This article has been accepted for publication in IEEE Transactions on Power Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TPWRS.2023.3260119

Ranking the Impact of Interdependencies on Power System Resilience using Stratified Sampling of Utility Data

Molly Rose Kelly-Gorham, Paul D.H. Hines, Ian Dobson

***Abstract*—It is well known that interdependence between electric power systems and other infrastructures can impact energy reliability and resilience, but it is less clear which particular interactions have the most impact. There is a need for methods that can rank the relative importance of these interdependencies. This paper describes a new tool for measuring resilience and ranking interactions. This tool, known as Computing Resilience of Infrastructure Simulation Platform (CRISP), samples from historical utility data to avoid many of the assumptions required for simulation-based approaches to resilience quantification. This paper applies CRISP to rank the relative importance of four types of interdependence (natural gas supply, communication systems, nuclear generation recovery, and a generic restoration delay) in two test cases: the IEEE 39-bus test case and a 6394-bus model of the New England/New York power grid. The results confirm industry studies suggesting that a loss of the natural gas system is the most severe specific interdependence faced by this region.**

***Index Terms*—Resilience, Critical Infrastructure, Interdependence**

I. Introduction

THERE are many known interdependencies between bulk power systems and other critical infrastructure (CI) systems that impact energy resilience. But which interdependencies are most important? To the authors' knowledge, this question remains unanswered (and mostly unasked). This paper presents a method to rank the effects of several interdependencies on the resilience of a power transmission system based on data collected by electric utilities.

A. Motivation

Extended, large power outages come with very high social costs. It is well known, but not particularly well quantified, that these costs are exacerbated by interdependence between CI systems. Communication towers, traffic lights, water filtration and supply systems, and even natural gas compressors, all of which become critical during times of emergency, need reliable electricity (to varying extents) to operate. And all of these systems are important to the operation and restoration of electric power systems. Developing policies that enhance societal resilience therefore requires that we deeply understand the interdependencies between electric power systems and other infrastructure systems. The interactions among power generation, the electricity transmission network, and other CI are often stressed even under normal operating conditions. Disturbances can easily push things over the edge and trigger cascades of failures that lead to large-scale disasters. These disturbances include earthquakes, fires, hurricanes, wind storms, trees interacting with power lines, line icing, snow storms, and flooding. While dependable electricity supply is already critical today, the ongoing electrification of energy services, such as transportation and heating, will make it even more important that we deeply understand how interdependence impacts energy resilience.

An important example of critical interdependence is natural gas and electricity. Natural gas generators need continuous access to gas through pipelines to generate electricity. Many natural gas compressor stations and extraction wells require electricity to operate [1]. During times of harsh cold conditions (cold snaps) gas demand for heating often takes priority over gas for electricity generation due to the contracts used by natural gas generators. Especially under severe cases, cold conditions can lead to load shedding as the only option for stabilizing the electric grid. The February 2021 Texas cold snap left 4.5 million customers without power. A primary cause of this event was natural gas generation failing due to cold temperatures [2].

A second example is communication systems. Communication towers are usually powered by the electric grid, often with battery systems for backup power. While the core communication systems used by the bulk power grid (SCADA) are typically separated from the public internet and cellular communication systems, and have their own battery backup systems, important interdependencies remain. During restoration of grid infrastructure, repair crews will use the public cellular networks to communicate with coordinators. Given that backup power systems for cellular towers can only operate for a limited time without access to grid power, there is an interdependence that will only arise during longer outages.

Nuclear generators are naturally interdependent on electric grids, because they require power from the grid to operate various safety systems. When power from the grid is seen by the plant as potentially unreliable (e.g., significant voltage fluctuations), nuclear plants are designed to disconnect from the grid and use local backup power systems to initiate a controlled plant shutdown. This creates an interdependence

M. Kelly-Gorham is with the Pacific Northwest National Laboratory, Richland, WA, 99352. email: mollyrose.kelly-gorham@pnnl.gov

P. Hines is with EnergyHub, Brooklyn, NY, 11217 USA. email: hines@energyhub.net

I. Dobson is with the Department of Electrical and Computer Engineering, Iowa State University, Ames, IA, 52242 USA. email: dobson@iastate.edu

Support from NSF award nos. 1735354, 1735513, and 2153163 is gratefully acknowledged.

Manuscript received [July 2022]; revised [November 2022]

that can both increase the size of a blackout by removing large generators from the supply, and the length of a blackout because once a nuclear plant initiates the shutdown process, it can be days to weeks before it is available to generate power again [3].

When CI systems have been damaged for extended time periods, restoring key elements of the power grid becomes even more challenging. This was seen in Puerto Rico after Hurricane Maria [4]. This concept, which we call *compounding risk over time*, is difficult to model precisely, but is clearly present in the history of large power grid failures, where large restoration projects become increasingly difficult as more infrastructure systems have been out for longer periods of time, leading to complicated and often unanticipated interactions.

B. Paper description

The Computing Resilience of Infrastructure Simulation Platform, CRISP, used in this paper, was initially presented in [5], [6], where it was used to evaluate transmission system resilience as a single infrastructure including renewables. In this paper, the CRISP framework is extended to evaluate transmission system resilience focusing on interactions with other infrastructures. A key feature of CRISP is its use of data collected by electric utilities about transmission (and here generation) outages to build probability distributions. These probability distributions combine the outcomes of the various processes involved in resilience, and sampling from these probability distributions can efficiently reproduce the response and resilience of the power system under stress. CRISP samples from these distributions to generate representative sets of line/generator outages and restoration times. This allows us to consider a large set of events caused by many different disturbances (e.g., fire, weather, human error) as the source of disturbance without needing to explicitly model each mechanism. CRISP then uses a grid model to track the progress of restoration and quantify the size, duration, impact, and risk of resilience events.

Conventional sampling methods such as Monte Carlo tend to under sample from extreme events, resulting in either massive computing requirements or an inaccurate estimate of the risk associated with large events [7]. To address this problem, this paper uses stratified sampling to find the low probability, high impact events in sufficient quantity to quantify the risk. Sampling a sufficient number of large events is already crucial to measuring the resilience of power systems, and when interdependencies between power systems and other CIs are included sampling methods become even more important.

This paper estimates the impact of several types of interactions and ranks them to find the effect on the resilience of two power system networks. As a first effort to incorporate interactions into the resilience quantification methods developed in [5], [6] these interaction models are appropriately simple. There are three specific interaction models: communications, natural gas supply, and nuclear restorations, as well as the more generic compounding risk over time.

C. Literature Review

The idea of measuring resilience, before its use in engineering, stems from the literature on Social-Ecological Systems, where interdependencies and the effects of different scales are seen as critical to measuring the resilience of these complex systems [8]. The field of civil engineering pioneered frameworks for resilience of infrastructure systems [9]. As reviewed in [10], the quantification of resilience of an infrastructure system requires measures of system performance under different stages of resilience events, including disaster prevention, damage propagation, and recovery. While there are many overlapping definitions of power system resilience, such as from IEEE [11], [12] and others [13], [14], the most succinct definitions are from CIGRÉ [15]: "Power system resilience is the ability to limit the extent, severity, and duration of system degradation following an extreme event" and from the US government [16] defining resilience as "the ability to anticipate, prepare for, and adapt to changing conditions and withstand, respond to, and recover rapidly from disruptions." All these definitions of resilience involve the power system's response to events or extreme events. The electrical engineering resilience literature focuses on understanding high impact low probability events that are not considered in standard reliability assessment [17]. In addition, a number of papers have shown that it is helpful to model resilience events using a sequence of stages, including initial failures, propagation, and restoration [14], [17]. CRISP analyzes all the phases of events as they dynamically evolve with special attention to sampling methods that are able to approximately quantify the most extreme events, and is therefore a method of assessing resilience.

In order to understand resilience in power systems, we need to understand the key elements of events that degrade resilience. These elements include the initial disturbance [18], which causes failures, which can then initiate cascades of outages [19], [20], which is followed by a restoration process [21]. Initiating outages are often modeled using fragility curves, with outage probabilities that are dependent on the nature of the external hazard [22]. Stratified sampling from utility data is applied in [23] to represent cascading outages in the electric transmission grid after an earthquake in order to optimize investment to harden the grid against earthquakes. Power system resilience to wind storms [24], earthquakes [9], and less frequently multiple hazards [25] are also explored in the engineering literature.

While the resilience of interdependent CI has been studied for decades [26], much of the problem remains open [27]–[29]. Several papers use multi-network models to study interdependent CI [30], [31]. Much of the work in this field studies the cascading failure between interdependent infrastructures [32], [33]. Vulnerable components considering the interdependence between CI are modeled in [34]. Others develop restoration strategies for multiple interdependent CI [30], [35]–[37]. Specific combinations of interdependent power and other CI are being investigated, for example, by finding the impact of specific events on interdependent communication and electricity infrastructure [14]. There are many ongoing challenges for modeling the resilience of these interdependent systems. For example, a review of communication systems and power systems [38] describes several mechanisms of interdependence in the case of smart grids and severe weather conditions that need further study. Many key interdependencies are increasing over time, such as transportation and electricity with the growth of electric vehicles. Indeed, benchmark problems are being formulated to encourage further research on this relationship [39]. Others are using data-based component failure models [40], and proposing creative ideas for utilizing interdependence to improve resilience [41]. Interdependence between natural gas and electric transmission networks is being studied not only in resilience settings [42], but also under normal operating conditions [43], [44] to understand interconnected energy systems that provide both heat and electricity [45]. Water and power networks are increasingly stressed by their interdependencies and are beginning to be investigated [46].

D. Contributions

This paper presents a novel method for ranking and approximately quantifying the impact of multiple interdependencies on the resilience of a bulk power system. The paper makes three primary contributions to methods for assessing the resilience of interdependent energy infrastructure:

1) The paper characterizes the outcomes of outage initiation, cascading, and restoration in the grid using the statistical patterns of historical data for transmission and generation. This method works directly from standard data already collected by utilities, and avoids many of the assumptions needed for simulation models.
2) The paper applies stratified sampling to the problem of resilience event modeling so that high impact low frequency events can be practically and efficiently sampled.
3) The paper approximately quantifies and ranks the resilience impact of interdependencies between electricity and gas, nuclear, and communications infrastructures, using data from the Northeastern US and 39-bus and 6394-bus grid models.

A ranking and approximate quantification can guide utility management, regulators, business, government, and the public towards investments that mitigate the highest-risk threats to resilience, and help avoid ineffective investments. Further, with these three innovations together, the methods in this paper can inform large-scale infrastructure investment priorities by providing decision makers with information about the relative importance of particular interdependencies that make society less resilient to various hazards.

In what follows, we present our methods in Sec. II, results in Sec. III, and conclusions in Sec. IV.

II. Methods

CRISP is a novel approach that quantifies the resilience of a power system to generator and transmission outages using a data-driven method that represents each process in resilience events from initiating disturbances through restoration. CRISP represents five specific processes: stress leading to initial failures, cascade of failures through the network, the post-disturbance degraded system state, restoration, and quantifying the impact of the event [5]. The model for each process combines real-world historical data from electric utility systems with a power system model. Many salient characteristics of historical power systems events will recur in the future. In particular, CRISP follows the *statistical patterns of previous events*, not specific historical events, and therefore CRISP generates events that have not previously happened, but are consistent with past patterns. The statistical patterns of the past responses of the grid (but not usually the detailed events) will recur to a large extent in the future. At the same time, it is the case that some new statistical patterns can emerge that are not manifested in historical data. Moreover, we do not rely on statistical patterns from previous events only; we also use a network model with an OPF to help determine the response of the grid. As a result, this method does allow us to test the impact of modifications to a system, as evidenced by the fact that we are able to test the response of our two test cases to different interdependence scenarios. After each simulation, CRISP delivers multiple measures of event impact, including event duration (hours), event size (MW), the number of components lost, and energy not served or ENS (MWh). Quantitatively describing the distribution of event duration, size, and ENS over a broad spectrum of initiating events is a valuable way to understand the resilience of a given power network.

CRISP models the combination of initiating outages and cascading outages by sampling from historical transmission line and generator outage data collected and reported by utilities. Using historical data in this way avoids the need to model cascading outages in detail, eliminating the need to make difficult assumptions about how particular cascading failures may propagate in a given power system [47], [48]. After sampling outages in this way, CRISP uses an optimal load shedding algorithm (see [5]) to calculate the size of the event in total load shed (MW). Restoration is modeled by sampling from historical data for both transmission line and generator restoration times, and then calculating the maximum load that can be served in each hourly time step until the system is fully restored, or 6 months have elapsed. Computations were performed, in part, on the Vermont Advanced Computing Center.

Fig. 1 shows a flow diagram for one event simulation in CRISP. The model begins with historical outage data and power system network data. From there the initial outages are sampled from the statistics of the historical outage data, and applied to the power network. The initial grid state for the event is found with optimal load shedding through an OPF. The same OPF is subsequently used to compute the load shed at each time step of the model in the loop shown in Fig. 1. The unavailable (outaged) components are restored once the time-step reaches the restoration time for that component. The sampled restoration times are subject to update in the interdependence model, where restoration times may be increased. When all load is restored, or the maximum time step is reached, the model exits the loop and reports the event duration (hours), initial event size (MW), and the ENS (MWh).

This article has been accepted for publication in IEEE Transactions on Power Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TPWRS.2023.3260119

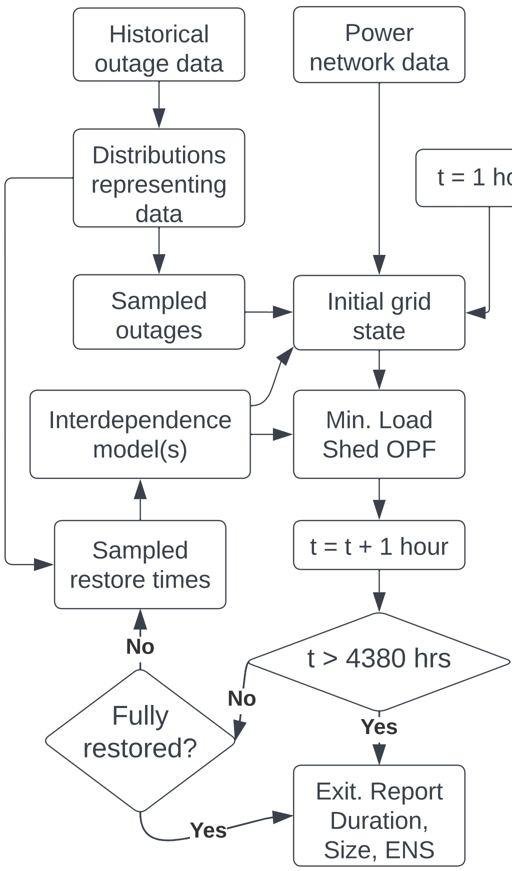

Fig. 1. Flow diagram for the CRISP model including interactions.

This section describes the data used in this paper, the statistical distributions used for line and generator outages, the stratified sampling method, and several important improvements made relative to previously published versions of CRISP. Key among the CRISP improvements is the ability to model internal and external interdependencies that impact power system resilience. This allows CRISP to quantify these interdependencies and to rank their impacts on power system resilience. We use CRISP to study the impact of four types of interactions: communications, natural gas supply, nuclear restoration, and a more abstract model of compounding risk over time. To rank interdependencies, an initial (baseline) resilience assessment of the power system is found through CRISP. CRISP then reassesses power system resilience for each interaction model using the same set of events as the baseline assessment. This allows one to compare the impacts of each interaction model on power system resilience and rank them.

A. Case studies and sources of data

This paper presents results from applying CRISP to two power system test cases based on data from the Northeastern US. The first case study uses the network data from the IEEE 39-bus case [49] covering New England in the US.[1] The network data for the second case study was developed from an Independent System Operator-New England (ISO-NE) PSS/E planning model designed to represent the 2018/19 winter-peak in Vermont and covers New England and New York in the US. We refer to this case as the 6394-bus NE-NY case.[2]

Additional generator properties, such as startup times, shutdown times, are included based on properties of the most similar generators in the 73-bus reliability test case [51]. We collected the forced outage rates (FOR) for different generator types from [52] via [53] for temperature conditions below -15°C, which range from FOR $= 1.9\%$ to FOR $= 21.2\%$. Restoration times come from Northeast Power Coordinating Council (NPCC) Generating Availability Data System (GADS) data from 2012-2015, as reported in [54].

The transmission line automatic outage data describing cascade size and spread comes from the New York Independent System Operator, NYISO, as processed from public data in [48]. The distribution of transmission line restoration times comes from automatic outages reported in the North American Electric Reliability Corporation's Transmission Availability Data System (TADS) for the US Eastern interconnection [55].

B. Number of line outages after cascading

Initiating a CRISP simulation requires the probability distribution of the total number of lines outaged after the initial outages and cascading. To build this distribution from the NYISO data, we first found the empirical probability of exactly one line outage to be $p_1 = 0.664$. Next, we found the Zipf distribution (or zeta distribution) for two or more lines out, which has the form

$$P[k \text{ lines out}] = \frac{1-p_1}{\zeta(s)-1}\frac{1}{k^s}, \quad k = 2, 3, ..., n_{\max} \qquad (1)$$

where the slope of the Zipf distribution on a log-log plot is $-s$, ζ is the Riemann zeta function, and $n_{\max} = 1695$ is the number of lines in the network. The Zipf portion of the distribution was fit by adapting the formula in [56] to the condition $k \geq 2$, giving the estimate $s = 3.59$. The purpose of using the Zipf distribution is to smooth the noisy data for large numbers of line outages with a plausible form of the distribution tail. (Note that the probability of (1) yielding more than 1695 lines out is of order 10^{-9} and is therefore neglected.)

Line restoration times from Eastern Interconnect data are fit to a log-normal probability distribution, as shown in Fig. 2. The resulting log-normal distribution with parameters $\mu = 4.77$ and $\sigma = 2.63$ is used to sample the line restoration times for both case studies.

C. Selecting which lines outage

After selecting the number of lines out after cascading, N, CRISP next chooses which specific lines go out by sampling

[1]The original IEEE 39-bus test case was updated to be $N-1$ secure by increasing line power flow limits and adding repeated lines where necessary. The modified test case has 55 lines and 10 generators and is available at [50].

[2]The original network was somewhat reduced in size by aggregating all "leaf" buses so that the resulting model includes only buses of degree 2 and higher. The resulting network contains 6394 buses. The HVDC lines were modeled as generators (positive or negative), with the power production/consumption limited between the operating point and the negative of that value. The line limits were increased so that the base case produces no flows above the line rating for each $N-1$ contingency using a DC power flow.

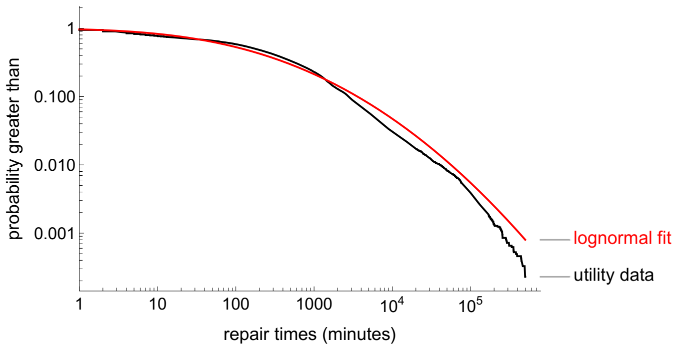

Fig. 2. Empirical distribution and log-normal fit of Eastern interconnect transmission line restoration times from TADS data.

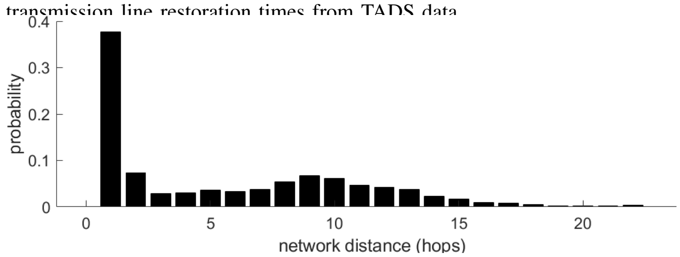

Fig. 3. Empirical distribution of network distances between all lines within a cascade using NYISO transmission line outage data.

from the empirical distribution of network distance between lines in the observed cascades from NYISO data shown in Fig. 3. In the first step, CRISP randomly samples one line in the network to be included in the cascade, L_1. At each subsequent step k, such that $2 \leq k \leq N$, CRISP samples a distance d from the empirical distribution shown in Fig. 3, and randomly selects another line L_k from the set of not-outaged lines that are a distance d from line L_1. This line is then included in the set of lines-to-go-out, k is advanced, and the process repeats until N lines are selected.

D. Generator outages

In this paper, CRISP models generator outages based on historical generator outage data under cold weather conditions. Specifically, we collected the forced outage rates (FOR) for different generator types from [52] via [53] for temperature conditions below –15°C, which range from FOR $= 1.9\%$ to FOR $= 21.2\%$. Because different extreme events will have dramatically different outcomes in terms of how many generators are out, we took the FOR data above and synthesized a broad-ranging generator outage probability mass function (PMF) to represent the number of generators out after a cascade. [3] The resulting PMF comes from an evenly weighted average of one million binomial generator outage distributions from one million evenly spaced outage rates in the range $1.9 \leq$ FOR ≤ 21.2, as shown in Fig. 4.[4] At the start of each

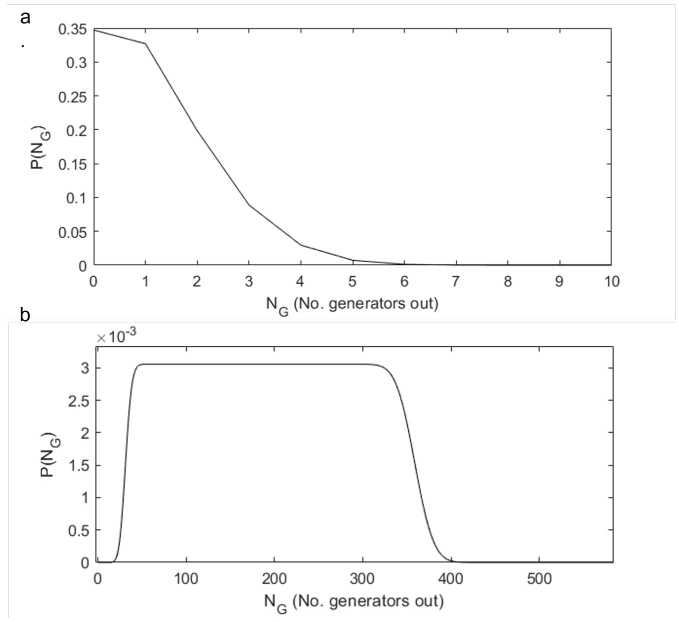

Fig. 4. Probability distributions of generator outages for (a) IEEE 39-bus case, and (b) 6394-bus NE-NY case.

simulation, CRISP samples from this PMF to determine the number of generator outages, and then chooses the particular generators that go out using a uniform distribution.

In order to build a distribution for generator restoration times, we use log-normal probability distributions fit to the NPCC GADS data for each of three types of outages: startup failures, unscheduled outages, and unscheduled deratings, which are distributed in the GADS data at 0.01%, 0.27%, and 0.72% respectively, as reported in [54]. For each generator restoration, CRISP chooses the type of outage according to the distribution above, and then samples the restoration time from the respective log-normal distribution.

E. Stratified Sampling

While large outages are rare, they contribute substantially to overall risk [57]. This comes from the fact that the distribution of outage sizes, measured by the number of lines out (1), has a heavy tail. One consequence is that straightforward, brute force Monte Carlo sampling of outage size is inherently ineffective because large cascades rarely appear in these samples. Thus, brute force Monte Carlo methods require an impractically large sample size to accurately evaluate the risk associated with large events.

To address this issue, we sample much more uniformly across the range of cascaded outage size. That is, instead of sampling directly from (1) and the generator outage distribution, we use stratified sampling [58, ch. 8] to sample more efficiently. The key idea is to divide the cascade sizes into bins or strata. We calculate the probability of each bin and then use CRISP to estimate the probability of an outcome, such as load shed of samples in each bin; that is, the load shed probability conditional on each bin. Then the load shed probability is obtained by weighting the probability of load shed conditional on each bin by the bin probabilities. This is described in more detail in the following.

[3] The PMF is based on the available generator FOR data. For generators with limited data such as wind and solar we use the same PMF.

[4] The PMF depends on the total number of generators in the case, as the total number of generators is a parameter of the binomial outage distribution. For the 39-bus system with 10 generators, Fig. 4a shows that the probability of NG generators being unavailable decreases from 0.35 for 0 generators unavailable to nearly 0 for 6 or more generators unavailable. For the larger 6394-bus NE-NY case with 1695 generators, Fig. 4b shows that the probability of 0 generators being unavailable is very small, and the probability increases until NG is approximately 50. Although the probability of an individual generator being unavailable is low, when there are many generators, the probability of a few or more generators being unavailable is high.

The cascade size is given by the number of lines out N and the number of generators out G. We choose bins $B_1, B_2, ..., B_K$ for the cascade size by specifying for each bin the range of numbers of lines out and the range of numbers of generators out:

$$B_k = \{N_k^{\min} \leq N \leq N_k^{\max} \text{ and } G_k^{\min} \leq G \leq G_k^{\max}\} \quad (2)$$

The bins partition the possible cascade sizes; that is, the bins are disjoint and they cover the full ranges of number of lines out from 1 to $n_{\max}$ and number of generators out from 0 to $g_{\max}$. The number of bins K is the product of the number of line bins and the number of generator bins.

Since the line and generator outages are assumed independent, the probability of each bin is easily computed from (1) and the generator outage distribution:[5]

$$\mathrm{P}[B_k] = \mathrm{P}[N_k^{\min} \leq N \leq N_k^{\max}]\ \mathrm{P}[G_k^{\min} \leq G \leq G_k^{\max}] \quad (3)$$

Given the number of lines out N and generators out G, CRISP can evaluate a sample of the load shed S and the ENS. Here we illustrate the process by estimating the distribution of the load shed S. One can similarly estimate ENS by substituting ENS for S. Suppose we take s samples from bin B_k and use CRISP to calculate s samples of load shed $S_1^{(k)}, S_2^{(k)}, ..., S_s^{(k)}$. Then we estimate the survival function of load shed assuming that the cascade size is in bin B_k as

$$\mathrm{P}[S > \ell \mid B_k] = \frac{1}{s} \sum_{j=1}^{s} I[S_j^{(k)} > \ell] \quad (4)$$

$I[\cdot]$ is the indicator function and the summation in (4) counts the number of samples with load shed larger than load level ℓ, where ℓ is a free variable ranging from zero to the total load. Then the survival function of load shed is given by weighting (4) estimated with CRISP with the bin probabilities calculated from (3):

$$\mathrm{P}[S > \ell] = \sum_{k=1}^{K} \mathrm{P}[S > \ell \mid B_k]\ \mathrm{P}[B_k] \quad (5)$$

Brute force Monte Carlo corresponds to stratified sampling with only one bin ($K = 1$), and one can see how having many bins over the range of cascade sizes and sampling equally from each bin distributes the sampling effort much more equally across the range of cascade sizes, since sampling from only one bin will greatly oversample the common small cascades, and greatly undersample the rare large cascades of most interest. Indeed, stratified sampling not only reduces the variance of the estimated quantities for the large cascades by increasing the number of large cascade samples, but enables much rarer and larger events to be sampled. Moreover, the bin probabilities are computed exactly from (3), which also gives lower variance estimates relative to empirical sampling from the entire distributions. To take full advantage of this, the bins are chosen so that the load shed does not vary too much over the bin [58, ch. 8.4].

[5] If the extreme event has significant correlations between transmission lines and generator failures, these are easily accommodated by computing the probability of each bin according to the joint distribution of transmission line and generator failures.

F. Interaction Models

A key to understanding the resilience of electric power systems is to understand the dependence of other CIs on the electric grid and its own dependence on those CIs. Here we describe models of several interdependencies between CI systems and the electric power grid.

1) Natural Gas Pipelines: Electric and natural gas networks are interdependent in many ways. To measure the impact of these interactions on the resilience of the power system, a probabilistic model of the natural gas interaction with the electric grid was created. Within this model, the natural gas delivery system (as a whole) will fail with a probability that is equal to the ratio of the number of lines outaged N to the total number of lines in the electric network. If the natural gas delivery system fails, then the model samples from a log-normal distribution with parameters μ and σ of 3 days and 1 day respectively to find the amount of time that the natural gas power plants will go without fuel supply.

2) Communication Infrastructure: This model represents interactions between the communication system and electric system when under stress, such as the fact that electric utility repair crews may not be able to use cellular systems during the restoration process if cell tower electricity supplies go down and their batteries run out. The model assumes each bus with load has a communication tower associated with it and each communication tower has a back-up battery. The model assumes that each tower's back-up battery has a duration drawn from a uniform probability distribution that ranges from 4 to 24 hours. CRISP samples from this distribution for each load to find the battery duration for each event. After finding the battery duration of each bus the model finds the initial demand unserved (load shed) at each node. If there is load shedding on that bus the probability of the tower failing is proportional to the load shed at that bus. To model the fact that if one or more towers fail, restoration in the vicinity of the failed towers becomes more difficult, CRISP increases (one time only) the restoration times of any lines connected to a bus with a failed communication system by a factor of 1.5.

3) Nuclear Restoration: Nuclear power plants have (appropriately) complicated safety procedures. This is particularly true when there are voltage fluctuations or when "Loss of off-site power" events occur, which typically trigger a long chain of shutdown procedures. Restoring normal plant operations after a safety shutdown event can take days to months. For example, a nuclear reactor can become poisoned with Xenon-136 gas if it is disconnected from the grid and must shutdown, and it can takes days for the gas to dissipate, which must occur before the plant restart process can begin. This nuclear poisoning can happen during blackouts and delay restoration of the nuclear plant. The long restoration times of nuclear plants can be seen in historical generator outage data [54].

In CRISP, this interaction occurs any time a nuclear generator has been switched off for three hours. If this does occur, then after the normal shutdown period, the nuclear plant is forced to wait t hours until starting up again, where t is sampled from the distribution of the mean time to recovery (MTTR) for nuclear generators as found in [54].

4) Compounding Risk Over Time: It is well known in power system restoration that when restoration periods become longer and critical services lack access to electricity, the restoration process itself becomes more difficult and delayed. One example is infrastructure employees being unable to drive to work when their gas tanks are empty and filling stations do not work, or when their cellular telephone batteries run out and charging is not easily available. Another example is increased traffic delays when traffic lights fail. There are many such interactions, and they may only emerge in longer and more widespread blackouts. Sometimes such interactions are surprising and only recognized when they occur. It is not feasible to model all the possible mechanisms of these interactions, but we can model the overall effect in which the restoration process during long and large events becomes increasingly severe, compounding over time. CRISP models this compounding risk over time effect as follows: The model checks each outaged line every 8 hours of restoration, and increases the restoration time by a factor of 1.05 with a probability that is the total fraction of load shed, at the current time step, in the local region (the nodes at either end of the line). The factor 1.05 came from the results of an early sensitivity study, where we found this model is very sensitive to the parameter chosen, with larger factors leading to unrealistically large impacts on the risk.

Each interdependence model depends on historical data in two ways. First, the models are impacted by the system state, and thus by the sampled generation and transmission outages. Second, each model has parameters selected based on literature about these particular interactions, and which could, in future work, be sampled from distributions representing detailed data about that particular interaction.

III. Results

This paper uses three metrics to describe the impact caused by resilience events: outage duration (hours), event size (max MW of load shed), and energy not served (ENS in MWh). We then use two different methods to visualize the statistical distributions that result from the full range of sampled events: the complementary cumulative distribution function (CCDF) of ENS and a new visualization that we call "Risk Boxes," which represent all three dimensions (duration, size and probability-weighted ENS). Together with our approach to stratified sampling, we consider these metrics and visualizations to be a useful way to describe the resilience of a particular system to extreme events. In contrast to conventional reliability analysis methods, such as Monte Carlo sampling, which tend to under sample from extreme events, the stratified sampling method used in CRISP captures the response of a given system to a wide range of high impact, low frequency events. Furthermore, our simulation method describes both the initial size of the outage and the restoration phase, which is critical to resilience analysis.

Figs. 5, 8, 9, and 11 show the CCDF of ENS for various test cases and scenarios. The CCDF allows one to see the full range of event impacts and their relative probabilities, and to compare the various interaction models. While the CCDF is useful, it is sometimes difficult to understand overall risk, which is a combination of cumulative probability and impact, in the CCDF. Also, the tail of the CCDF has significant uncertainty, due to the relatively small number of large events sampled and the small probabilities involved. Furthermore, the CCDF of ENS only describes the total energy associated with an event, which means that short duration, large size events are viewed similarly to long duration, small size events.

In order to provide further insight into system resilience, we introduce a second visualization technique, which we call Risk Boxes, that shows the overall risk measured as probability times impact (ENS). Each Risk Boxes plot has 9 boxes placed along coordinates of event sizes (MW) and event durations (hours). The plots show the risk from blackout events of different sizes (small, medium, large) and durations (short, medium, long) along coordinates for size and duration. The summed risk for each range of size and duration is indicated by the area of the box, which is normalized by the area of the largest box in the figure. The resulting Risk Boxes are shown in Figs. 6, 7, and 10. The combination of these two visualization methods (CCDF and Risk Boxes) provides additional insight into the resilience of a given system beyond either metric alone.

The following subsections discuss the results for two different test cases: (1) a modified version of the IEEE 39-bus test case [50] and (2) the 6349-bus NE-NY case.

A. IEEE 39-bus Test Case

In order to implement stratified sampling for the IEEE 39-bus case, we created one bin for each possible combination of lines out (1 to 55) and generators out (0 to 10), giving 605 bins. Stratified sampling was then applied with 100 samples drawn from each bin.

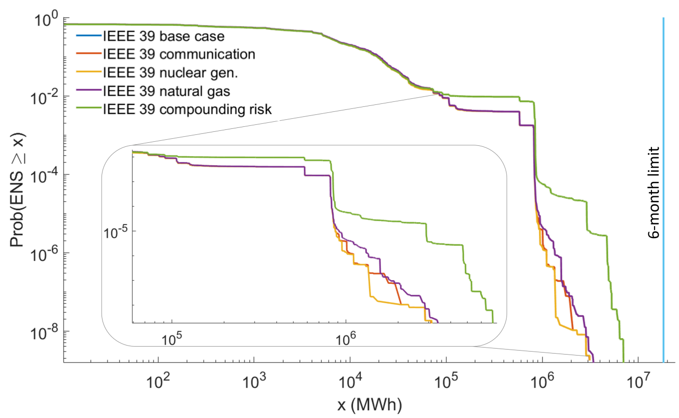

Fig. 5. The CCDF of ENS for the IEEE 39-bus test case for the base case and each of the interaction models. The inset shows more detail for the distribution of large event sizes.

1) 39-bus test case results: Figs. 5 and 6 compare the 4 interactions and the base case of no interactions for the IEEE 39-bus test case. Fig. 5 shows the CCDF of ENS. The small and medium ENS events show no significant difference from the base case with no interactions, while for some interaction models the ENS of the large events increased from the base case. Fig. 6 shows the Risk Boxes for the 39-bus

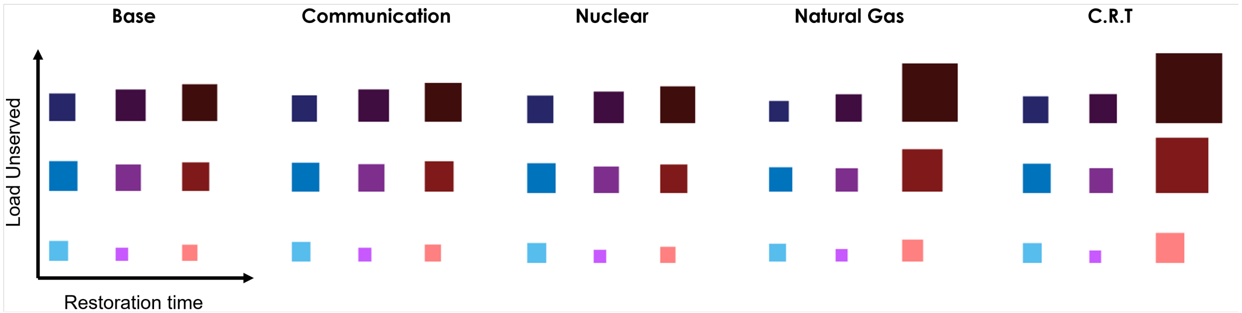

Fig. 6. Risk Boxes for the IEEE 39-bus case. The vertical position of each box indicates events with a small, medium, or large amount of MW load shed and the horizontal position of each box indicates events with a small, medium, or large outage duration. The area of each box indicates the relative risk associated with the events associated with that box. Each box has the shade and color associated with the MW shed and duration of those events. The shade of the color is lightest for small MW shed and darkest for large MW shed. The short duration events are blue, the medium duration events are purple, and the large duration events are red. The threshold between small and medium sizes events is 4000 MW, between medium and large event sizes is 6000 MW, between short and medium duration events is 48 hrs, and between medium and long duration events is 250 hrs.

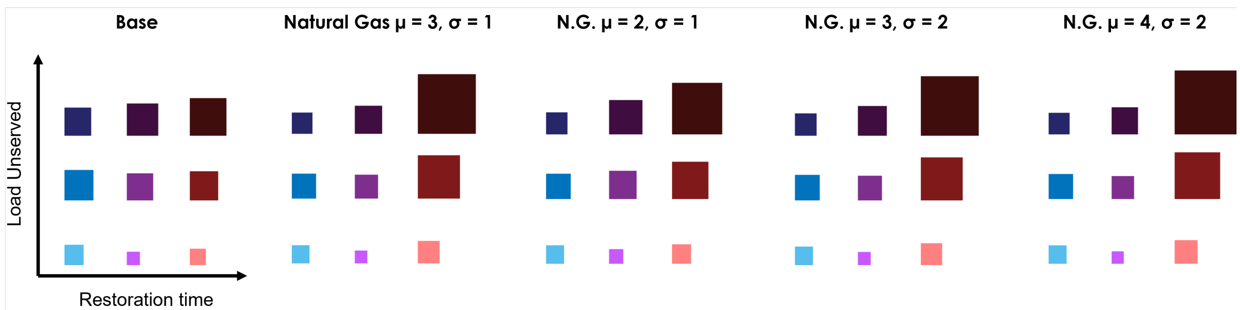

Fig. 7. Risk Boxes showing parameter sensitivity analysis results for the natural gas interaction model applied to the 39-bus test case, using the same method and thresholds as in Fig. 6.

case with the three specific interaction models and the general compounding risk over time model applied and compared with the base case. The results clearly indicate that natural gas interaction substantially increases the risk of long events over small, medium, and large event bins. In addition, the compounding risk over time interaction substantially increases the risk of long events of all sizes. The communication and nuclear restoration interactions had very little impact on resilience. Furthermore, Fig. 5 suggests that the nuclear restoration interaction does not substantially impact resilience. However, there are visible increases in the ENS of large events from the communication, natural gas, and compounding risk interactions. The compounding risk over time interaction clearly illustrates the potential impact of continued lengthening of restoration times and is the most severe interaction for the 39-bus test case.

2) Sensitivity Analysis: Each of the interaction models has various parameters that naturally impacted the outcomes. Here we discuss results from sensitivity analysis of key parameters in the interaction models for the 39-bus test case.

The assumed model parameters in the nuclear generation interaction are minimal since that model mainly used data presented in [53]. While there was one modeling parameter of the number of hours a nuclear plant needed to be out before this interaction could be initiated, changing that number from 3 hours to either 2 or 1 hours showed no effect on the results.

The communication model sensitivity analysis examined the two model parameters: the range of battery durations and the multiplication factor on the restoration time. The range of battery durations was switched from 4 to 24 hours to 1 to 4 hours. The 1 to 4 hour battery parameter results show that the shorter battery life made little difference, or even slightly improved the resilience from the original battery interaction model. The other parameter in the communication model is the factor 1.5 by which the restoration times of the adjacent lines are multiplied. This factor was increased to 1.75 and 2 for the sensitivity analysis, and both increases to this factor decrease the resilience of the power system as expected.

The log-normal distribution used to model the natural gas interaction has two parameters: μ and σ. Increases to μ or σ decreased the resilience of the power system. As shown in Fig. 7, the assumed values of μ and σ did impact the results of the study. As shown in Fig. 8, changes in μ and σ had the most impact on large and small ENS event impacts for the higher probability, medium events.

We also tested various values for the parameters in the compounding risk over time model. Changing the time parameter, set to 8 hours in the original model, by a few hours did not significantly impact the results. However, changing the factor parameter, set to 1.05 in the original model, had a significant effect on this interaction model.

3) Comparison With Monte Carlo Sampling: To validate and assess the performance of stratified sampling, we compared the results from stratified sampling to conventional Monte Carlo sampling. Fig. 9 shows results from stratified sampling with bins from 1 to 55 lines and 0 to 10 generators and standard Monte Carlo sampling, both with 60 500 samples on the IEEE 39-bus case. In order to quantify the uncertainty in each of these methods we used a bootstrapping technique using sampling with replacement from the 60 500 samples.

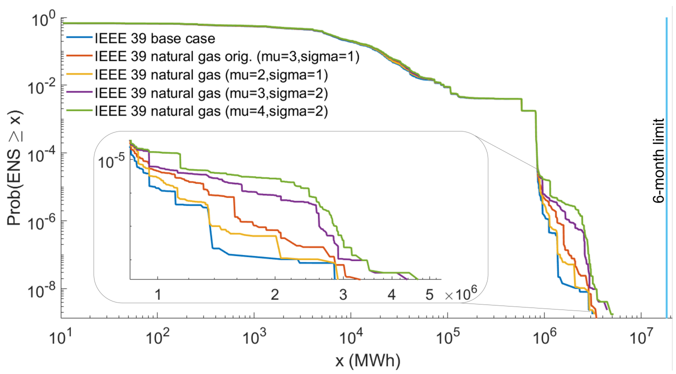

Fig. 8. CCDFs of ENS for the natural gas interaction model sensitivity analysis on the IEEE 39-bus case for the full range of events. Inset shows detail for large events.

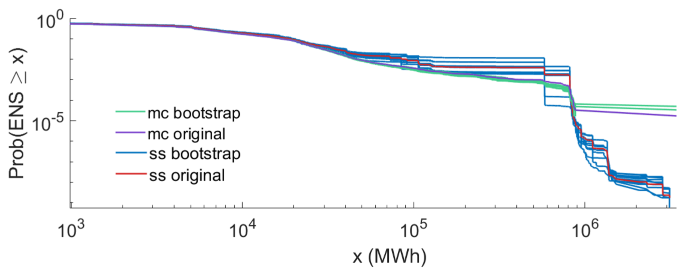

Fig. 9. Comparison of Monte Carlo sampling and stratified sampling results for the IEEE 39-bus case. To show the variance in each, we show 10 bootstrapping runs for each method. The figure shows event impacts $\geq 10^3$ MWh; the data for smaller event impacts is identical for both methods.

The results clearly show that for small events the Monte Carlo method provides less uncertainty, while for large events the stratified sampling method shows data for far more of the tail in the distribution. The benefit of stratified sampling is the added range of event impact, and the trade off for that added range is lower accuracy for smaller events. Stratified sampling allows the sampling to be deployed much more effectively across the full range of event impacts.

B. 6349-bus NE-NY Case

To implement stratified sampling for the 6394-bus NE-NY case, bins were chosen to cover a large range of line and generator outages and include rare, massive contingencies.[6] The bins were selected based on the expected wide range of outcomes from different lines and generators being sampled. The generator outages were modeled only up to 400 generators outaged at once, since the probability of more than 400 generators being out at one time is very small.

Fig. 10 shows the Risk Boxes for the three specific interaction models: communications, natural gas, and nuclear restoration, as well as a base case with no interactions. The results for the NE-NY test case clearly show the enormous impact that a natural gas shortage could have on the system. Most notably, the natural gas interaction substantially increases the risk associated with long duration and large load shed events. The nuclear restoration interaction once again shows no significant effects on the risk of the power system to any size or duration of events, while the communication interaction model shows a small increase in the risk of long events of all sizes. Fig. 11 shows the CCDF of ENS for each interaction over the full range of event impacts. Fig. 11 also shows that the natural gas interaction ranks as the most severe in decreasing resilience to large events, in the NE-NY case.

[6] The bins for the NE-NY case are as follows: line bins are $(N^{\min}, N^{\max})$ = [(1,1), (2,2), (3,3), (4,4), (5,10), (11,20), (21,40), (41,80), (81,100), (101,150), (151,200), (201,250), (251,300), (301,350), (351,400), (401,450), (451,500), (501,550), (551,600), (601,1695)]; and generator bins are $(G^{\min}, G^{\max})$ = [(0,0), (1,1), (2,5), (6,10), (11,20), (21,50), (51,60), (61,70), (71,80), (81,90), (91,100), (101,130), (131,160), (161,200), (201,230), (231,260), (261,300), (301,330), (331,360), (361,400)].

C. Discussion

The large number of extreme resilience events identified by the stratified sampling method, clearly shows the advantages of this method over conventional sampling. Stratified sampling captures low probability and extremely high impact events that Monte Carlo would not find without an unreasonable amount of sampling. This is due to the simple fact that stratified sampling samples more frequently from the extreme initiating events that are more likely to result in large blackout events, and less frequently from smaller initiating events. The captured extreme events provide us with results that we can visualize in new, insightful ways. The Risk Boxes used here give us insight into resilience, because they specifically highlight changes in risk from low probability, high impact events. In particular, the Risk Boxes show that interdependencies most substantially affect the risk of large scale and long duration blackouts.

One of the bases for the approach in this paper is that there are robust patterns in the statistics of transmission system failures. There is both theoretical and observational evidence for the robustness of the power law behavior in the distribution of observed ENS that is shown in Figs. 5, 8, 9, 11, and 12. The power law character for various measures of blackout impact has been documented across different power grids in [57], [59], [60], and its robustness has been theoretically explained as a consequence of engineering upgrades using methods of self-organized criticality in [59]. It is plausible that the distribution of distances between cascaded transmission lines and the distribution of generator outages observed over many years are characteristic of the grid and the generator fleet respectively and that these patterns will persist for future events (indeed, similar sorts of assumptions are routinely made for reliability statistics). At the very least, the forms of these statistics are arguably as reliably predictive as simulation models of failures.

The results in this study are presented as conditional probabilities, in which we describe the probability of a certain event, given that at least one transmission line is out. While the outage probability distributions are based on historical data, there were naturally very few observations for very large initiating events, making it difficult to accurately estimate the probabilities for the very large sets of transmission and generator outages, such as would result from a large winter storm and/or a cascading blackout. Therefore, the probabilities presented here should be treated as relative values for ranking; there is significant uncertainty in the actual small probabilities. Nevertheless, it may be instructive to map these conditional probabilities to annual probabilities that are more easily interpreted. In order to do so, we first need to know how often events occur. The transmission line outage data used here is

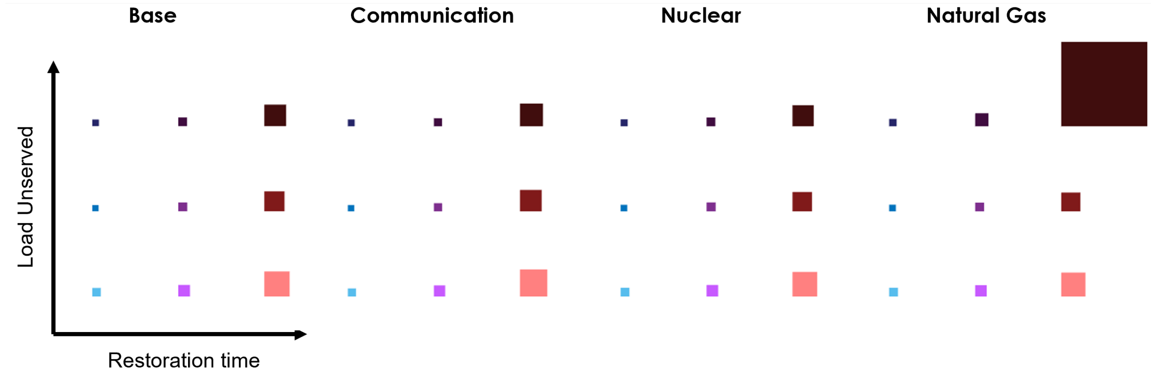

Fig. 10. Risk Boxes showing the relative risk for events of various sizes and durations of 3 specific interactions compared with the base case simulated on the 6349-bus NE-NY case, using the same method and thresholds as in Fig. 6.

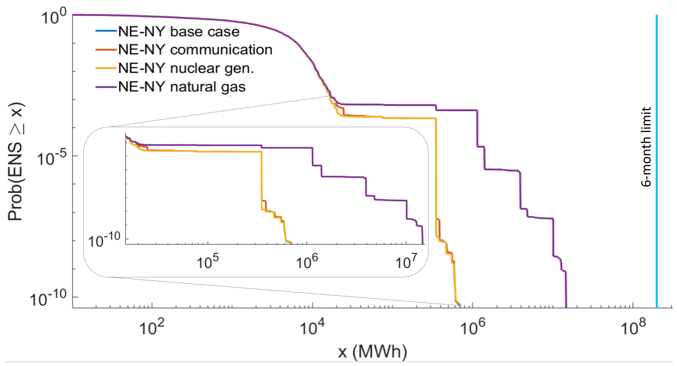

Fig. 11. CCDF of ENS for the 3 specific interaction models compared with the base case on the 6394-bus NE-NY case. Inset shows detail for large events.

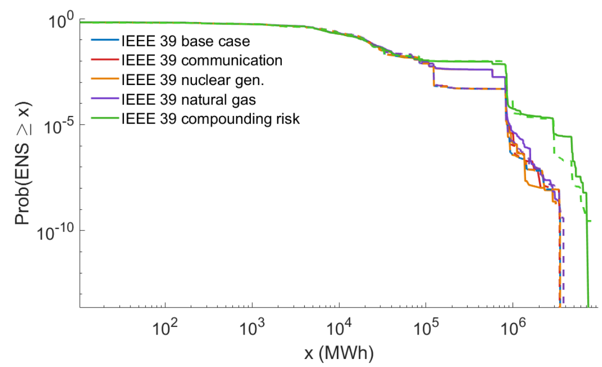

Fig. 12. The CCDF of ENS for the IEEE 39-bus test case for the base case and each of the interaction models are shown in solid lines and in dashed lines. The solid lines use line data for the entire year and the dashed lines use line data for the winter months.

based on 9,600 automatic line outages that occurred over 12 years, which is 800 unplanned outages per year [48]. So for this dataset, a one-in-a-million event ($\Pr = 10^{-6}$) corresponds to an annual probability of 0.0008.

We found it somewhat surprising that the communication interaction model resulted in only a small reduction in system resilience, despite the fact that communication systems and electric power systems are increasingly intertwined. This is likely the result of the fairly limited mechanisms through which communication failures could extend power outages in our model. Future power systems, with a larger number of small, internet-connected devices, like wifi-connected, behind-the-meter DERs and demand response programs, will be more interdependent with communication networks and these modeling assumptions may need to be revisited.

The results of the compounding risk over time interaction model on the IEEE 39-bus case surprised us, as well, with the large impacts of small compounding factors to decrease resilience. This generic compounding risk model nicely illustrates the potentially extreme effects that the many mechanisms delaying restoration taking place in concert can have.

Finally, in both test cases the natural gas interaction clearly ranked as the most severe specific interaction reducing power system resilience. This sheds light on the enormous potential social costs that could plausibly result from a long natural gas system outage in the Northeastern US. This reliance on natural gas for a very large portion of electricity supply in the Northeast region is well known, and has been highlighted in congressional testimony as a major risk [61]. The results in this paper suggests an approach to quantifying the magnitude of this risk and highlight the importance of prioritizing solutions that could mitigate this particularly severe risk. Our paper gives some quantification of this regional energy security risk and highlights the importance of identifying solutions to this risk.

D. Tailoring statistics to a specific threat or season

One goal of this paper is to show how sampling from statistics derived from historical data can represent the grid's response to resilience threats. To do this for grid resilience threats from multiple interactions with different infrastructures, we have used general historical data. The line outage statistics used in the other parts of this paper were obtained from line outage data collected year round (not from any particular season). However, it is also feasible to do the sampling from statistics derived from data that is tailored to a specific threat, or a specific season. Accordingly, we recalculate the statistics of Section IIB for the data from the New York State winter months (November - March). Note that the generation outage statistics already assumed cold weather. For the winter probability distribution of the total number of lines outaged after the initial outages and cascading, the empirical probability of exactly one line outage becomes $p_1 = 0.690$, and the Zipf distribution log-log slope becomes $s = 4.29$. For the winter line restoration times, the fitted log-normal distribution parameters for the Eastern interconnection become $\mu = 4.99$ and $\sigma = 2.59$.

Fig. 12 compares the CCDF of ENS obtained for the year-round data to that obtained for the winter months. While some small differences can be seen, the overall results and interaction rankings remain the same. That is, the overall conclusions from this paper were not sensitive to seasonal changes in the outage distributions of the transmission lines. Since we are studying the impact of interactions on the electrical grid, we did not consider the seasonality of the impact on building heating systems. Clearly, a winter outage would have higher total societal cost due to the impact on heating systems.

IV. Conclusions

This paper demonstrates a new method, called CRISP, to rank the risk associated with interdependencies between power systems and other infrastructures, based on sampling from the statistics of historical data that is commonly available to electric transmission utilities. We used the method to model and rank the risk associated with three specific types of interdependencies: communications, nuclear restoration, and natural gas. Natural gas supply failure clearly ranked as the most severe specific risk for the Northeastern US, confirming the importance of natural gas supply security in the region. We also modeled generic interactions due to several mechanisms by which restoration time delays can compound. The results show that compounding risk can have a severe impact on resilience.

In order to understand the impact of interdependence on power system resilience, it is essential to capture low probability, high impact events. The results in this paper clearly show that stratified sampling is a useful tool for describing the risk associated with low probability events. Conventional sampling methods, such as Monte Carlo, do not sufficiently sample from large, long-duration blackout events to accurately describe this risk. Many samples of small blackouts do not capture the most significant risks. This paper uses a new visualization tool, called Risk Boxes, with box area indicating risk placed along dimensions of load shed and duration. This tool allows one to see how blackout size and duration change under different circumstances.

The new methods advanced in this paper suggest a number of directions for future work. First, while we combined the best data that is generally available for the region of our case study from a number of sources, the method would benefit from systematically using the proprietary models and data that are routinely available to a single utility or system operator. Second, CRISP represents interdependencies with high-level models that approximate complicated interactions; these high-level models could be improved by developing statistical models from detailed physical interaction models, and/or data from observed interactions. For example, several detailed physics-based models of the gas and electricity interaction have recently been developed [43], [44]; it would be useful to train a statistical model on the results from such physics-based models. Other interactions could be modeled, especially when the method is applied to other regions. For example, the interactions between water and electricity infrastructure could be modeled by using the statistics of large increases or decreases in hydro resources and their impact on energy system resilience, and the impact of energy on water (and waste water) system resilience. Similarly, this approach could be used to study interactions between transportation systems and electrical systems, given the increased coupling that is brought about through transportation electrification. There is also an emerging need to extend the communication systems interaction model to better understand the implications of the growing reliance on distributed energy resources that use the public internet for communications. In addition, future work could specifically quantify the impact of policies or technologies that could improve resilience.

V. Acknowledgments

The authors gratefully thank the New York Independent System Operator for making publicly available the data which can be processed [48] into the line outage size and spread statistics. The authors gratefully thank the North American Electric Reliability Corporation for sharing the anonymized outage data from the Transmission Availability Data System. The analysis and any conclusions are strictly those of the authors and not of the NYISO or NERC.

References

[1] A. Zlotnik et al., “Coordinated scheduling for interdependent electric power and natural gas infrastructures,” *IEEE Trans. Power Systems*, vol. 32, no. 1, pp. 600–610, 2017.

[2] C. W. King et al., “The timeline and events of the February 2021 Texas electric grid blackouts - a report by a committee of faculty and staff at the university of Texas at Austin,” July 2021.

[3] A. Blankstein, T. Perry, and S. Grad, “Blackout shuts down San Onofre nuclear reactors,” *LA Times*, sep 2011. [Online]. Available: https://latimesblogs.latimes.com/lanow/2011/09/blackout-san-onofre-reactor.html

[4] A. Kwasinski, F. Andrade, M. Castro-Sitiriche, and E. O’Neill-Carrillo, “Hurricane Maria effects on Puerto Rico electric power infrastructure,” *IEEE Power and Energy Technology Systems Journal*, vol. PP, pp. 1–1, 02 2019.

[5] M. R. Kelly-Gorham, P. Hines, and I. Dobson, “Using historical utility outage data to compute overall transmission grid resilience,” *Modern Electric Power Systems Conf.*, September 2019, Wrocław Poland.

[6] M. R. Kelly-Gorham, P. Hines, K. Zhou, and I. Dobson, “Using utility outage statistics to quantify improvements in bulk power system resilience,” *Power Systems Computation Conf.*, July 2020, Porto Portugal.

[7] P. Rezaei, M. J. Eppstein, and P. D. H. Hines, “Rapid Assessment, Visualization, and Mitigation of Cascading Failure Risk in Power Systems,” in *48th Hawaii Int. Conf. System Sciences*, Jan. 2015, pp. 2748–2758.

[8] B. Walker, C. S. Holling, S. Carpenter, and A. Kinzig, “Resilience, adaptability and transformability in social–ecological systems,” *Ecology and Society*, vol. 9, no. 2, 2004.

[9] M. Bruneau et al., “A framework to quantitatively assess and enhance the seismic resilience of communities,” *Earthquake Spectra*, vol. 19, no. 4, pp. 733–752, 2003.

[10] M. Ouyang and L. Dueñas-Osorio, “Time-dependent resilience assessment and improvement of urban infrastructure systems,” *Chaos*, vol. 22, no. 033122, 2012.

[11] A. Stankovic et al., “Methods for analysis and quantification of power system resilience,” *IEEE Trans. Power Systems*, oct 2022, early and open access.

[12] ——, “Methods for analysis and quantification of power system resilience,” Piscataway, NJ, USA, pp. 1—-110, 2022.

[13] M. Panteli, D. N. Trakas, P. Mancarella, and N. D. Hatziargyriou, “Power systems resilience assessment: hardening and smart operational enhancement,” *Proc. IEEE*, vol. 105, no. 7, pp. 1202—-1213, may 2017.

[14] C. Nan and G. Sansavini, "A quantitative method for assessing resilience of interdependent infrastructures," *Reliability Engineering & System Safety*, vol. 157, pp. 35–53, 2017.

[15] E. Ciapessoni et al., "Defining power system resilience," *Electra*, oct 2019.

[16] B. Obama, "Preparing the United States for the impacts of climate change," p. 66819–66824, 2013.

[17] M. Panteli et al., "Metrics and quantification of operational and infrastructure resilience in power systems," *IEEE Trans. on Power Systems*, vol. 32, no. 6, pp. 4732–4742, Nov 2017.

[18] C. Coffrin et al., "Relaxations of AC minimal load-shedding for severe contingency analysis," 10 2017.

[19] F. Wenli, L. Zhigang, H. Ping, and M. Shengwei, "Cascading failure model in power grids using the complex network theory," *IET Generation, Transmission, Distribution*, vol. 10, no. 15, pp. 3940–3949, 2016.

[20] P. D. H. Hines, I. Dobson, and P. Rezaei, "Cascading power outages propagate locally in an influence graph that is not the actual grid topology," *IEEE Trans. Power Systems*, vol. 32, no. 2, pp. 958–967, Mar. 2017.

[21] Y. Tan, A. K. Das, P. Arabshahi, and D. S. Kirschen, "Distribution systems hardening against natural disasters," *IEEE Trans. Power Systems*, vol. 33, no. 6, pp. 6849–6860, 2018.

[22] S. Dunn et al., "Fragility curves for assessing the resilience of electricity networks constructed from an extensive fault database," *Natural Hazards Review*, vol. 19, Feb 2018.

[23] B. Cheng, L. Nozick, and I. Dobson, "Investment planning for earthquake resilient electric power systems considering cascading outages," *Earthquake Spectra*, Feb. 2022.

[24] M. Panteli et al., "Power system resilience to extreme weather: Fragility modeling, probabilistic assessment, and adaption measures," *IEEE Trans. Power Systems*, vol. 32, no. 5, pp. 3747–3757, 2017.

[25] Y. Zhou, M. Panteli, B. Wang, and P. Mancarella, "Quantifying the system-level resilience of thermal power generation to extreme temperatures and water scarcity," *IEEE Systems Journal*, vol. 14, no. 1, pp. 749–759, 2020.

[26] M. Amin, "Toward secure and resilient interdependent infrastructures," *Infrastructure Systems*, vol. 8, no. 3, pp. 67–75, 2002.

[27] M. Ouyang, "Review on modeling and simulation of interdependent critical infrastructure systems," *Reliability Engineering and System Safety*, vol. 121, p. 43–60, 2014.

[28] X. Liu, B. Chen, C. Chen, and D. Jin, "Electric power grid resilience with interdependencies between power and communication networks – a review," *IET Smart Grid*, vol. 3, no. 2, pp. 182–193, 2020.

[29] X. Lu et al., "An open source modeling framework for interdependent energy-transportation-communication infrastructure in smart and connected communities," *IEEE Access*, vol. 7, pp. 55 458–55 476, 2019.

[30] C. Zhang, J. Kong, and S. P. Simonovic, "Modeling joint restoration strategies for interdependent infrastructure systems," *PloS one*, vol. 13, no. 4, p. e0195727, 2018.

[31] J. Kong, S. P. Simonovic, and C. Zhang, "Sequential hazards resilience of interdependent infrastructure system: A case study of greater toronto area energy infrastructure system," *Risk Analysis*, vol. 39, no. 5, pp. 1141–1168, 2019.

[32] N. Goldbeck, P. Angeloudis, and W. Y. Ochieng, "Resilience assessment for interdependent urban infrastructure systems using dynamic network flow models," *Reliability Engineering & System Safety*, vol. 188, pp. 62–79, 2019.

[33] M. Korkali et al., "Reducing cascading failure risk by increasing infrastructure network interdependence," *Scientific Reports*, vol. 7, 2017.

[34] C. Jaworsky, C. Spataru, and K. Turitsyn, "Vulnerability assessment for interdependent gas and electricity networks," in *48th Hawaii Int. Conf. System Sciences*, 2015, pp. 2766–2775.

[35] Y. Ge, L. Du, and H. Ye, "Co-optimization approach to post-storm recovery for interdependent power and transportation systems," *Modern Power Systems and Clean Energy*, vol. 7, no. 4, pp. 688–695, 2019.

[36] D. B. Karakoc et al., "Community resilience-driven restoration model for interdependent infrastructure networks," *Int. Journal of Disaster Risk Reduction*, vol. 38, no. 101228, 2019.

[37] J. Sun and Z. Zhang, "A post-disaster resource allocation framework for improving resilience of interdependent infrastructure networks," *Transportation Research Part D*, vol. 85, p. 102455, 2020.

[38] L. Martins et al., "Interdependence between power grids and communication networks: A resilience perspective," in *13th Int. Conf. Design of Reliable Communication Networks (DRCN)*, 2017, pp. 1–9.

[39] H. A. A. Moham, "A synthetic case study for analysis of the rising interdependency between the power grid and e-mobility," *IEEE Access*, vol. 7, pp. 58 802–58 809, 2019.

[40] H. Zhang et al., "Resilience assessment of interdependent energy systems under hurricanes," *IEEE Trans. Power Systems*, vol. 35, no. 5, pp. 3682–3694, 2020.

[41] C. Shao et al., "Integrated planning of electricity and natural gas transportation systems for enhancing the power grid resilience," *IEEE Trans. Power Systems*, vol. 32, no. 6, pp. 4418–4429, 2017.

[42] S. D. Manshadi and M. E. Khodayar, "Resilient operation of multiple energy carrier microgrids," *IEEE Trans. Smart Grid*, vol. 6, no. 5, pp. 2283–2292, 2015.

[43] L. A. Roald et al., "An uncertainty management framework for integrated gas-electric energy systems," *Proceedings of the IEEE*, vol. 108, no. 9, pp. 1518–1540, 2020.

[44] A. Antenucci and G. Sansavini, "Gas-constrained secure reserve allocation with large renewable penetration," *IEEE Trans. Sustainable Energy*, vol. 9, no. 2, pp. 685–694, 2018.

[45] E. A. Martinez Cesena, E. Loukarakis, N. Good, and P. Mancarella, "Integrated electricity-heat-gas systems: Techno-economic modeling, optimization, and application to multienergy districts," *Proceedings of the IEEE*, vol. 108, no. 9, pp. 1392–1410, 2020.

[46] J. M. Gonzalez et al., "Spatial and sectoral benefit distribution in water-energy system design," *Applied Energy*, vol. 269, p. 114794, 2020.

[47] H. Ren, I. Dobson, and B. Carreras, "Long-term effect of the n-1 criterion on cascading line outages in an evolving power transmission grid," *IEEE Trans. Power Systems*, vol. 23, pp. 1217–1225, 09 2008.

[48] N. K. Carrington, I. Dobson, and Z. Wang, "Transmission grid outage statistics extracted from a web page logging outages in Northeast America," in *North American Power Symposium*, Nov. 2021.

[49] S. Tamronglak, S. H. Horowitz, A. G. Phadke, and J. S. Thorp, "Anatomy of power system blackouts: preventive relaying strategies," *IEEE Trans. Power Delivery*, vol. 11, no. 2, pp. 708–715, Apr. 1996.

[50] "Modified IEEE 39 bus test case," https://github.com/phines/infrastructure-risk/tree/master/CRISP/data/case39.

[51] C. Barrows, "Reliability test system - grid modernization lab consortium," aug 2019. [Online]. Available: https://github.com/GridMod/RTS-GMLC

[52] S. Murphy, L. Lavin, and J. Apt, "Resource adequacy implications of temperature-dependent electric generator availability," *Applied Energy*, vol. 262, no. 114424, pp. 1–14, 2020.

[53] S. Murphy, F. Sowell, and J. Apt, "A time-dependent model of generator failures and recoveries captures correlated events and quantifies temperature dependence," *Applied Energy*, vol. 253, no. 113513, pp. 1–11, 2019.

[54] S. Murphy, J. Apt, J. Moura, and F. Sowell, "Resource adequacy risks to the bulk power system in north america," *Applied Energy*, vol. 212, p. 1360–1376, 2018.

[55] S. Ekisheva et al., "Impact of extreme weather on North American transmission system outages," in *IEEE PES General Meeting*, Jul. 2021.

[56] A. Clauset, C. Shalizi, and M. Newman, "Power-law distributions in empirical data," *SIAM Review*, vol. 51, no. 4, pp. 661–703, 2009.

[57] B. A. Carreras, D. E. Newman, and I. Dobson, "North American blackout time series statistics and implications for blackout risk," *IEEE Trans. Power Systems*, vol. 31, pp. 4406–4414, Nov 2016.

[58] A. B. Owen, *Monte Carlo theory, methods and examples*. book in draft, 2013, available at https://statweb.stanford.edu/˜owen/mc/.

[59] I. Dobson, B. Carreras, V. Lynch, and D. Newman, "Complex systems analysis of series of blackouts: cascading failure, critical points, and self-organization," *Chaos*, vol. 17, no. 2, p. 026103, June 2007.

[60] P. Hines, J. Apt, and S. Talukdar, "Large blackouts in north america: historical trends and policy implications," *Energy Policy*, vol. 37, pp. 5249–5259, 2009.

[61] ISO-NE, "Operational fuel-security analysis," Jan. 2018, https://www.iso-ne.com/static-assets/documents/2018/01/20180117_operational_fuel-security_analysis.pdf.