be used to study the propensity of cascading failures to occur on a given input power transmission network. Any input network including buses, lines, substations, transformers, and generators can be used for the cascading failure process, resulting in the identification of the most vulnerable substations in a contingency scenario. By coupling this process with failure identification in other infrastructures, this model can also be used in-part to identify critical interdependencies and vulnerabilities between other systems and power failures.

# References


[1] Guo, H., Zheng, C., Iu, H. H. C., & Fernando, T. (2017). A critical review of cascading failure analysis and modeling of power system. In Renewable and Sustainable Energy Reviews (Vol. 80, pp. 9–22). Elsevier Ltd. https://doi.org/10.1016/j.rser.2017.05.206

[2] Ouyang, M. (2014). Review on modeling and simulation of interdependent critical infrastructure systems. In Reliability Engineering and System Safety (Vol. 121, pp. 43–60). Elsevier Ltd. https://doi.org/10.1016/j.ress.2013.06.040

[3] Marcos, F. E. P., Domingo, C. M., Román, T. G. S., Palmintier, B., Hodge, B. M., Krishnan, V., de Cuadra García, F., & Mather, B. (2017). A review of power distribution test feeders in the United States and the need for synthetic representative networks. In Energies (Vol. 10, Issue 11). MDPI AG. https://doi.org/10.3390/en10111896

[4] U.S. Census Bureau. Census 2020, Zip Code Tabulation Areas Gazetteer File. https://www2.census.gov/geo/docs/maps-data/data/gazetteer/2020_Gazetteer/2020_Gaz_zcta_national.zip

[5] U.S. Department of Homeland Security. Homeland Infrastructure Foundation-Level Data (HIFLD) Power Plants shapefile. https://hifld-geoplatform.opendata.arcgis.com/datasets/power-plants-2/explore

[6] Birchfield, A. B., Xu, T., Shetye, K. S., & Overbye, T. J. (2018). Building synthetic power transmission networks of many voltage levels, spanning multiple areas. Proceedings of the Annual Hawaii International Conference on System Sciences, 2018-January, 2766–2774. https://doi.org/10.24251/hicss.2018.349

[7] Birchfield, A. B., Schweitzer, E., Athari, M. H., Xu, T., Overbye, T. J., Scaglione, A., & Wang, Z. (2017). A metric-based validation process to assess the realism of synthetic power grids. Energies, 10(8). https://doi.org/10.3390/en10081233

[8] Birchfield, A. B., & Overbye, T. J. (2020). Planning Sensitivities for Building Contingency Robustness and Graph Properties into Large Synthetic Grids. 53rd Hawaii International Conference on System Sciences.

[9] Heckenbergerová, J., Musilek, P., & Filimonenkov, K. (2013). Quantification of gains and risks of static thermal rating based on typical meteorological year. International Journal of Electrical Power and Energy Systems, 44(1), 227–235. https://doi.org/10.1016/j.ijepes.2012.07.005

[10] Carneiro, J. S. A., & Ferrarini, L. (2011). A probabilistic protection against thermal overloads of transmission lines. Electric Power Systems Research, 81(10), 1874–1880. https://doi.org/10.1016/j.epsr.2011.05.011

[11] J. S. A. Carneiro and L. Ferrarini, "Analysis and Design of Overload Protections for HV Lines with a Probabilistic Approach," Proceedings of the 10th International Conference on Probablistic Methods Applied to Power Systems, Rincon, PR, USA, 2008, pp. 1-6.

[12] T. Brown, J. Hörsch, D. Schlachtberger, PyPSA: Python for Power System Analysis, 2018, Journal of Open Research Software, 6(1), arXiv:1707.09913, DOI:10.5334/jors.188

[13] Jonas Hörsch, Fabian Hofmann, David Schlachtberger, and Tom Brown. PyPSA-Eur: An open optimisation model of the European transmission system. Energy Strategy Reviews, 22:207-215, 2018. arXiv:1806.01613, doi:10.1016/j.esr.2018.08.012.

[14] Gegner, K. M., Birchfield, A. B., Xu, T., Shetye, K. S., & Overbye, T. J. (2016). A methodology for the creation of geographically realistic synthetic power flow models. 2016 IEEE Power and Energy Conference at Illinois, PECI 2016, 1–6. https://doi.org/10.1109/PECI.2016.7459256

[15] Jin Lu, Xingpeng Li, Hongyi Li, Taher Chegini, Carlos Gamarra, Y. C. Ethan Yang, Margaret Cook, and Gavin Dillingham, "A Synthetic Texas Backbone Power System with Climate-Dependent Spatio-Temporal Correlated Profiles", arXiv, Feb. 2023.

[16] Kezunovic, M. (2005). Fundamentals of Power System Protection. In The Electrical Engineering Handbook. Chapter 9. Elsevier Inc. https://doi.org/10.1016/B978-012170960-0/50058-X

[17] M. Bartos et al., "Impacts of rising air temperatures on electric transmission ampacity and peak electricity load in the United States," Environ. Res. Lett., vol. 11, no. 11, 2016, doi: 10.1088/1748-9326/11/11/114008.