Benchmarking and Validation of Cascading Failure Analysis Tools

IEEE Working Group on Understanding, Prediction, Mitigation and Restoration of Cascading Failures

Abstract—Cascading failure in electric power systems is a complicated problem for which a variety of models, software tools, and analytical tools have been proposed but are difficult to verify. Benchmarking and validation are necessary to understand how closely a particular modeling method corresponds to reality, what engineering conclusions may be drawn from a particular tool, and what improvements need to be made to the tool in order to reach valid conclusions. The community needs to develop the test cases tailored to cascading that are central to practical benchmarking and validation. In this paper, the IEEE PES working group on cascading failure reviews and synthesizes how benchmarking and validation can be done for cascading failure analysis, summarizes and reviews the cascading test cases that are available to the international community, and makes recommendations for improving the state of the art.

Index Terms—Cascading failure, validation, risk analysis, power systems reliability.

I. INTRODUCTION

ACCORDING to the North American Electric Reliability Corporation (NERC), a cascading blackout is "the uncontrolled successive loss of system elements triggered by an incident at any location" [1]. Because not all cascading outages (a sequence of interdependent component outages) result in a blackout (a large, unplanned loss-of-load), this paper uses the term cascading failure to represent any sequence of independent and dependent outages, regardless of whether a blackout ensues. Cascading failures are typically triggered by one or more disturbance events, such as a set of transmission line or generator outages. Triggering events can result from a variety of exogenous threats, such as earthquakes, weather-related disasters, hidden failures, operator errors, and even deliberate acts of sabotage. Since power systems are generally operated to be N-1 secure, most historical cascades have been triggered by multiple outages in combination, motivating the need for probabilistic analysis. The dependent outages in a cascade can result from a wide variety of different mechanisms including thermal overloads, voltage instability, and angular instability [2].

Because the resulting blackouts can be large and costly, utilities are increasingly required by reliability regulators to systematically study and manage cascading outage risk in their system. For example, NERC planning standards [3] require that, "Each

Transmission Planner and Planning Coordinator shall investigate the potential for cascading and uncontrolled islanding in its planning assessment studies." Specifically, NERC requires utilities to complete simulation studies that address each of the following types of cascades:

- Overloads where a component exceeds the phase protective relay settings, assumed to be in accordance with PRC-023-2 [4], or a rating established by the operator (overload cascading);
- Multiple generators pull out of synchronism with one another (angular instability cascading);
- Poor transient voltage response due to insufficient dynamic reactive resources (voltage instability cascading).

In addition, new standards are in development in Europe (see [5], [6], and [7]) and in the USA (NERC standard TPL-007-1 [8]) that will require analysis of additional exogenous threats such as geomagnetic disturbances.

In response to increasing regulations and several large cascading blackouts [9]–[11], a growing number of tools are being developed in industry and academia to address this analysis need. Given that these tools are increasingly being used to make large investment decisions, and the critical importance of managing the risk of massive cascading blackouts, it is important that cascading failure analysis tools be tested to ensure that they provide accurate and useful information. Doing so requires verification (ensuring that tools perform correctly), validation (checking the accuracy of the results), and benchmarking (a systematic, reproducible validation procedure).

Practical benchmarking and validation also require the use of standard, published sets of test case data that in some way represent a particular power system (hereafter "test case"). Existing test cases are scattered across multiple continents and are often difficult to find or access. Comprehensive information about these cases is sorely lacking. In this paper, we describe and reference a wide variety of international test cases (both public and nonpublic) and provide details on how to access them. Also, validation studies often (and should) make use of historical data from power system operations. Here we describe several sources for this type of data.

Thus motivated, the three goals of this paper are: (1) to discuss existing approaches to the problem of benchmarking and validating cascading outage data and simulators; (2) to provide guidance for practitioners and researchers seeking to objectively evaluate a particular cascading failure analysis technique or software tool; and (3) review the available test cases and other data for cascading failure analysis. This paper brings together current research and expert opinions from members of

the IEEE PES Working Group on Cascading Failure, building on prior work in which the group addressed methodologies and tools [12], [13].

The paper is organized as follows: Section II presents the definitions and main requirements for benchmarking and validation. Section III outlines recommendations for effective benchmarking. Section IV briefly reviews several published benchmark/validation studies. Section V critically reviews sources of non-test-case data that can be used in cascading failure analyses. Section VI describes a wide variety of public and nonpublic power system test cases, outlines how these cases are, or are not, useful for cascading failure analysis, and provides guidance for obtaining the cases. Section VII discusses the need for new test cases. Finally, Section VIII summarizes the conclusions of the Working Group for cascading failure benchmarking and validation.

II. Benchmarking and Validation: Concepts and Requirements

This section defines terms and describes important attributes of benchmarking and validation studies.

A. Definitions

Benchmarking is a process for measuring the performance of a tool, such as a software program or a business process, using a trusted procedure and/or dataset, in a way that allows one to compare the performance of one tool to another. Using trusted data and procedures when comparing tools allows for relatively objective comparisons. For example, the LINPACK benchmark [14], which is used to rank supercomputers, is a package of data, software libraries, and procedures that, when used correctly, allows one to compare the computational performance of different computer systems.

Because cascading failure analysis (unlike power flow analysis, for example) is a relatively immature power systems application area, and because there are many uncertainties and challenges in modeling and simulating cascading failure, there are few complete benchmarks for cascading failure analysis; this paper outlines what does exist, and suggests ways to improve the state of the art going forward.

Benchmarking is closely related to the processes of verification and validation [15]. *Verification refers to the process of checking to see that a tool solves the problem that it was intended to solve.* In the context of a software tool (as is the case with most cascading failure analysis systems), verification involves checking to ensure that the tool produces the answers that it should get (without numerical instability or memory errors, etc.), given its internal assumptions, over a wide range of possible operating conditions [16]. Since verification is more about avoiding software errors, rather than fundamental methodology, this paper focuses more on *validation*, which *is the process of checking a system to ensure that it obtains answers that are correct, according to a set of criteria for correctness* [16]. Verification involves checking the correctness of the tool (often in the relatively narrow sense of being free of bugs), whereas validation typically involves the combination of the tool with some type of test data to evaluate the correctness of the answers provided by a method. Benchmarking typically brings the two processes together to create a reproducible process for validating and comparing different approaches to the same problem.

Within a power systems context validation is necessary both for software tools, which integrate concepts about cascading, mathematical representations of those concepts, and ultimately a software encoding of the mathematics, and for the datasets to which cascading failure analysis is applied. The latter is particularly important for industry practitioners who need to ensure that a particular dataset is an accurate (or at least useful) representation of a particular network.

There are many useful approaches to benchmarking and validation. The following are a few examples:

1. *Checking for internal validity*. Internal validation involves checking the assumptions that go into a tool to determine if they are realistic. Simulation studies involve some simplification and approximation of physical processes, so internal validation involves determining which assumptions are likely to produce misleading results, and which ones are appropriate given the purpose of the tool.

2. *Comparing simulation results with real data*. The ability to reproduce reality with sufficient exactness is the usual ultimate goal in validation. However, in the case of cascading outages, which encounter many thresholds for discrete actions such as tripping a line or not tripping a line, similar tools with similar data (or even the real power system on successive days) may behave differently under very similar conditions. One tool may trip the line and another very similar tool may not and this can have a large effect on the way that a particular cascade evolves. Therefore it is usually too stringent to require an exact match of the simulation to real blackout data. There are two approaches to solve this. One approach is based on engineering judgment and asks whether the simulated cascade is one of the plausible sequences of events given an engineer's experience with real events. The second approach gathers the statistics of real and simulated cascades and asks whether they have similar statistical characteristics. For example, the distributions of final blackout sizes or how much the cascades propagate can be compared. The quantities compared should correspond to the conclusions that will be drawn from using the tool.

3. *Comparing the performance of one tool with another tool (cross-validation)*. Once good performance metrics (benchmarking processes) have been established, measuring the similarities and differences between tools can provide valuable insight into the relative merits of the tools. Often this process will show that different tools are useful for solving different types of problems.

4. *Checking for reproducibility*. Given the same method, assumptions and data, it is important that a tool be able to lead one to similar conclusions from one run to the next. Because many cascading failure simulation tools include some random variables, the results will sometimes be somewhat different among trials; in these cases it is important to know how many runs are needed to produce statistically reliable results. It is also important to check that minor changes in data and assumptions do not produce dramatically different results.

5. *Sensitivity analysis*. When a model has many input parameters, it is important to know how results change with changes in the inputs. A sensitivity analysis tests the impact of perturbations to inputs to identify those parameters that have substantial impact on outcomes.

Section II-D discusses the application of these approaches to power systems in additional detail.

A key component of benchmarking and validation is a set of trusted test cases/datasets with known properties to which each tool can be applied. Because there has been little discussion in the literature about the merits of various test cases for cascading failure analysis, Section V of this paper focuses on public data about power system outages, since these are particularly valuable in benchmarking and validation. Sections V–VI focus on the use and origins of power system test cases.

B. The Challenges of Cascading Failure Simulation

Most approaches to cascading failure analysis involve the use of a cascading outage simulation tool. Developing and, ultimately, validating these tools is a substantial challenge because of the numerous, diverse set of mechanisms by which all real cascades propagate. For each additional mechanism of cascading included in a model one needs to make assumptions about how a system will react to extreme, rarely observed operating conditions.

Potential mechanisms that might be modeled include an array of traditional instability and protection phenomena including cascading overloads interrupted by relays, hidden relay failures, voltage collapse, dynamic instability, and inter-area oscillations. These have been discussed substantially in the research literature and reviewed by IEEE and CIGRE working groups [17], [18]. In addition, cascading failures typically involve an array of communication, control, economic and societal factors. All of these mechanisms occur at diverse time scales, further complicating the modeling process. Human operators play a particularly important, and difficult-to-model role. Operators' inadequate situational awareness was an important factor in a number of recent disturbances (e.g., Europe in 2006 [10] and North America in 2003 [9] and 2011 [11]). On the other hand, operator actions can also reduce risk in a system; appropriate mitigating actions by operators can arrest the spread of large blackouts. Modeling operator actions is a substantial challenge.

Other key complicating factors are the uncertainties of the system state, and the stochastic nature of both the triggering (exogenous) events that lead to the start of a cascade (day, time, weather, etc.), and the interdependent (endogenous) events involving control, dynamics, and protection through which a cascade propagates. Deterministic models can be useful when one wants to reproduce a historical event, or to study a well understood operating condition such as a state estimator case, whereas probabilistic models become increasingly important as uncertainty further increases, such as in a planning context.

While there has been progress in modeling some of these mechanisms (both triggering and propagating, deterministic and stochastic), the relative importance of the various mechanisms is largely unknown. The amount of modeling detail required to accurately represent each of the mechanisms in a way that leads to useful engineering conclusions remains an open question. However, what is clear is that validation and benchmarking are critically important in order to improve existing models and to identify areas where more research and development are needed.

C. Importance and Users of Validation and Benchmarking

The many complicated mechanisms involved in cascading precipitate an even greater need for ensuring that analysis methods are valid. Benchmarking and validation are necessary to determine which aspects of real blackouts are reproduced by different types of models, what sort of conclusions can reasonably be drawn from a particular tool, and what limitations exist for a particular methodology. There will always be a gap between simulation and reality; benchmarking and validation are needed to understand this gap, to interpret the results, and to determine the extent to which the results can inform engineering decisions. Understanding this gap is also necessary in order to improve the next generation of modeling and simulation tools.

Validation and benchmarking are important for a variety of stakeholders in the electricity industry. Validation is important for researchers testing new ideas to understand the implications of, for example, new modeling approaches. For software developers and vendors, validation allows potential clients to understand the limitations of, and gain confidence in a particular tool. And utilities and system operators need to validate both case data and tools to ensure that investment and operating decisions are based on sound data and models.

D. Approaches to Validating Cascading Failure Simulations

Several different approaches can be employed to effectively validate power system simulations. Within more established, restricted and well understood problems, such as power flow and standard contingency analysis, there is a measure of consensus in the power systems engineering community regarding the amount of detail needed to answer particular, well-defined questions. For problems of this sort, it makes sense to validate simulations by evaluating the accuracy of the component models and measuring the extent to which models align with measurements. However, this type of consensus does not yet exist for cascading failure simulation and analysis. As a result, validation approaches that are entirely patterned on the traditional approaches for more established problems are not generally practical; there is a need for new approaches or at least significant extensions to existing ones.

One of the key reasons for this lack of consensus is the diversity of ill-understood and difficult-to-model mechanisms involved in cascading failure, as explained in Section II-B. In the following we describe several different approaches that can be usefully employed to validate a particular cascading failure analysis method.

1) Internal Validation and Sensitivity Analysis: In all cases, engineering judgment is needed to determine if a cascading failure analysis tool is internally valid or invalid, in that the modeling assumptions are inappropriate to the questions that the model is trying to answer. For example, if a tool focuses on

studying how cascading overloads propagate through a transmission system, power-flow models (and not abstract topological ones; see [19]) are needed in order to provide valid answers.

An important tool for internal validation is sensitivity analysis [20]. If a tool is intended to produce a particular statistical outcome (e.g., Loss of Load Probability or Expected Energy Unserved), sensitivity analysis can be used to determine if a particular modeling assumption has an important impact on the outcome variable of interest. Assumptions that do not significantly impact the outcome do not merit as much attention as those that impact the outcome statistic, and thus could lead to erroneous conclusions.

2) Validation: Comparing Models to Real Data: In the case of cascading failure, validation against real data is particularly important, because complete internal validation is infeasible. Modeling all mechanisms of cascading failure in great detail is not computationally feasible. But what types of comparisons are appropriate and feasible for validating cascading failure analyses?

One approach is to compare simulated event sequences to historical cascade sequences. Many simulations are specifically designed for the purpose of reproducing a particular historical cascade (e.g., [21]). Doing so can help in understanding the cascading mechanisms that contributed to a particular blackout and developing the lessons learned. Moreover, inaccuracies in particular system model can be revealed. In this case, the validation process is clear: the simulation is validated for that particular blackout when it closely reproduces the sequence of events of that blackout. The stochastic nature of initial and operating conditions that must be considered when modeling potential future blackouts, or when studying blackout risk in general, does not apply for this type of event reproduction, since the initial operating conditions are known. Reproduction studies of this sort typically require tuning of component thresholds and settings to reproduce the exact sequence that occurred.

However, the ability of a simulation to reproduce a particular historical cascade does not necessarily imply that it is valid for cascading failure analysis in general, across all operating conditions. Uncertainty in the many thresholds involved in cascading (e.g., the current at which a line will sag into vegetation), operating conditions, and operator actions mean that validation of a simulator against one event is insufficient to argue that a model is valid for all conditions. Because of these uncertainties, two different tools, both appropriately validated, can produce different event sequences for the same triggering disturbance. Even in a real power system, similar initiating events can lead to different outcomes on different days. In addition, power network data (line impedances, generator set points, etc.), which are core to all simulation tools, always include some inaccuracy. An exact deterministic match between a cascading simulator and data from particular observed cascades is not a necessary condition for validation. While the tuning of simulations to observed cascades is useful, it is not clear that this tuning will consistently improve a simulator's ability to facilitate good decisions about other cascades. An additional challenge to this type of validation is that the complete data from historical cascades are rarely, if ever, made broadly available, which means that very few tools can be extensively compared with historical event sequences.

On the other hand, comparing the statistical properties of simulated cascades to the statistical properties of historical cascades is feasible, and provides valuable insight. Indeed there are distinctive patterns in the observed statistics of historical cascading blackouts, which can be reproduced in simulators [21]–[25]. Thus one approach to validation is to run a simulator with a suitable sampling of initial states and initiating events, gathering the resulting statistics, and then comparing the simulation's statistics with the historical statistics [26]. A good statistical match does not necessarily prove that the simulation is valid, but it is a positive indication that the simulation captures important features of cascading.

In particular, one use for blackout size data is to develop empirical probability distributions of blackout sizes [21], [23], [27], against which simulated data can be compared. An important feature of these distributions is the heavy tail, which indicates that larger blackouts are more likely than predicted by conventional risk analysis methods. The heavy tails result largely from cascading failures, in which small disturbances propagate additional outages, which progressively weaken the system and ultimately produce large blackouts.

Conclusion that are more definite can emerge when there is a gap between simulated and observed statistics. If the gap is too large, then the simulation is not validated. For example, a tool cannot be considered valid for long cascades if it cannot at least qualitatively reproduce the heavy tails that are consistently found in the distribution of historical blackout sizes. Another useful statistical measure is the observed frequency of cascades of various sizes. On the other hand some simulators are not designed to fully simulate long cascades, but are designed instead to stop after cascade sizes cross a particular threshold (at which point the sequence of events can become highly uncertain; see, e.g., [28], [29]). In these cases, comparing the frequency of cascades larger than 10 (for example) dependent events to historical data may also be a useful approach. Knowing the nature of the statistical gap between a simulator and historical data can drive improvements to the model, such as choosing which mechanisms to model in greater detail. Standard statistical tests can be applied to quantify the gap between real and simulated probability distributions. One avenue of further research would employ formal computer model calibration techniques such as in [30] and references therein.

3) Cross-Validation: Comparing Models to Each Other: Another useful approach to validation is to compare the statistical properties of one model with those of another. Observed differences between models can be useful in understanding the relative importance of different modeling assumptions. One approach, proposed in [31] and [32], is to compare the extent to which cascading event sequences in two different models include the same endogenous events, when subjected to the same sets of initiating events. Doing so allows one to test the importance of particular modeling assumptions. If a parameter variation substantially changes the agreement between a new model being tested and a reference model, that parameter is important and should be studied in detail.

III. CHARACTERISTICS OF A HIGH-QUALITY CASCADING FAILURE VALIDATION/BENCHMARK ANALYSIS

In response to increasing attention to the cascading failure problem, there are an increasing number of tools available from both research and industry organizations; however, it is difficult to know how these tools have been validated. Publicly releasing the results of cascading failure validation studies can be tremendously valuable to both practitioners and researchers. This section offers the working group consensus recommendations for publishing the validation and/or benchmarking studies. Our objectives are to provide a checklist of items that should ideally be included in such publications and to speed the dissemination of ideas and accelerate acceptance and validation. This working group is in turn committed to providing venues for presenting and publishing these ideas through its committee meetings, website [33], and panel sessions.

The core elements of a cascading failure validation study are typically: (1) one or more power system case datasets, (2) cascade failures models embedded into simulation software, and (3) data against which the simulation results are to be compared. Validation is important for all three of these stages: case data, models, and comparison data.

With respect to test case data, studies should use either or (preferably) both of public power-flow cases and real industry-provided cases. The use of public test cases makes it possible for other researchers to reproduce and verify the results of particular studies. Real-world test cases allow one to evaluate the suitability of a tool for practical industry application. In all cases, studies should clearly document data sources. If data access is restricted, then the process for obtaining the data should be outlined in the study. In situations where new or supplemental data are added to publicly available test cases, the best practice is to publish the data on the Internet, noting any restrictions that are necessary for a particular dataset. Often, it is necessary to add new types of data to existing test cases in order to test a particular tool. For example, component outage probabilities and switching configurations can be key elements of cascading, but the relevant data are not typically included with standard test cases. In such cases, carefully documenting the methods used to augment the test case is necessary.

In addition, the working group recommends that studies include the following elements:

- Studies should clearly state the objectives of the analysis, and the types of conclusions that can be drawn from the results. For example, the study should state whether the objective is to give a set of credible cascades under stated conditions or whether it is to estimate the probability or risk of certain cascades.
- Studies should include some tests on larger power systems (e.g., hundreds to thousands of buses), or at least a description of how the methods can be scaled up to larger systems. Cascading is a large-scale phenomenon; the size of a test network can have important impacts on the outcomes [34].
- Studies should clearly describe the rationale for selecting the particular test cases or data used for the study. The report should clearly explain the advantages of the particular data used, and the shortcomings of data that were not used. In some cases, limitations in data availability may constrain the available choices.
- The benchmark should include, or at least reference, detailed descriptions of the modeling procedures used to come to the report's conclusions. The benchmark should state which initiating and cascading failure mechanisms are modeled and indicate how each is modeled. Ideally the detail should be sufficient so that an informed researcher can statistically duplicate the results (keeping in mind the fact that cascading failure is necessarily a stochastic phenomenon, making it unlikely that every detail will be precisely replicable). Including appendices with example event sequences resulting from particular initiating disturbances is one way to do this. In other cases, the results to be reproduced might be measures of risk, such as a probability distribution function for blackout sizes [26], for well-documented cases such as the IEEE RTS [35].
- If the benchmark makes any claims about probability or risk, it must sample from the sources of uncertainty and estimate a probability distribution of the blackout size for comparison with real data and other simulations. Blackout size measures include line outages, load shed, and energy unserved. Publishing the statistics of other quantities (for example, propagation or cascade spreading) is encouraged. Note that even "deterministic" simulations can sample the operating conditions and initial outages to estimate a probability distribution of blackout size. When sampling is used (which should be the case in the vast majority of studies), the benchmark should carefully explain the sampling methods. The report should specify how the method samples from the potential operating conditions, initial faults, and the progress of the cascades.
- A clear distinction should be maintained between models attempting to reproduce in detail characteristics of specific historic cascade events, and those that aim to assess overall risk from cascading events on a planning timescale.
- Particularly for probabilistic tools, it is important that a tool be able to identify internally consistent pre-contingency conditions, such as would result from a security constrained optimal power flow or similar dispatch routine [36].
- Finally, as previously mentioned, benchmarks should clearly list their data sources, including test case data, outage data (e.g., TADS [37]). If non-public data were used, the authors should include as much detail as possible about where the data came from and what procedures (e.g., Critical Energy Infrastructure Information [38]) are needed to obtain similar data.

One of the most important characteristics of a benchmarking/validation study is to provide quantitative metrics that allow future analysts to compare the statistics of different studies, and qualitative descriptions that facilitate similar comparisons. Conventional reliability statistics, such as Loss of Load Probability, are not particularly good measures of cascading failure risk because of the many uncertainties involved and the fact that cascade sizes can span several orders of magnitude. However, useful comparisons can be made. Detailed qualitative descriptions that compare simulated cascades

to real or other simulated ones can allow one to evaluate the credibility of cascade sequences, in terms of including familiar or reasonable outage interactions, appearance of previously known grid weaknesses, and the overall degree of degradation in the grid due to cascading relative to the initiating contingencies. Descriptions of the patterns in which cascades spread given the cascading mechanisms modeled can be useful. Useful quantitative features for validation studies include probability distributions of cascade size, average amount of propagation, lengths of cascades, and statistics of cascade spread. References [21], [27], [28] provide useful examples of measurable comparison statistics. Further research on formal statistical approaches for comparing model outputs with historic data would be of value.

IV. EXAMPLE BENCHMARKING STUDIES

The Working Group has identified several papers that contain notable elements of a good self-published benchmark/validation study. The following subsections briefly review several of these analyses.

A. Reproduction of Existing Cascading Blackout Event

The cascading blackout on August 10, 1996 blackout resulted in a loss of 30,390 MW of load and affected 7.5 million customers in western North America. Reference [21] describes efforts to reproduce the events of this massive cascading failure using a transient stability simulation tool (GE PSLF) [39]. This study is notable for several reasons. First, it clearly documents the process of comparing a simulation model to data from a historical cascade. Second, the study illustrates the challenges of reproducing a historical cascade with a particular model. As documented in the paper, the authors needed to make substantial adjustments to the component models before they were able to accurately reproduce data from the disturbance. The paper illustrates the type of insight gained from reproduction studies.

B. Validation via Comparison to Time-Domain Simulations

References [40], [41] discuss the validation of the cascading simulator of a probabilistic tool for operational risk assessment called PRACTICE [42], [43] using a time domain power system model of the Italian EHV transmission system from the early 2000's. In this case, cascading event sequences identified by the simulator were compared to corresponding sequences of branch outages obtained from a detailed dynamic model. This comparison was completed for a large set of single and multiple contingencies and for two different initial operating conditions (peak daytime, and nighttime) [40]. The two sets of simulated outages matched well, especially during the early stages of cascading, where slower overloading events were the primary mechanism of failure. The fact that the simulated outages did not match well for events during later stages of the cascade highlights the challenges of modeling the many mechanisms of cascading that occur after the early stages.

C. Extension of Traditional Methods

The benchmarking study reported in [28] describes a method for extending traditional reliability planning methods to address NERC requirements for multiple-contingency analysis. Notably the proposed method classifies contingencies that cause limit violations into those that are not likely to initiate cascading failure, and those that "cannot be eliminated as potential causes for widespread outages." The paper describes the data sources for the study, as well as information about how the data might be obtained, and provides enough information about the methodology that an informed reader might duplicate the results. Power systems conferences provide a good venue for publishing studies of this sort. The panel format allows for presentations to supplement printed material. For example, the presentation associated with [28] provided detailed statistical results and an example of a cascading chain that could reasonably be excluded from cascading and another that could not.

D. Benchmarking With a US Western Interconnect Model

TRELSS is a software package for cascading failure analysis [44]. References [45], [46] describe results from an extreme events research team, which developed a $\sim$16,000-bus Western Interconnection power flow model for cascading failure simulation, generated a significant number of initiating events ($\sim$33,000) to systematically generate cascading scenarios (NERC Category-D events [47]), and simulated/evaluated the cascading sequences that followed them using TRELSS. Methods were developed for identifying critical event sequences based on their occurrence in many simulated blackouts and ranking initiating events in order of severity. Notable features of the benchmark include the modeling of protection control groups and voltage problems. Protection control groups approximate the effect of the protection system when there is a fault with the simultaneous outage of predetermined groups of components. Voltages were modeled using a quasi-steady state AC power flow model.

E. Benchmarking With Eastern Interconnect Models

The Potential Cascading Modes (PCM) cascading failure simulation software has been tested during several demonstration projects with US utilities. In [48] PCM was used to automate the process of sequential AC contingency analysis in order to identify initiating events that may lead to cascading outages due to thermal overloads and voltage violations. The project was a large-scale demonstration project using a circa 2007 US Eastern Interconnection model with approximately 50,000 buses. The project studied 250 NERC Category B (single) contingencies and approximately 31,000 Category C (multiple) contingencies [47]. While the results were consistent with prior manual analysis of some of the extreme events, PCM also identified some potentially cascade-initiating contingencies that were previously unknown to the participating utilities.

A second project [49], which also used an Eastern Interconnect model, focused on identifying and analyzing optimal remedial actions needed to prevent cascades or mitigate their effects. Two types of computations were performed: (1) Determining measures to prevent cascading, and (2) Mitigating the consequences of cascading outages after they have occurred. The results indicated that all identified potential cascading modes may be prevented using the existing controls in the network, which was consistent with manual analysis previously performed by the utility.

F. Validation of OPA on WECC Data

The OPA cascading blackout simulation [21] was validated on a 1553-bus WECC network model by determining OPA parameters from WECC data and then comparing the blackout statistics obtained with OPA to historical WECC data from NERC and BPA TADS [26]. The blackout statistics compared were the distribution of blackout size and the propagation and distribution of line outages. Reasonable agreement was verified, and attributed to the modeling of the complex system feedback modeled in OPA by which the power grid upgrades in response to blackouts.

V. SOURCES OF CASCADING OUTAGE DATA

As suggested previously, comparing a model's results to historical data from real power systems is a useful validation method. However, obtaining good data is often not trivial [50]. Here we discuss some of the types of data that are available.

A. Historical Blackout Size Data

The North American Electric Reliability Corporation (NERC) had previously made public data for reportable blackouts in North America since 1984. These data indicate that there are approximately 13 very large blackouts (above ~ 300 MW) per year. The measures of blackout size in the NERC data include load shed (MW) and number of customers affected. Blackout duration information is also available, but the data quality is less certain. A lightly processed version of these data used in [23] is available on the Internet [24]. Blackout size data of this sort have provided key insights in the study of cascading large blackouts. The NERC data have been analyzed in [21], [23] and used for validation of a cascading failure simulation in [26]. International data on the distribution of blackout size is reviewed in [27].

The NERC data result primarily from U.S. government reporting requirements. The thresholds for reporting a blackout include uncontrolled loss of 300 MW or more of firm system load for more than 15 minutes from a single incident, load shedding of 100 MW or more implemented under emergency operational policy, loss of electric service to more than 50,000 customers for 1 hour or more, and other criteria detailed in the U.S. Department of Energy forms EIA-417 and OE-417. As with all real data, the NERC data have some limitations, including missing and incorrect data. In addition, reporting practices have changed somewhat over time, which may impact observed trends in the data.

B. Transmission Line Outage Data

Transmission owners in the USA are required to report higher voltage transmission line and transformer outage data to NERC for the Transmission Availability Data System (TADS). The TADS data describe the element, time, and cause for major component outages that occur within NERC regions. More than a decade of this type of transmission component outage data is publicly available from BPA [51]. In addition, NERC publishes aggregated quantities based on the TADS data [52]. The TADS outage cause codes include such initiating causes or factors such

as weather, lightning, foreign interference, equipment failure, power system condition, human error, and unknown.

One use for this type of data is to group outages from a period of time (such as a year) into cascades according to the outage times, and then analyze the results to estimate the extent to which outages propagate [25]. The average annual propagation of line outages is a new metric of cascading and can be used to quantify the effect of cascading on the distribution of the number of lines outaged [25]. In the BPA data, this propagation increases as a cascade continues and then appears to level off. Quantifying the way that the propagation of line outages behaves in real data provide a way to validate cascading failure simulations. For example, the observed propagation can be compared to the corresponding propagation of line outages in cascades produced by a simulation [26].

C. Canadian Data

The Canadian Electricity Association (CEA) adopted a proposal to create a facility for centralized collection, processing and reporting of reliability and outage statistics for electrical generation, transmission and distribution equipment in 1975. The transmission segment of the Equipment Reliability Information System (ERIS) program includes transmission equipment outage statistics for equipment with operating voltages of 60 kV and above and was implemented in 1978. Reference [53] indicates two main purposes of data collected in the ERIS system; the first is to assess past performance of typical transmission elements, and the second is to estimate its future performance. CEA outage data statistics were also used to analyze common-mode and dependent outage events in the bulk transmission system [54].

D. WECC Data

The Western Electricity Coordinating Council (WECC) Transmission Reliability Data (TRD) collection system was initiated in 2006 and collects both forced and scheduled outages for all circuits (transmission lines and transformers) configured ≥ 200 kV. The TRD database contains outage data history and inventory data for each WECC participating member utility. The collected data are used to support WECC Reliability Criteria and Performance Category Upgrade Request Process (PCUR) and form the basis for the State of the Interconnection reports produced in WECC in 2012 and 2013 [55]. Detailed analysis of TRD data was performed in [56], [57]. Reference [56] presents concepts associated with the statistical validation of performance indices obtained from outage data and inventory data in the TRD system. Reference [57] presented performance indices of bulk transmission system elements (lines and transformers) with emphasis on common-mode and dependent outage events.

E. Reports on Historical Outages

Since there are many detailed and useful reports on historical outages, we do not summarize them here, but refer to [9], [11], [58]–[60]. It is especially useful to read these reports to get an impression of the variety and complexity of mechanisms involved in cascading. Inspiring examples of reproducing the details of particular outages include [21], [61].

TABLE I
DESCRIPTIVE STATISTICS FOR COMMON PUBLIC POWER SYSTEMS TEST CASES

Test case name /Source	No. of Buses	No. of Lines	No. of Transformers	No. of Gen. Units	Gen. Cap, MW	Load, MW	Line Limits?	Dyn. Data?	Ref.
New Brunswick	85	98	24	395	4,320	3,156	Yes		[62]
IEEE RTS 1979	24	34	5	33	3,405	2,850	Yes	No	[63]
IEEE RTS 1996	73	104	16	99	10,215	8,550	Yes	No[a]	[35]
IEEE 14	14	20	-	5	272	259	No	No	[64]
IEEE 30	30	41	-	6	300	283	No	No	[64]
IEEE 39	39	34	12	10	7,665	6,150	No	Yes	[65],[66]
IEEE 57	57	80	-	7	1,279	1,250	No	No	[64]
IEEE 118	118	177	9	54	3,800	3,668	No	No	[64]
IEEE 162[b]	162	238	46	17	17,440	15,381	No	Yes	[64], [67]
IEEE 300	298	304	105	69	22,930	22,470	No	No	[64]
NETS – NYPS 68	68	66	20	16	17,830	17,620	No	Yes	[68]
Poland system [c]	2,383	2,896	170	327	29,593	24,558	Yes	No	[69]
ICPS [a] 11	11	11	-	1	72	57	Yes	No	[70], [71]
ICPS [a] 13	13	13	-	1	3,016	1,800	Yes	No	[70], [71]
ICPS [a] 43	43	43	-	1	8,012	6,965	Yes	No	[70], [71]
Reduced WECC	200	294	67	31	62,200	53,800	Yes	Yes	[72],[73]
IEEJ East 10 [d]	47	78	22	10	81,430	80,000	No	Yes	[74]
IEEJ East 30	107	136	55	30	73,540	72,600	No	Yes	[74]
IEEJ West 10	47	32	10	10	107,930	107,800	No	Yes	[74]
IEEJ West 30	115	99	30	30	100,460	100,200	No	Yes	[74]
Nordic 32	32	52	17	23	10,738	10,368	No	Yes	[75]

[a] Ref. [76] proposes a set of dynamic data for this case.
[b] There are two variants for this test case, a 162-bus/17-gen system and a 145-bus/50-gen system.
[c] There are seven variations of this test system included with MATPOWER (see Sec. VI.B.3)
[d] The gen. capacity and load for the IEEJ Cases (Japan) in Table I correspond to daytime conditions. Those test cases also include nighttime conditions.

VI. POWER SYSTEM TEST CASES AND CASCADING ANALYSIS

Most cascading failure analyses involve the use of power system test case data. Validating a particular tool will thus usually involve the use of a particular set of test case data, which typically represents a particular power system operating at one or more states. While criteria for modern test cases have been suggested and the desirability of providing access is recognized [50], to our knowledge these criteria have not been applied to existing test case or used to develop new cases. Cases that meet the basic criteria proposed would be candidates for inclusion in a cascade failure benchmark protocol. In this section we discuss a variety of public and non-public test case data sources, suggest how these datasets can be accessed, and (where applicable) their potential for use in cascading failure analysis.

A. Small, Publicly Available Test Cases

A number of test cases have been published and released publically over the last several decades. These test cases were mainly developed as standardized datasets to test and compare results from different approaches and methodologies. Some, but not all, of these cases are useful for certain types of cascading failure analysis. It is important to note that many of the test cases were originally developed in order to benchmark a specific power system problem, other than cascading failure. For example, the IEEE RTS 1996 focuses on system reliability analysis, whereas the IEEE 118- and 300-bus test cases were designed for testing power flow algorithms.

One challenge for the applicability of many of these test cases to cascading failure analysis is network size. Because cascading is inherently a large-scale power systems problem, most types of cascading failure analysis require larger test case (e.g., at least

100 buses). Another problem for many public test cases is the lack of coordinated line rating limits. However, because public test cases facilitate reproducibility, these cases continue to be used for research and development.

For completeness, this section introduces all of the most common public test cases, which are summarized in Table I.

1) The IEEE 1979 and 1996 Reliability Test Systems (RTS): The IEEE 1979 RTS 24-bus test system is a reference network that was extensively used to test or compare methods for system reliability analysis [35], [63]. The IEEE 1996 RTS 73-bus test system interconnects three identical RTS 24-bus test systems [35]. The IEEE 1979 test case has been used to evaluate cascading outage models that include protection system elements (such as relay failure or wide area monitoring) [77]–[80] as well as to assist Monte Carlo type simulations for power system vulnerability assessment [81]–[86]. More recent papers have explored similar topics with the relatively larger IEEE 1996 test case and leveraged its size to illustrate islanding and intelligent control in the context of cascading outages [87]–[92]. It is often a useful starting point for research, given that the case includes line ratings and reliability data, however the system is quite robust by default and thus often requires some modification before cascading failure data can be acquired from the case.

2) The IEEE 14-, 30-, 57-, 118-, 162-, and 300-Bus Test Cases: The IEEE 14-, 30-, 57-, 118-, 162- and 300-Bus Test Cases [64]–[67] represent different snapshots of a portion of the American Electric Power System (in the Midwestern US) as it was in the early 1960's. The 300-bus test case represents a system that interconnects three control areas.

The smaller cases in this group have been used to explore structural vulnerability of power systems, static security margins, and the role of DC systems in cascading failures [93]–[98].

The larger cases have proven useful in order to assist probabilistic approaches to the analysis of cascading failures, interaction models [99]–[105], and intelligent islanding solutions [106], [107]. While these models have been used for cascading failure analysis, the lack of transmission line flow limits mean that limits must be synthesized for cascading failure analysis, which may limit the usefulness of these cases for some types of analysis. There have been recent efforts to include typical dynamic data in most of these models [108].

3) The IEEE 39-Bus Test System: The IEEE 39-Bus Test Case is an approximate representation of the New England 345 kV system [65], [66]. The test system includes dynamic data of the generators with exciters and it was originally developed to explore an energy function analysis for transient stability. This test case has also been endowed with a protection system and used to study hidden failure impact on cascading propagation and to demonstrate intelligent control techniques for vulnerability assessment [106], [107], [109]–[112].

B. Public Test Cases Based on Industry Data

Here we describe industry-grade test cases that can be useful for cascading failure analysis.

1) New Brunswick (NB) Test System: In 1987 CIGRE Study Committee 38 published the Power System Reliability Analysis Application Guide, which describes various reliability approaches, techniques and data requirements [113]. In 1992 the CIGRE Task Force 38-03-10 conducted research based on findings in [62] and compared various software tools for power system reliability analysis using the New Brunswick Power test system. The published report [62] presents a complete example, including the data required, the assumptions made, and the techniques available for the analysis. By 1996 the New Brunswick system was used to compare nine different reliability models with and without these network reinforcements.

2) The NETS-NYPS 68-Bus Test System: The NETS-NYPS 68-Bus Test case [68], [114] represents a reduced order equivalent of the interconnected New England test system (NETS) and New York power system (NYPS). There are five geographical regions. Generators G1 to G9 represent the NETS generation, G10 to G13 represent the NYPS generation, and G14 to G16 are dynamic equivalents of the three neighbor areas connected to the NYPS.

3) The MATPOWER Polish Test Cases: The MATLAB-based toolbox MATPOWER [69] includes some of the IEEE reliability test cases described above and also provides several larger steady-state cases based on the Polish network. Dr. Roman Korab from the Silesian University of Technology originally provided these data. The Polish test cases represent the 110 kV, 220 kV and 400 kV networks for the following snapshots:

- Poland 2003/2004: Winter peak and off-peak (2746 buses), summer peak (2736 buses) and off-peak (2737 buses).
- Poland 2007/2008: Summer peak (3120 buses), winter peak (3375 buses including equivalents for the connections with the German, Czech and Slovak networks).

Because the data are public and because of their relatively large size, these cases have been used by a number of authors for cascading failure analysis (e.g., [36], [115]–[117]). The case was

recently extended for dynamic simulation with synthetic machine data that was generated according to rules based on conversion rules used in the Siemens PSS/E program [32]. This is a notable example of a test case that may be effectively used for a wide variety of cascading failure validation studies.

4) 32-Bus Nordic and CIGRE Test Case: The 32-bus Nordic test system [75], [118] had 23 generators and was originally developed by CIGRE in 1995 to test long-term dynamics and it was later modified to study voltage stability. A detailed description of the CIGRE 32-Bus Test Case can be found in [118]–[120]. There are two different voltage levels, 130 kV and 400 kV, and dynamic and static data can be found in [121].

5) ICPS 11-, 13- and 43-Bus Test Systems: These three Ill-Conditioned Power Systems (ICPS) of 11, 13, and 43 buses are used primarily to test methodologies and programs for solving ill-conditioned systems or determine the existence of load flow solutions [70], [71].

6) WECC Reduced 200-Bus System: This system was used to demonstrate practical use of the Generation Restoration Milestones (GRM) methodology [72] and to examine the effects of replacing conventional generation by wind and solar generation on the grid voltage performance [73].

7) Japanese IEEJ Bulk Power System Models: The Institute of Electrical Engineers Japan (IEEJ) has developed four Japanese test systems [74], which include generator dynamic data. The 50 Hz system models (East 10-machine East 30-machine systems) represent the looping system in the Tokyo area. The 60 Hz system models (West 10-machine and West 30-machine systems) represent the longitudinal grid structure connecting the west area and the east area. These Japanese test systems include two different load conditions (daytime and night time). Table I summarizes the basic information for the daytime conditions.

C. Published Test Cases With Restricted Access

In addition to published cases that have free access, there are some well-known cases with restricted access of one kind or another. Frequently these might be cases used by a particular vendor. A program license or some other permission might be required to access these cases. In other cases these could be government cases, special study group cases, or utility cases where membership or approval of the group is required for access. Due to these limitations, the published works on cascading failure analysis that leverage these test systems are scarce. The following subsections give background information and example applications for the test systems summarized in Table II.

1) GE PSLF Test Cases: These cases are supplied as part of GE PSLF program installation [39].

2) PowerWorld Test Cases: PowerWorld software ships with three test cases (5, 7, and 10 buses). The 5-bus case was used to test a power flow redistribution algorithm designed to mitigate cascading blackouts [122], [123].

3) NPGC Test Case: The NPGC (Northeast Power Grid of China) system consists of Heilongjiang, Jilin, Liaoning, and the northern part of Inner Mongolia of China [103], [124], [125]. The system covers an area of more than $1.2 \times 10^6 \mathrm{km}^2$ and serves more than 100 million people. Most of the hydropower plants are located in the east and most of the thermal power plants

TABLE II
SUMMARY OF TEST CASES WITH RESTRICTED ACCESS

Test case name /Source	No. of buses	No. of lines	No. of xformers	No. of gen. units	Gen. cap. MW	Load MW	Line limits?	Dyn. Data?	Ref
GE PSLF 18	18	18	6	4	3,093	2,949	Yes	Yes	[39]
GE PSLF 56	56	56	37	18	3,349	3,255	Yes	Yes	[39]
PowerWorld 5	5	2	3	2	506	500	Yes	Yes	[122]
PowerWorld 7	7	11	0	5	767	760	Yes	Yes	[122]
PowerWorld 10	10	16	0	7	970	880	Yes	Yes	[122]
NPGC	568	429	264	129	18,497	18,365	Yes	No	[103]
POM 4900	4,911	5,072	922	454	44,531	43,752	Yes	Yes	[127]
TRELSS	2,182	3,311	355	268	147,605	143,536	Yes	No	[44]

are located in the west and Heilongjiang province. The major consumers are in the middle and south of Liaoning province. Hence the power is transmitted from the west and the east to the middle and from the north to the south.

This system has been studied using two different cascading failure simulators: the improved OPA model [124] and the OPA model with slow process [125]. The models are calibrated to obtain blackout frequency similar to the NPGC system. The blackout size distribution of the NPGC system obtained from the two models also matches well a statistical analysis of historical blackout data in China [126]. The NPGC test case has been also used to validate the Galton-Watson branching process model for estimating the statistics of cascades of line outages and discretized load shed [103].

4) POM 4900-Bus Test Case [127]: The POM 4911-bus test case represents the 12 control areas of the Texas Interconnection.

5) TRELSS Test Case [44]: The TRELSS 2182, 12-area bus test case is a reduced 1992 summer case for the eastern USA interconnection. The data were included as part of installation of the EPRI TRELSS Program. Since this dataset was explicitly developed for cascading failure analysis, it is a particularly useful test case for cascading analysis. Reference [128] explores the distribution of initial failures for this test case.

D. Obtaining Real System Models

Ultimately, it is important to test cascading analysis tools on validated representations of real power systems. In some cases, models of power systems used in transmission operation and planning are available. Accessing models in Great Britain, the United States, and Australia are discussed in [50]. Here we further and briefly describe the development and availability of a sample of real system models from around the world.

1) United States: Prior to the attack on the World Trade Center on September 11, 2001, as a part of "open access" in the USA, basic power flow data and maps were available on the public internet for download. While this practice has ceased, significant amounts of data can be obtained through open access, provided proper procedures are followed. Specifically, power flow (positive sequence) grid models, system maps, and switching diagrams are available to anyone who can demonstrate a valid need, is willing to sign a nondisclosure agreement, and passes a Federal Energy Regulatory Commission (FERC) screening process. The ability to access the data is established in the Code of Federal Regulations at 18 C.F.R. § 388.113. The

procedure for obtaining data can be found at [38]. Reliability data can be accessed via the references in Section V.

2) Brazil: The Brazilian power system model used by the system operator and utilities consists of approximately 5000 buses, 7000 branches and 1500 generators. This includes all buses from 230 kV to 750 kV and some lower voltage buses. Peak load is approximately 80 GW. The generation is mainly hydro and typically distant from load centers. Consequently, the system depends on long high voltage transmission lines and DC links. Cascading effects can be mainly triggered by multiple outages of these lines. The integrity of the Brazilian electrical network heavily relies on special protection schemes [129], which are mainly used for load and generation shedding. The power flow and dynamic models are available online in the National System Operator website [130].

The power flow data are formatted for locally used software, but can be easily exported to other applications. On the other hand, most of the dynamic models are user-defined. Although both the model description (control blocks interconnections) and respective parameters are published, exporting these models to other power system applications is not a trivial task. SPS data is currently unavailable. However, the probability of a SPS failure to operate as designed should be taken into account in the cascading modeling as it potentially has a high impact on the reliability of the Brazilian power system. The effect of SPS reliability on the probability of cascading outages has been evidenced by recent blackouts in Europe, such as the Irish disturbance of August 2005 [131] and the Nordic disturbance of December 2005, where a spurious operation of SPS and/or a SPS failure to operate when required contributed to their development.

Reliability indices for the Brazilian power systems from 2007 to 2011 can also be found in the ONS website [130]. These indices include total number of perturbations and the total number and amounts of losses of load.

3) Italy: A model of the Italian transmission power system, which was used to validate the PRACTICE cascading simulator [43], [40]–[42], has been implemented in a time domain simulator used by the Italian TSO [132] by exploiting data from previous research projects. The model represents the EHV (400/220 kV) Italian transmission system of early 2000's with some equivalents for the neighboring countries: it consists of about 1400 electrical nodes, 1000 lines, 700 transformers and 300 generating units. The dynamic models include TSO-customized models for prime movers, control, protection and defense systems and a load model, which captures the typical

behavior of the sub-transmission networks connected to the HV (132/150 kV) side of EHV/HV transformers. Both a peak and an off-peak operating scenario are available.

The above data are available for research purposes inside RSE, which is not authorized by the TSO to let others access them. More recent network models are being created starting from the available data, by adding recent grid reinforcements. Towards this goal, RSE is consulting the grid development plans published by the TSO and publicly available at [133] they provide reliability indicators of the grid over the years, the time schedule of new grid reinforcements and the actual state of progress of scheduled improvements. The same documents also report the connection of new renewable and conventional power plants to the grid. At a European level, a significant source of information is the Ten-Year Network Development Plan of ENTSO-E [134], which describes the major projects to strengthen the European Network in the medium and long term.

4) Building a Real System Model in Europe: Reference [135] proposes an approximate power flow model of the first synchronous area of the interconnected power system of Continental Europe to study the effects of cross-border trades. The model is built by combining a simple knowledge of power system engineering standards and typical values with publicly available information, which includes national generation, peak load, power flow exchange, cross-border line information, generation/substation lists, and geographic information on population and industry from public websites. More recently, the Working Group "System Protection and Dynamics" (SPD) of ENTSO-E developed a Dynamic Study Model (DSM) [136] for the main global dynamic phenomena (frequency transients and oscillation modes) among the areas of the system. A 2020 peak demand case presented in [136] includes 26 areas, 21,382 nodes and 10,829 generators. The DSM uses standard dynamic models for loads, generators and their control devices. The standard dynamic model parameters are tuned using measurement of system events and experts' knowledge. The DSM has some limitations for dynamic analyses of cascading, as it does not include component protections, defense plans or particular control schemes (like over/underexcitation controls, over/underfrequency control), or realistic load dynamic modeling. The model has undergone simplification and anonymization for data security reasons and its use is recommended exclusively under the supervision of a SPD group expert, to balance the need to perform research activities with the need to defend the system against cascading failures potentially triggered by anti-social elements. The Initial Dynamic Models from ENTSO-E can be accessed if a confidentiality agreement is signed.

VII. EMERGING REQUIREMENTS FOR CASCADING ANALYSES

As cascading failure analysis becomes more common in the electricity industry, new requirements for these analyses will certainly emerge. Here we briefly mention several areas where additional improvements are needed in future benchmarking and validation studies.

A. Improved Test Cases

No existing test case provides all of the information that one would ideally want to perform a complete cascading failure benchmark study. There is substantial need for collaborative work to generate new test cases or improve upon existing ones to support a wide range of cascading failure analysis. In the view of the working group, the following data would be particularly valuable in such test cases:

- Generator cost, or other dispatch criteria
- Facility Ratings (power flow limits)
- Protection system/relay data
- Branch outage probability data (see below)
- Breaker failure and bus section fault probabilities
- Detailed node-breaker topology data (see below)
- Power system loading, hazards, and weather

There is a distinct need for test cases that provide probabilistic data and thus allow utilities to explore the potential benefits of probabilistic/risk-based approaches to security and cascading failure analysis [5]. There is also a need for publically available test cases that have been thoroughly evaluated specifically for the problem of cascading failure analysis. The TRELSS [44] test case is a useful, but not easily accessible, example of this. The development and public release of such test cases is an important topic for future work.

B. Use of Cases With Node-Breaker Representation

For many years the majority of system planning studies have used "bus-branch" models. While these models are adequate for most studies they have important limitations. For instance, basic bus-branch data do not enable one to determine the substation breaker configuration, and will thus limit one's knowledge of the system's response to contingencies. An alternative is to use "node-breaker" representations, which are increasingly used for studies of cascading and variable energy resource integration. For example, WECC has begun transferring "node-breaker" models used for state estimation to its TSOs to perform operational studies. This allows one to use the same nomenclature in both offline and online system models, which enables full automation of the creation and processing of contingencies.

C. Wide-Area Protection and Smart-Grid Systems

As smart-grid technology such as phasor measurement units, dynamic line ratings, and real-time demand management systems become more common, there will be an increasing need to model these systems within cascading failure studies. However, wide area protection schemes that make use of these systems can be very complicated; incorporating such systems into cascading failure models is an important topic for future research and development. Moreover the increasing penetration of intermittent generation based on renewable energy sources and the higher frequency of extreme weather events calls for the probabilistic assessment of the power system resilience to these phenomena. In particular, weather data and load forecasting can serve as probabilistic inputs to both on-line and off-line cascading analyses.

VIII. CONCLUSION

Cascading outages, being a combination of many different interactions, is a very complicated problem for which many methods of simulation and analysis are emerging. While each of these tools may produce plausible results and there is some

commonality with respect to producing sequences of potential cascade scenarios, there is no consistency in results and the actionable conclusions are not well determined. The mechanisms that need to be modeled and the required details of the model that are necessary to produce useful and consistent results are not understood. For example, will sequences of steady state solutions produce an adequate result or are dynamics necessary? The required art in simulation is not at all settled, with open questions on the tradeoffs between speed and accuracy, sampling appropriately from the uncertainties, generating plausible cascades, estimating the cascading and blackout severity, and most importantly, what decisions can be justified based on the results. For example, are statistical projections of blackout frequency and extent from simulations adequate to make investment decisions?

Benchmarking and validation are essential to guide and further the current developments in cascading analysis. In this paper, the working group has discussed and surveyed the current state of the art and made recommendations to facilitate progress and good practice in benchmarking and validation. Much of the practical implementation of benchmarking and validation hinges on the available data and test cases. We give a detailed account of the available data in this paper. We also critically and systematically surveyed the international state of the art in cascading failure test cases and indicated key requirements for further improvements. This will enable and encourage the community to access and use these test cases as well as guide further improvements so that cascading failure models, analyses and simulations can be properly tested, benchmarked, and verified.

REFERENCES

[1] *Reliability Terminology*, NERC, Aug. 2013 [Online]. Available: http://www.nerc.com/AboutNERC/Documents/Terms%20AUG13.pdf

[2] IEEE Task Force on Understanding, Prediction, Mitigation and Restoration of Cascading Failures, "Risk assessment of cascading outages: Methodologies and challenges," *IEEE Trans. Power Syst.*, vol. 27, no. 2, pp. 631–641, May 2012.

[3] *Transmission System Planning Performance Requirements*, NERC Standard TPL-001-4, 2013.

[4] *Transmission Relay Loadability*, NERC Standard PRC-023-2, Jul. 2012.

[5] E. Ciapessoni, D. Cirio, A. Pitto, and M. Sforna, "An integrated framework for power and ICT system risk-based security assessment," *Int. J. Eng. Res. Applications*, vol. 4, no. 1, pp. 1–6, Jan. 2014.

[6] K. R. W. Bell, "Issues in integration of risk of cascading outages into utility reliability standards," in *Proc. IEEE Power and Energy Soc. General Meeting*, Detroit, MI, USA, Jul. 24–29, 2011, pp. 1–7.

[7] E. Karangelos, P. Panciatici, and L. Wehenkel, "Whither probabilistic security management for real-time operation of power systems?," in *Proc. IREP Symp. Bulk Power System Dynamics and Control—IX*, 2013, pp. 1–17.

[8] *Transmission System Planned Performance for Geomagnetic Disturbance Events*, NERC Standard TPL-007-1, Dec. 2014.

[9] U.S.-Canada Power System Outage Task Force, U.S. Dept. of Energy and National Resources, Canada "Final report on the August 14th blackout in the United States and Canada,", Apr. 2004.

[10] UCTE, "Final report system disturbance on 4 November 2006," Union for the Co-ordination of Transmission of Electricity, Tech. Rep., 2007.

[11] FERC/NERC, "Arizona-Southern California outages on September 8, 2011: Causes and recommendations," Fed. Energy Regulatory Commission and North American Electric Reliability Corporation, , Tech. Rep., Apr. 2012.

[12] IEEE Task Force on Understanding, Prediction, Mitigation and Restoration of Cascading Failures, "Initial review of methods for cascading failure analysis in electric power transmission systems," in *Proc. IEEE PES General Meeting*, Jul. 2008, pp. 1–8.

[13] Task Force on Understanding, Prediction, Mitigation and Restoration of Cascading Failures, "Survey of tools for risk assessment of cascading outages," in *Proc. IEEE PES General Meeting*, Detroit, MI, USA, 2011, pp. 631–641.

[14] J. J. Dongarra, P. Luszczek, and A. Petitet, "The LINPACK benchmark: Past, present and future," *Concurrency and Computation: Practice and Experience*, vol. 15.9, pp. 803–820, 2003.

[15] *A Guide to the Project Management Body of Knowledge*, 4th ed. Newtown Square, PA, USA: Project Management Inst., 2008.

[16] W. L. Oberkampf and C. J. Roy, *Verification and Validation in Scientific Computing*. Cambridge, U.K.: Cambridge Univ., 2010.

[17] IEEE Task Force Report, "Blackout experiences and lessons, best practices for system dynamic performance, and the role of new technologies," IEEE, Final Report, May 2007.

[18] CIGRE Working Group C1.17, "Planning to manage power interruption events," Paris, Tech. Brochure 433, Oct. 2010, CIGRE.

[19] P. Hines, E. Cotilla-Sanchez, and S. Blumsack, "Do topological models provide good information about vulnerability in electric power networks?," *Chaos*, vol. 20, no. 3, 2010, Art. ID 033122.

[20] M. Morgan and M. Henrion, *Uncertainty: A Guide to Dealing With Uncertainty in Quantitative Risk and Policy Analysis*. Cambridge, U.K.: Cambridge Univ., 1990.

[21] D. Kosterev, C. Taylor, and W. Mittelstadt, "Model validation for the august 10, 1996 WSCC system outage," *IEEE Trans. Power Syst.*, vol. 14, no. 3, pp. 967–979, Aug. 1999.

[22] B. A. Carreras, D. E. Newman, I. Dobson, and A. B. Poole, "Evidence for self-organized criticality in electric power system blackouts," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 51, no. 9, pp. 1733–1740, Sep. 2004.

[23] P. Hines, J. Apt, and S. Talukdar, "Large blackouts in North America: Historical trends and policy implications," *Energy Policy*, vol. 37, no. 12, pp. 5249–5259, 2009.

[24] Supplementary information for "Large blackouts in North America: Historical trends and policy implications," [Online]. Available: http://www.uvm.edu/~phines/publications/2009/blackouts.html

[25] I. Dobson, "Estimating the propagation and extent of cascading line outages from utility data with a branching process," *IEEE Trans. Power Syst.*, vol. 27, no. 4, pp. 2146–2155, Nov. 2012.

[26] B. A. Carreras, D. E. Newman, B. A. Carreras, I. Dobson, and N. S. Degala, "Validating OPA with WECC data," in *Proc. 46th Hawaii Int. Conf. Syst. Sci.*, Maui, HI, USA, Jan. 2013, pp. 2197–2204.

[27] I. Dobson, B. A. Carreras, V. E. Lynch, and D. E. Newman, "Complex systems analysis of series of blackouts: Cascading failure, critical points, and self-organization," *Chaos*, vol. 17, June 2007, Art. ID 026103.

[28] S. S. Miller, "Extending traditional planning methods to evaluate the potential for cascading failures in electric power grids," in *Proc. IEEE PES General Meeting*, Jul, 2008, pp. 1–7.

[29] M. J. Eppstein and P. D. H. Hines, "A "random chemistry" algorithm for identifying collections of multiple contingencies that initiate cascading failure," *IEEE Trans. Power Syst.*, vol. 27, no. 3, pp. 1698–1705, Aug. 2012.

[30] MUCM Project, "Managing uncertainty in complex models," [Online]. Available: http://www.mucm.ac.uk/

[31] R. Fitzmaurice, E. Cotilla-Sanchez, and P. Hines, "Evaluating the impact of modeling assumptions for cascading failure simulation," in *Proc. IEEE PES General Meeting*, 2012, pp. 1–8.

[32] J. Song, E. Cotilla-Sanchez, P. Hines, and G. Ghanavati, "Dynamic modeling of cascading failure in power systems," *IEEE Trans. Power Syst.*, vol. PP, no. 99, 2015, DOI: 10.1109/TPWRS.2015.2439237.

[33] IEEE Cascading Failure Working Group Web Site [Online]. Available: http://sites.ieee.org/pes-camscftf

[34] B. A. Carreras, D. E. Newman, and I. Dobson, "Does size matter?," *Chaos*, vol. 24, no. 2, 2014, Art. ID 023104.

[35] IEEE Task Force, "The IEEE reliability test system—1996," *IEEE Trans. Power Syst.*, vol. 14, no. 3, pp. 1010–1020, Aug. 1999.

[36] P. Rezaei, P. Hines, and M. Eppstein, "Estimating cascading failure risk with random chemistry," *IEEE Trans. Power Syst.*, vol. 30, no. 5, pp. 2726–2735, Sep. 2015.

[37] "North American Electric Reliability Corporation, Transmission Availability Data System Automatic Outage Metrics and Data," 2009 Report, June 14, 2010 [Online]. Available: www.nerc.com

[38] Critical Energy Infrastructure Information (CEII) [Online]. Available: http://www.ferc.gov/legal/ceii-foia/ceii.asp

[39] General Electric (PSLF) [Online]. Available: http://site.ge-energy.com/

[40] E. Ciapessoni, D. Cirio, and A. Pitto, "Cascading simulation techniques in Europe: The PRACTICE experience," in *Proc. IEEE PES General Meeting*, Vancouver, BC, Canada, Jul. 20–25, 2013.

[41] E. Ciapessoni, D. Cirio, and A. Pitto, "Cascadings in large power systems: Benchmarking static vs. time domain simulation," in *Proc. IEEE PES General Meeting*, Jul. 27–31, 2014, pp. 1–5.

[42] E. Ciapessoni, D. Cirio, E. Gaglioti, S. Massucco, A. Pitto, and F. Silvestro, "Risk evaluation in power system contingency analyses," in *Proc. CIGRE Session*, Paris, France, Aug. 24–29, 2010.

[43] E. Ciapessoni, D. Cirio, S. Massucco, and A. Pitto, "A risk-based methodology for operational risk assessment and control of power systems," in *Proc. 17th Power Syst. Computation Conf.*, Stockholm, Sweden, Aug. 2011.

[44] Electric Power Research Institute (TRELSS) [Online]. Available: http://www.epri.com/

[45] R. C. Hardiman, M. T. Kumbale, and Y. V. Makarov, "An advanced tool for analyzing multiple cascading failures," in *Proc. 8th Int. Conf. Probability Methods Applied to Power Systems*, Ames, IA, USA, Sep. 2004, pp. 629–634.

[46] M. Morgan *et al.*, California Energy Commission, Extreme Events Phase 2. 2011, publication number: CEC-MR-08-03.

[47] *System Performance Following Loss of Two or More BES Elements*, NERC Standard Standard TPL-003-0.

[48] N. Bhatt *et al.*, "Assessing vulnerability to cascading outages," in *Proc. PSCE*, Seattle, WA, USA, 2009, pp. 1–9.

[49] M. Koenig, P. Duggan, J. Wong, M. Y. Vaiman, M. M. Vaiman, and M. Povolotskiy, "Prevention of cascading outages in con edison's network," in *Proc. IEEE PES T&D Conf. Expo.*, 2010, pp. 1–7.

[50] K. R. W. Bell and A. N. D. Tleis, "Test system requirements for modelling future power systems," in *Proc. IEEE PES General Meeting*, Jul. 2010, pp. 1–8.

[51] BPA, "Bonneville power administration transmission services operations & reliability website," Bonneville Power Administration Transmission Services Operations [Online]. Available: http://transmission.bpa.gov/Business/Operations/Outages [Online]. Available:

[52] NERC, "Transmission Availability Data System, Automatic Outage Metrics and Data," Tech. Rep., Jun. 14, 2010.

[53] R. Billinton, "Canadian experience in the collection of transmission and distribution component unavailability data," in *Proc. Conf. Probabilistic Methods Applied to Power Systems*, Ames, IA, USA, 2004, pp. 268–273.

[54] J. Schaller, "Common mode event perspectives from the canadian electricity association equipment reliability information system," in *Proc. IEEE PES General Meeting*, San Diego, CA, USA, 2012, pp. 1–6.

[55] "2012 state of the interconnection," Western Electricity Coordinating Council 2012 [Online]. Available: www.wecc.biz

[56] M. Papic, J. J. Bian, and S. Ekisheva, "A novel statistical-based analysis of WECC bulk transmission reliability data," in *Proc. IEEE PES General Meeting*, Vancouver, BC, Canada, 2013, pp. 1–5.

[57] B. Keel, M. Papic, and D. Tucker, "Western electricity coordinating council experience in the collection of transmission common-mode and dependent outages," in *Proc. IEEE PES General Meeting*, San Diego, CA, USA, 2012, pp. 1–8.

[58] Union for the Co-ordination of Transmission of Electricity (UCTE) Final Rep., Nov. 4, 2006 [Online]. Available: www.ucte.org

[59] North American Electric Reliability Council, "1996 System Disturbances," 2002.

[60] "IEEE PES PSDP task force on blackout experience, mitigation, and role of new technologies, blackout experiences and lessons," Best Practices for System Dynamic Performance, and the Role of New Technologies Jul. 2007, IEEE Special Publication 07TP190.

[61] V. Venkatasubramanian and Y. Li, "Analysis of 1996 western american electric blackouts," in *Proc. Bulk Power Syst. Dynamics and Control—VI*, Cortina d'Ampezzo, Italy, Aug. 2004.

[62] CIGRE Task Force 38-03-10, "Power system reliability analysis, vol. 2 composite power system reliability evaluation," in *Proc. CIGRE*, Paris, France, 1992.

[63] IEEE Task Force, "IEEE reliability test system," *IEEE Trans. Power App. Syst.*, vol. PAS-98, pp. 2047–2054, Nov./Dec. 1979.

[64] IEEE (AEP), "Power system test case archive," [Online]. Available: http://www.ee.washington.edu/research/pstca/

[65] T. Athay, R. Podmore, and S. Virmani, "A practical method for the direct analysis of transient stability," *IEEE Trans. Power App. Syst.*, vol. PAS-98, no. 2, pp. 573–584, 1979.

[66] M. A. Pai, *Energy Function Analysis for Power System Stability*. Boston, MA, USA: Kluwer Academic, 1989.

[67] The IEEE Stability Test Systems Task Force of The Dynamic System Performance Subcommittee, "Transient stability test systems for direct stability methods," *IEEE Trans. Power Syst.*, vol. 7, no. 1, pp. 37–43, Feb. 1992.

[68] G. Rogers, *Power System Oscillations*. Boston, MA, USA: Kluwer Academic, 2000.

[69] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, "Matpower: Steady-state operations, planning and analysis tools for power systems research and education," *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 12–19, Feb. 2011.

[70] S. C. Tripathy, G. D. Prasad, O. P. Malik, and G. S. Hope, "Load-flow solutions for ill-conditioned power systems by a newton-like method," *IEEE Trans. Power App. Syst.*, vol. PAS-101, no. 10, pp. 3648–3657, Oct. 1982.

[71] K. Okumura, K. Terai, and A. Kishima, "Solution of ill-conditioned load flow equation by homotopy continuation method," in *Proc. IEEE Int. Symp. Circuits Syst.*, Jun. 1991, pp. 2897–2899.

[72] Y. Hou *et al.*, "Computation of milestones for decision support during system restoration," *IEEE Trans. Power Syst.*, vol. 26, no. 3, pp. 1399–1409, Aug. 2011.

[73] K. Yagnik and V. Ajjarapu, "Consideration of the wind and solar generation reactive power capability on grid voltage performance," in *Proc. IEEE PES General Meeting*, 2012, pp. 1–7.

[74] Institute of Electrical Engineers of Japan [Online]. Available: http://www2.iee.or.jp/ver2/pes/23-st_model/english/index020.html

[75] T. Van Cutsem *et al.*, "Description, Modeling and Simulation Results of a Test System for Voltage Stability Analysis," IEEE Working Group on Test Systems for Voltage stability analysis, Tech. Rep., Jul. 2010, Version 1.

[76] K. Walve, "Nordic32-A CIGRÉ Test System for Simulation of Transient Stability and Long Term Dynamics," Svenska Kraftnat Tech. Rep., 1993.

[77] Z. Bie and X. Wang, "Evaluation of power system cascading outages," in *Proc. PES General Meeting*, 2002, vol. 1, pp. 415–419.

[78] X. Yu and C. Singh, "A practical approach for integrated power system vulnerability analysis with protection failures," *IEEE Trans. Power Syst.*, vol. 19, no. 4, pp. 1811–1820, Nov. 2004.

[79] M. Zima and G. Andersson, "Wide area monitoring and control as a tool for mitigation of cascading failures," in *Proc. 8th Int. Conf. Probabilistic Methods Applied to Power Syst.*, Sep. 2004, pp. 663–669.

[80] D. S. Kirschen, D. Jayaweera, D. Nedic, and R. N. Allan, "A probabilistic indicator of system stress," *IEEE Trans. Power Syst.*, vol. 19, no. 3, pp. 1650–1657, Aug. 2004.

[81] F. Yang, S. Meliopoulos, G. Cokkinides, and Q. B. Dam, "Bulk power system reliability assessment considering protection system hidden failures," in *Proc. IREP Symp.*, Aug. 2007, pp. 3408–3421.

[82] E. Ciapessoni, S. Grillo, S. Massucco, A. Pitto, and F. Silvestro, "Operational risk assessment and control: A probabilistic approach," in *Proc. ISGT Europe*, 2010, pp. 1–8.

[83] A. Delgadillo, J. M. Arroyo, and N. Alguacil, "Power system defense planning against multiple contingencies," in *Proc. 17th Power Syst. Computation Conf.*, Aug. 2011, pp. 1–7.

[84] M. Panteli, P. A. Crossley, D. S. Kirschen, and D. Sobajic, "Assessing the impact of insufficient situation awareness on power system operation," *IEEE Trans. Power Syst.*, vol. 28, no. 3, pp. 2967–2977, Aug. 2013.

[85] X. Li and H. Wang, "Operation risk assessment of wind farm integrated system influenced by weather conditions," in *Proc. IEEE PES Gen Meeting*, 2013, pp. 1–5.

[86] Q. Chen and L. Mili, "Composite power system vulnerability evaluation to cascading failures using importance sampling and antithetic variates," *IEEE Trans. Power Syst.*, vol. 28, no. 3, pp. 2967–2977, Aug. 2013.

[87] Q. Chen, Y. Lin, and J. D. McCalley, "The risk of high-order transmission contingencies," in *Proc. IEEE PES General Meeting*, 2007, pp. 1–7.

[88] N. Zaag, J. F. Restrepo, H. Banakar, and F. D. Galiana, "Analysis of contingencies leading to islanding and cascading outages," in *Proc. IEEE PES General Meeting*, 2007, pp. 63–67.

[89] D. A. Panasetsky and N. I. Voropai, "A multi-agent approach to co-ordination of different emergency control devices against voltage collapse," in *Proc. IEEE Bucharest Power Tech Conf.*, Jun. 2009, pp. 1–7.

[90] E. Johansson, K. Uhlen, G. Kjolle, and T. Toftevaag, "Reliability evaluation of wide area monitoring applications and extreme contingencies," in *Proc. 17th Power Syst. Computation Conf.*, Aug. 2011, pp. 1–7.

[91] N. I. Voropai, V. G. Kurbatsky, N. V. Tomin, and D. A. Panasetsky, "Preventive and emergency control of intelligent power systems," in *Proc. 3rd ISGT*, 2012, pp. 1–7.

[92] M. Negnevitsky, N. Voropai, V. Kurbatsky, N. Tomin, and D. Panasetsky, "Development of an intelligent system for preventing large-scale emergencies in power systems," in *Proc. IEEE PES General Meeting*, 2013, pp. 1–8.

[93] E. Zio, C. A. Petrescu, and G. Sansavini, "Vulnerability analysis of a power transmission system," in *Proc. Int. Probabilistic Safety Assessment and Manag. Conf.*, May 2008, pp. 1–8.

[94] C. J. Kim and O. B. Obah, "Vulnerability assessment of power grid using graph topological indices," *Int. J. Emerging Electric Power Syst.*, vol. 8, no. 6, 2007.

[95] A. Rabiee and M. Parniani, "Optimal reactive power dispatch using the concept of dynamic VAR source value," in *Proc. IEEE PES General Meeting*, 2009, pp. 1–5.

[96] Y. Koc, M. Warnier, R. E. Kooij, and F. M. T. Brazier, "A robustness metric for cascading failures by targeted attacks in power networks," in *Proc. 10th IEEE Int. Conf. Networking, Sensing and Control*, 2013, pp. 48–53.

[97] W. Lian, L. Zhang, S. Zhao, and D. Zhang, "A method for cascading failure simulation based on static security analysis," in *Proc. IEEE PES General Meeting*, 2009, pp. 1–7.

[98] Y. Zhong, X. Zhang, S. Huang, X. Yu, Z. Zhang, and Q. Zhu, "Cascading failure model of AC-DC system and blackout mechanism analysis," in *Proc. IEEE PES General Meeting*, 2014, pp. 1–5.

[99] B. H. Chowdhury and S. Baravc, "Creating cascading failure scenarios in interconnected power systems," in *Proc. IEEE PES General Meeting*, 2006.

[100] M. Rahnamay-Naeini, Z. Wang, A. Mammoli, and M. M. Hayat, "Impacts of control and communication system vulnerabilities on power systems under contingencies," in *Proc. IEEE PES General Meeting*, 2012, pp. 1–7.

[101] M. Rahnamay-Naeini, Z. Wang, A. Mammoli, and M. M. Hayat, "A probabilistic model for the dynamics of cascading failures and blackouts in power grids," in *Proc. IEEE PES General Meeting*, 2012, pp. 1–8.

[102] Y. Wang and R. Baldick, "Case study of an improved cascading outage analysis model using outage checkers," in *Proc. IEEE PES Gen. Meeting*, 2013, pp. 1–5.

[103] J. Qi, I. Dobson, and S. Mei, "Towards estimating the statistics of simulated cascades of outages with branching processes," *IEEE Trans. Power Syst.*, vol. 28, no. 3, pp. 3410–3419, Aug. 2013.

[104] J. Qi and K. Sun, "Estimating the propagation of several cascading outages with multi-type branching processes," *Cornell Univ. Library Phys. Soc.*, vol. 1, pp. 1–8, May 2014.

[105] J. Qi, K. Sun, and S. Mei, "An interaction model for simulation and mitigation of cascading failures," *IEEE Trans. Power Syst.*, vol. 30, no. 2, pp. 804–819, May 2014.

[106] L. Ding, F. M. Gonzalez-Longatt, P. Wall, and V. Terzija, "Two-step spectral clustering controlled islanding algorithm," *IEEE Trans. Power Syst.*, vol. 28, no. 1, pp. 75–84, Feb. 2013.

[107] J. Quirós-Tortós, R. Sánchez-García, J. Brodzki, J. Bialek, and V. Terzija, "Constrained spectral clustering based methodology for intentional controlled islanding of large-scale power systems," *IET Gener. Transm. Distrib.*, vol. 9, no. 1, pp. 31–42, Jan. 2015.

[108] P. Demetriou *et al.*, "Dynamic IEEE test systems for transient analysis," *IEEE Syst. J.*, vol. PP, no. 99, pp. 1–7, Jul. 2015.

[109] A. G. Phadke and J. S. Thorp, "Expose hidden failures to prevent cascading outages," *IEEE Comput. Applications in Power*, pp. 2–23, Jul. 1996.

[110] A. Tiwari and V. Ajjarapu, "Contingency assessment for voltage dip and short term voltage stability analysis," in *Proc. IEEE PES General Meeting*, 2007, pp. 1–8.

[111] S. F. Mahdavizadeh, M. A. Sandidzadeh, and M. R. Aghamohammadi, "Dynamic classification of power systems via Kohonen neural network," *Int. J. Comp. Eng. Res.*, vol. 2, no. 1, pp. 46–55, Jan.–Feb. 2012.

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

[113] CIGRE WG 03 of SC 38, "Power system reliability analysis. Application guide," in *Proc. CIGRE*, Paris, France, 1987.

[114] B. Pal and B. Chaudhuri, *Robust Control in Power Systems*. New York, NY, USA: Springer, 2005.

[115] 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," [Online]. Available: http://arxiv.org/abs/1508.01775 preprint arXiv:1508.01775 [physics.soc-ph];

[116] P. A. Kaplunovich and K. S. Turitsyn, "Statistical properties and classification of N-2 contingencies in large scale power grids," in *Proc. 47th IEEE Hawaii Int. Conf. Syst. Sci.*, 2014, pp. 2517–2526.

[117] J. Quirós-Tortós *et al.*, "A sectionalizing methodology for parallel system restoration based on graph theory," *IET Gener. Transm. & Distrib.*, vol. 9, no. 11, pp. 1216–1225, Aug. 2015.

[118] M. C. Stubbe, Long Term Dynamics Phase II Final Report Cigre, Tech. Rep. Task Force 38.08.08, Mar. 1995.

[119] G. Demartini *et al.*, "Co-ordinated economic and advance dispatch procedures," *IEEE Trans. Power Syst.*, vol. 11, no. 4, pp. 1785–1791, Nov. 1996.

[120] "Long Term Dynamics, Phase II," Final Report, CIGRÉ TF 38.02.08 Tech. Rep., 1995, CE/SC 38 GT/WG 02, Ref. No. 102.

[121] O. Samuelsson, "Power system damping-structural aspects of controlling active power," Ph.D. dissertation, Dept. Ind. Electr.. Eng. Autom., Lund Inst. Tech., Lund, Sweden, 1997.

[122] PowerWorld Corporation [Online]. Available: http://www.power-world.com/products/simulator/overview

[123] A. A. Yusuff, A.-G. A. Jimoh, and J. Munda, "Mitigation of cascade blackout in power systems by using widest path and power flow redistribution algorithms," in *Proc. Int. Conf. Advances Energy Eng.*, Jun. 2010, pp. 129–133.

[124] S. Mei, F. He, X. Zhang, S. Wu, and G. Wang, "An improved OPA model and blackout risk assessment," *IEEE Trans. Power Syst.*, vol. 24, no. 2, pp. 814–823, May 2009.

[125] J. Qi, S. Mei, and F. Liu, "Blackout model considering slow process," *IEEE Trans. Power Syst.*, vol. 28, no. 3, pp. 3274–3282, Aug. 2013.

[126] X. Weng, Y. Hong, A. Xue, and S. Mei, "Failure analysis on China power grid based on power law," *J. Control Theory and Applications*, vol. 4, no. 3, pp. 235–238, Aug. 2006.

[127] V&R Energy Systems Research [Online]. Available: http://www.vren-ergy.com/index.php/powersystemsoftware/pom.html

[128] F. Xia and A. Meliopoulos, "A method for probabilistic simultaneous transfer capability analysis," in *Proc. IEEE/PES Summer Meeting*, Portland, OR, USA, Jul. 1995, Paper 95 SM 514-0 PWRS.

[129] P. Gomes and G. Cardoso, Jr., *Reducing Blackout Risk by System Protection Schemes—Detection and Mitigation of Critical System Conditions*. : CIGRE Publications, 2006, C2-201.

[130] [Online]. Available: http://www.ons.org.br/avaliacao_condicao/casos_eletromecanicos.aspxPower flow cases can be downloaded from http://www.ons.org.br/plano_ampliacao/casos_ refer_ regime_ permanente.aspx (Arquivo no formato cartão), or from http://www.ons.org.br/avaliacao_ condicao/c_referencia_ perma-nente_anual.aspx (those under Formato Cartão) respectively for future and past scenarios. The dynamic models can be downloaded from

[131] ESB and National Grid, "Report on investigation into system disturbance of august 5th 2005," Dec. 2005.

[132] C. Bruno *et al.*, "CRESO and SICRE: Modern environments for power system static and dynamic analysis," in *Proc. PowerGrid Europe*, Cologne, Germany, May 26–28, 2009.

[133] [Online]. Available: www.terna.it

[134] European Network of Transmission System Operators for Electricity (ENTSO-E), "Ten-Year Network Development Plan," , , Report, Jul. 2012.

[135] Q. Zhou and J. W. Bialek, "Approximate model of European interconnected system as a benchmark system to study effects of cross-border trades," *IEEE Trans. Power Syst.*, vol. 20, no. 2, pp. 782–788, May 2005.

[136] A. Semerow *et al.*, "Dynamic study model for the interconnected power system of continental Europe in different simulation tools," in *Proc. IEEE Power Tech*, Jun. 2015, pp. 1–6.