Transmission-Capacity Expansion for Minimizing Blackout Probabilities

John Shortle
Systems Engineering & Operations Research, George Mason University, Fairfax, VA 22030
Steffen Rebennack
Division of Economics & Business, Colorado School of Mines, Golden, CO 80401
Fred W. Glover
OptTek Systems, Inc., Boulder, CO 80302

***Abstract*—The objective of this paper is to determine an optimal plan for expanding the capacity of a power grid in order to minimize the likelihood of a large cascading blackout. Capacity-expansion decisions considered in this paper include the addition of new transmission lines and the addition of capacity to existing lines. We embody these interacting considerations in a simulation optimization model, where the objective is to minimize the probability of a large blackout subject to a budget constraint. The probability of a large-scale blackout is estimated via Monte-Carlo simulation of a probabilistic cascading blackout model. Because the events of interest are rare, standard simulation is often intractable from a computational perspective. We apply a variance-reduction technique within the simulation to provide results in a reasonable time frame. Numerical results are given for some small test networks including an IEEE 14-bus test network. A key conclusion is that the different expansion strategies lead to different shapes of the *tails* of the blackout distributions. In other words, there is a trade-off between reducing the frequency of small-scale blackouts versus reducing the frequency of large-scale blackouts.**

***Index Terms*—Transmission expansion, rare-event simulation, splitting, simulation optimization, cascading blackouts.**

I. Introduction

The objective of this paper is to determine an optimal plan for expanding the transmission capacity of a power grid in order to minimize the likelihood of a large cascading blackout. Because simulation efficiency is a concern when rare events are considered, rare-event simulation techniques are used within the simulation optimization to yield meaningful results within a reasonable computing time.

There is a rich body of literature on transmission expansion and planning. For a survey of issues, see [1], [2], [3]. We do not attempt to review all literature here, but rather cite some representative papers related to mathematical optimization of transmission expansion. The unique aspect of our work is that the formulation utilizes a rare-event simulation of cascading blackouts.

A typical formulation for capacity expansion minimizes cost subject to demand constraints. The decision variables are the added capacities of the transmission lines. *Perfect reliability* of the system is often assumed via the demand constraints. That is, these constraints enforce that line capacities exceed the line loads based on an assumed demand profile, so line failures and cascading blackouts implicitly do not occur. Alternatively, some formulations allow demand to exceed capacity, but unmet demand is curtailed (line failures do not occur), and the unmet demand is counted against the objective function. Examples of these types of models include [4], [5], [6], [7], [8], [9]. Other formulations generalize the demand constraints to an "$N-1$" policy. That is, the constraints specify that line capacities be at least as large as the carried load following any possible single component failure (cf. [6], [10]).

A variety of mathematical techniques have been used to solve these types of optimization problems, such as non-linear programming [11], mixed-integer programming [5], ordinal optimization [12] (which identifies solutions that are "good enough"), heuristic methods [9], and so forth. Other extensions include evaluation of multiple objectives [1], [6], [8], [13], vulnerability analysis [14], and long-term system dynamics [10], [15]. Grid reliability has also been analyzed via Markov-process models of equipment (e.g., [16], [17]). These models consider failures, maintenance, and repairs of individual equipment, but typically do not consider the dynamics of cascading blackout dynamics – how equipment failures impact power flows leading to downstream failures.

The unique aspect of this work is the manner in which we treat reliability within the optimization. Here, we allow for the possibility of line failures and subsequent cascading dynamics. We model these as possible, yet unlikely, events. Because the associated probabilities of interest may be quite small, say one in 10^6, estimating these probabilities can be challenging from a computational perspective (Section III-B). Furthermore, within the context of optimization, rare-event probabilities must be evaluated multiple times over a large number of design alternatives.

We specifically pay attention to the tails of the distributions. [18] notes that evaluation of the *entire* blackout distribution is important: Some mitigation strategies reduce the likelihood of small-scale blackouts but increase the likelihood of large-scale blackouts, and vice-versa. Thus, the overall distribution must be considered in determining whether or not a policy has an overall benefit or not – not just a mean value. Historically, the frequency of blackouts (e.g., those affecting $10,000$ or more customers) has followed a *power-law* distribution [19]. This means that the frequency of outages affecting y or more customers (where y is large) is roughly proportional to $y^{-\alpha}$, for some constant $\alpha > 0$. Compared with, for example, a normal distribution, a power-law distribution implies a much greater likelihood of an extremely large event. This paper considers the direct *short-term* impact of upgrades. While the tails of the blackout distribution can be modified in the short term by expanding the network, [18] suggests that this may be difficult to sustain in the *long term* because self organization drives the system to evolve towards a power tail.

The basic formulation of our capacity-expansion problem is to minimize the probability of a large-scale blackout subject to cost constraints (see Section II). To evaluate the probability of a large-scale blackout, we implement a stochastic model of blackout dynamics [20] similar to a number of models in the literature (cf. [21], [22]). The model treats line failures as instantaneous events: When one line fails, the power flow is instantly redistributed throughout the network; the resulting flows may result in new line failures, and the process repeats until there are no new line failures.

As noted in [23], when considering rare events, it is very difficult to fully model all relevant details of cascading dynamics. The philosophy here is to model representative components in a simple way with a goal of understanding higher level effects. A similar philosophy is used in other works such as [15], [21], [22], [24]. Here we consider basic elements such as network topology, network characteristics (bus loads, physical line characteristics, etc.), probabilistic line failures (that depend on the line load and the rated line capacity), physics of power flows, and an "$N-1$" rule. Many other complexities are not considered – for example, transient effects, human error, long-term dynamics (over several years), timing of events (e.g., the time it takes for an overloaded line to rise in temperature), weather, instabilities, nonlinear dynamics, and so forth. Thus, we acknowledge that this is not a full-scale model but still believe the model is useful for gaining insight into rare-event dynamics of more complex systems.

The main contributions of this paper are the following: First, we formulate the capacity expansion problem with respect to the minimization of a rare-event blackout probability. Rather than treating reliability as a *deterministic* demand constraint or objective-function cost (cf. [4]–[10]), we explicitly model the stochastic nature of blackout dynamics. Cascading line failures are modeled as possible, but unlikely, events within a discrete-event simulation. Second, we integrate a rare-event simulation technique within a simulation-optimization framework. While the simulation-optimization approach [25] and the rare-event model [26] are not new, the integration of these two approaches is new. Third, numerical experiments provide insights on how capacity-expansion decisions impact rare-event blackout probabilities over the *complete* cumulative distribution function (CDF) of blackout size. Based on the model, large-scale blackouts tend to be avoided by focusing resources on strengthening *one* path to each node in the network (i.e., a tree-like structure). In contrast, small blackouts tend to be avoided by solutions that strengthen multiple paths to each node (e.g., something closer to a mesh structure). Fourth, we provide an analytical model that explains numerical results observed in the full model.

II. Optimal Capacity Expansion Problem

The problem we solve takes the following general form:[1]

$$\begin{aligned} &\text{minimize:} && \Pr\{\text{blackout size} \geq \text{threshold}\}, \\ &\text{subject to:} && \text{expansion cost} \leq \text{budget}. \end{aligned} \tag{1}$$

"Blackout size" can be measured in a number of ways such as the number of failed transmission lines or the total load shed. Here, we measure blackout size with respect to the total load shed, which provides a more direct measure of the impact to customers. Estimating the objective function, Pr{blackout size ≥ threshold}, is accomplished via Monte-Carlo simulation of a probabilistic blackout model (see Section III-A). When the blackout-size threshold is large, rare-event techniques are needed to efficiently estimate the objective function (see Section III-B).

To specify the optimization problem in (1), we first give some definitions. The existing baseline network is defined by the following parameters:

- n number of buses (nodes) in network,
- p_i real load [MW] at bus i,
- d_{ij} Euclidean distance between buses i and j,
- b_{ij} imaginary part of element (i,j) of admittance matrix,
- a_{ij} existing transmission capacity for line (branch) i–j,
- $\mathbb{S}$ set of potential lines to add to network $= \{(i,j)\,|\,\text{no line exists between buses } i \text{ and } j\}$.

We assume throughout this section that $i<j$. Expansion costs are assumed to consist of fixed costs and variable costs. The associated parameters are:

- c_f fixed cost to add a line [per distance]
- c_v variable cost to add capacity [per MW per distance]
- r minimum capacity of a new line [MW]
- T total capacity-expansion budget

The *decision variables* of the capacity-expansion problem are:

- $\mathbf{B}_{ij}$ new line connecting buses i and j (0=no, 1=yes),
- $\mathbf{A}_{ij}$ additional transmission capacity [MW] for line i–j.

The specific formulation of the general problem in (1) is:

$$\text{minimize:} \quad \Pr\{\text{blackout size} \geq \text{threshold}\,|\,\mathbf{A}_{ij}, \mathbf{B}_{ij}\}, \tag{2}$$

subject to:

$$\sum_{(i,j)\in\mathbb{S}} c_f d_{ij} \mathbf{B}_{ij} + \sum_{i=1}^{n} \sum_{j=i+1}^{n} c_v d_{ij} \mathbf{A}_{ij} \leq T \tag{3}$$

$$\mathbf{B}_{ij} r \leq \mathbf{A}_{ij} \leq \mathbf{B}_{ij} R_{ij}, \qquad (i,j)\in\mathbb{S} \tag{4}$$

$$\mathbf{B}_{ij} \in \{0,1\}, \qquad (i,j)\in\mathbb{S} \tag{5}$$

$$\mathbf{A}_{ij} \geq 0, \qquad 1 \leq i < j \leq n. \tag{6}$$

The probability in the objective function (2) depends on the decision variables $\mathbf{A}_{ij}$ and $\mathbf{B}_{ij}$ (it also implicitly depends

[1] An alternate formulation is to minimize cost subject to a blackout constraint. We choose the approach in (1) since the optimization engine (OptQuest) expects a computationally expensive black box to evaluate the objective function.

on the other model parameters such as n, p_i, c_f, etc.). The relationship between this probability and $\mathbf{A}_{ij}$ and $\mathbf{B}_{ij}$ is complex. We do not specify this relationship explicitly, but rather define it implicitly via the model and assumptions discussed in Section III-A. Estimates for the objective function are obtained by running Monte-Carlo simulations of the model with $\mathbf{A}_{ij}$ and $\mathbf{B}_{ij}$ as inputs. Accuracy of the simulation output is controlled by the optimization engine (OptQuest, see Section III). The engine conducts multiple replications of the objective function at each design point to achieve a specified confidence level in the output.

Equation (3) is the budget constraint. Capacity-expansion costs are computed as follows: If there is an existing line connecting buses i and j, then the cost of increasing this line by capacity $\mathbf{A}_{ij}$ is $c_v d_{ij} \mathbf{A}_{ij}$. If there is not an existing line connecting buses i and j, then the cost of adding a new line with capacity $\mathbf{A}_{ij}$ is $c_f d_{ij} + c_v d_{ij} \mathbf{A}_{ij}$.

Equation (4) enforces relationships between the binary and continuous variables. The right-hand side states that line capacity cannot be added to a non-existent line; that is, if $\mathbf{B}_{ij} = 0$, then $\mathbf{A}_{ij}$ must also be 0 ($R_{ij} \equiv T/c_v d_{ij}$ is an upper bound on the capacity expansion for line i–j if the entire budget were devoted to that line). The left-hand side states that a new line must have a minimum capacity r.

Some implicit assumptions of the model are: (1) Generator and line costs are independent of node location. This assumption can easily be relaxed by redefining the cost parameters with more subscripts; the formulation structure is still the same. (2) Line capacities can be added in arbitrarily fine increments. This can be relaxed by introducing binary or integer variables. (3) Line-capacity enhancements do not impact the susceptance and conductance of the line. This assumption is relaxed in our last numerical example. Several additional assumptions related to evaluation of the objective function are stated in Section III-A.

III. Optimization Framework

Figure 1 shows the overall implementation of the optimization. The highest level is controlled by OptQuest, a simulation-optimization engine [25]. The engine uses meta-heuristics and mathematical optimization techniques to search for near optimal solutions. To use the engine, the analyst defines decision variables and constraints and supplies a function that implements the objective function. The objective function is treated as a black box by OptQuest.

In our case, we implement the objective function (2) as a Monte-Carlo simulation. The decision variables ($\mathbf{A}_{ij}$ and $\mathbf{B}_{ij}$) are controlled by OptQuest as it searches the decision space, and these are inputs to the simulation. In our problem, standard Monte-Carlo simulation is too slow. Thus, a rare-event splitting technique is implemented on top of the underlying blackout simulation to improve efficiency. Splitting has been applied to cascading blackouts in several papers (e.g., [26], [27]). We now describe the three components of Figure 1 in more detail, starting from the bottom up.

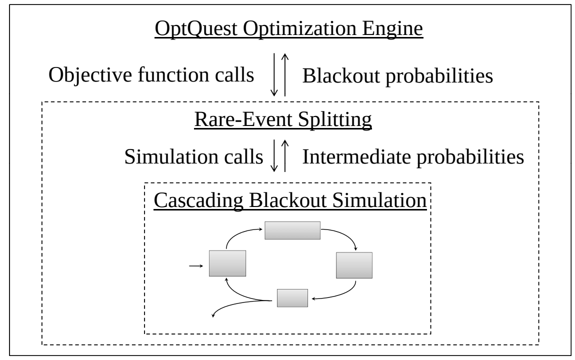


Fig. 1. Overall optimization framework.

A. Cascading Blackout Simulation

Figure 2 shows the basic logic for a stochastic blackout model [26]. The simulation begins by randomly selecting a line to trip. Then it cycles through a loop in which line failures in the previous iteration may lead to line failures in the next iteration. This cascading effect continues until there are no new line failures. The main steps of the model are:

- **Check network connectivity.** When a transmission line fails, it is possible for the network to split into separate islands. Disconnected islands are mathematically independent, so subsequent steps in the loop are repeated once for each island. This step handles the bookkeeping needed for tracking multiple islands throughout the simulation.[2]
- **Match generated power and load.** This step matches total generated power within an island with the total load within an island. If load exceeds generator capacity, then load is shed. Conversely, if generator capacity exceeds load, then generated power is adjusted to match the load.
- **Solve linearized power-flow equations.** The standard power-flow equations are non-linear [28],[29]. A common approximation (cf. [21]) is their linearization to yield the so-called DC power-flow equations:

$$p_i = \sum_{j=1}^{n} b_{ij}(\theta_i - \theta_j), \qquad i = 1, \ldots, n, \tag{7}$$

  where p_i is the net real power injected into bus i, θ_i is the phase angle of the voltage at bus i, and b_{ij} is the imaginary part of element (i, j) of the admittance matrix.
- **Check for line trips.** Each line remains in a working state as long as the power flow on the line remains below the line capacity. The line fails as soon as the load exceeds the capacity.

In this model, the capacity of line i–j is assumed to be a fixed, but unknown *random* quantity C_{ij}. The mean value of C_{ij} is the existing capacity a_{ij} plus the additional capacity $\mathbf{A}_{ij}$. In each iteration, a working line i–j fails with probability

$$\Pr\{C_{ij} \le x_{ij} | C_{ij} > x'_{ij}\} = \frac{[F_{ij}(x_{ij}) - F_{ij}(x'_{ij})]^+}{1 - F_{ij}(x'_{ij})}, \tag{8}$$

[2]The model does not consider *active* islanding in which the system deliberately separates itself into islands to contain the blackout, though such a feature could be embedded within the optimization framework.

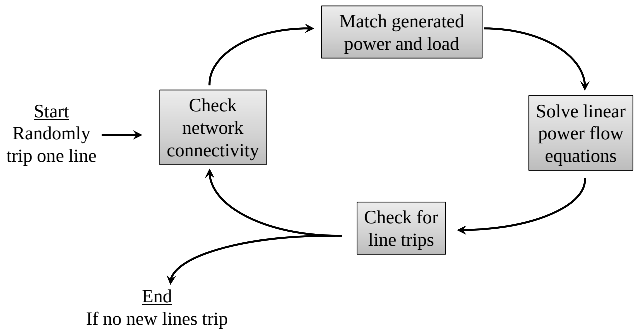


Fig. 2. Overall logic of a stochastic cascading line-failure model [26].

where F_{ij} is the CDF of C_{ij}, x_{ij} is the power flow across line i–j, x'_{ij} is the largest *previously observed* power flow across line i–j, and the function $[s]^+ \equiv \max(s, 0)$. The function $[\cdot]^+$ accounts for the possibility that the load on a line may decrease from one iteration to the next.

A stochastic model for line capacity can be motivated in several ways. First, vertical sag depends on stochastic factors such as wind and ambient temperature, as does vegetation growth (e.g., [30]). Second, uncertainty in protective devices can be a source of randomness in line capacity (e.g., [21]). Third, stochastic capacity can be a surrogate for stochastic load. Since $\Pr\{\text{capacity} \le \text{load}\} = \Pr\{\text{capacity} - \text{load} \le 0\}$, putting a distribution on capacity can be a surrogate for putting a distribution on the *difference* between capacity and load.

B. Rare-Event Splitting

To illustrate the challenge in simulating rare events, consider an event that occurs with probability 10^{-6}. The number of trials required to estimate this probability with a relative error of 10% using standard simulation is about 10^8 (relative error $\equiv$ standard deviation of estimator divided by its mean). Conducting one trial per second, it would take about 3 years to complete the experiment.

Rare-event techniques can be used to improve the efficiency. One approach is *importance sampling*. This approach has been used to evaluate cascading blackout probabilities in [21], [22], [30]. Here, we use an alternate technique called *splitting* (cf. [31], [32], [33]). The basic idea of splitting is to interrupt the simulation whenever it gets "near" the rare event of interest. From the interruption point, the simulation is split into multiple independent replications. In this way, a higher fraction of simulation time is allocated to sample paths that are more likely to reach the rare event.

To illustrate, suppose that a large-scale blackout is defined as an event where the shed load is L or more. We can break this problem into two easier problems as follows:

(a) Estimate the probability that the shed load reaches $L/2$,
(b) Estimate the probability that the shed load reaches L *given* that the shed load has reached $L/2$.

By interrupting the simulation when the shed load reaches $L/2$, it is possible to control the relative allocation of sample paths devoted to estimating the probabilities in (a) and (b). Roughly speaking, the benefit is achieved because we can spend equal amounts of time in both parts. In contrast, standard simulation spends virtually all of its time simulating (a), so the probability estimate for (b) is not very good.

More generally, we can define a larger number of intermediate levels and simulate the probability of reaching each level in succession starting from the previous level. The final estimate is the product of the intermediate probabilities. For further details of the splitting methodology and its application to cascading blackouts, see [20], [26].

C. OptQuest

OptQuest uses meta-heuristics and mathematical optimization to search the solution space and select decision variables ($\mathbf{A}_{ij}$ and $\mathbf{B}_{ij}$) that are passed to the underlying simulation. OptQuest's main optimization engine is based on the scatter search methodology coupled with tabu search strategies (e.g., [34], [35], [36], [37]). The optimization technologies within OptQuest include both a catalog of search procedures and a separate catalog of solution generation methods. The system also employs additional technologies to complement the default search mechanisms, including design of experiments, cross entropy and linear and mixed integer programming. OptQuest is built under the assumption that a computationally expensive black box is used to evaluate the objective function. So, prediction models are used to estimate the value of the objective function prior to calling the objective-function evaluator and to assist in establishing search directions.

Because the problems treated by OptQuest are generally too complex to have closed form mathematical representations, there is no way to define a problem relaxation that will enable a determination of whether the solutions returned by OptQuest are optimal. (No classical mathematical approach, short of total enumeration of a potentially astronomical number of alternatives, can provide such a guarantee of optimality in these cases either.) In special instances, as where the problem reduces to a linear or mixed integer programming problem, then optimality is determined in the same way as by current state-of-the-art methods based on classical algorithms.

IV. Numerical Results

This section discusses numerical experiments applied to several network topologies – a grid network, a tree-like network, and a small real network. The grid and tree networks are idealized topologies. Nevertheless, they are similar to real networks in the sense that each bus is incident to, on average, about three lines, which is approximately the average for large networks in practice [24].[3] Similar idealized networks structures were also considered in [38]. The fact that multiple distinct network topologies have similar incidence properties indicates that the degree of a bus is only one of many properties that impacts the behavior of a network.

[3]For example, the IEEE 118-bus network has 179 lines with an average incidence of $2 \times 179/118 \approx 3.0$. In the grid network (Figure 3), the average incidence is 3.2. In the tree network (Figure 6), each node is incident to exactly three lines.

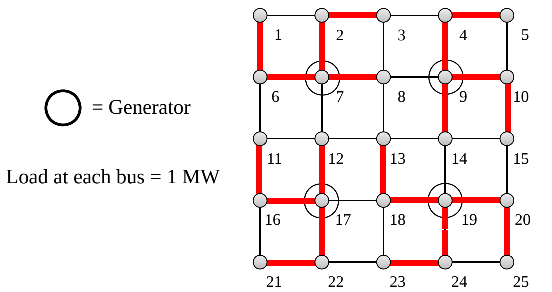


Fig. 3. Grid network; thick lines show a sample set of minimal trees.

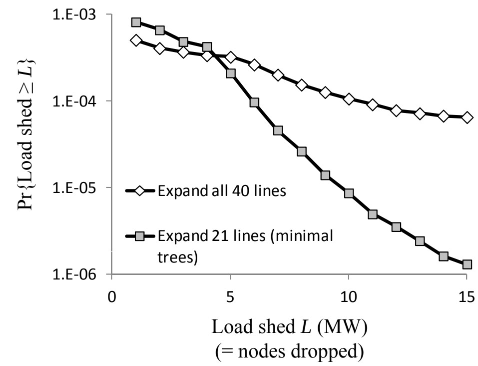


Fig. 4. Comparison of capacity expansion solutions.

A. Grid Network

Figure 3 shows a 5×5 grid network. The network contains 25 buses, 40 transmission lines, and 4 generators (the thick lines will be explained in a moment). Each bus has a load of 1 MW. The capacity of each line follows a normal distribution with mean μ and standard deviation $\sigma = .5$ MW.

The line-capacity parameter μ is set according to the following "stochastic $N-1$" rule: Loop through all possible single line failures; for each single line failure, determine the resulting power flows on the network; set μ equal to twice the largest power flow observed on *any* line following *any* single line failure. This means that, following any initial line failure, all lines will be operating at or below 50% of their *expected* capacities. Now, this $N-1$ rule does not mean that it is *impossible* for a second line to fail. Rather, for a particular line, the *actual* ratio of load to capacity is stochastic. Thus, for each run of the simulation, there are typically a few lines where, by chance, the actual loadings are higher, say at 90% utilization or more. After the initial trip, if the observed power flow on any line exceeds its *actual* capacity, then the line fails, and the blackout continues.

In subsequent numerical examples, we consider more congested networks where μ is set at a 75% utilization threshold rather than a 50% utilization threshold. More generally, it is possible to apply the methodology to arbitrary network loadings – for example, networks in which most lines are operating well below their expected capacity, but where a few lines are operating near their expected capacity. Such networks can be evaluated by changing the bus loads p_i or line capacities a_{ij} in the baseline network. (Our numerical tests do not include such networks, so the overall conclusions of the paper may or may not extend to such cases.)

In using OptQuest, the analyst can propose solutions to the optimization engine. These solutions are incorporated into the set of candidate solutions used by the search heuristics. We provide the following trial solutions to the OptQuest engine:

1) Uniform expansion of existing lines. The budget is divided evenly to increase the capacity of existing lines. Each existing line's capacity is increased by an amount such that the associated cost is T/n.
2) Expansion of minimal trees. The budget is divided evenly among a subset of existing lines that form trees from the generators to the loads. The basic idea is to focus resources on strengthening exactly *one* existing path to each node. A sample set of minimal trees is shown by the thick lines in Figure 3. These lines connect each node to exactly one generator using a minimal number of hops. Because of the symmetry of the problem, there are multiple ways to define such trees.

Figure 4 shows a sample comparison of the two suggested capacity-expansion solutions. The blackout-size distributions are simulated using splitting without any optimization. The figure shows that the uniform-expansion solution (40-line expansion) is slightly better than the minimal-tree solution (21-line expansion) at preventing *smaller* blackouts. But, the minimal-tree solution is much better at preventing larger blackouts.

This illustrates that different expansion strategies lead to different shapes of the *tails* of the blackout distributions. There is a potential trade-off between reducing the frequency of small-scale blackouts versus reducing the frequency of large-scale blackouts. This has also been observed in [18]. A reduction in small-scale blackouts has a more immediate benefit, while a reduction in large-scale blackouts requires a potentially longer time horizon to observe the benefit.

Now we discuss results obtained from solving the optimization problem, (2)-(6). Table I shows sample parameters used in the numerical experiments. The first set of parameters are parameters in the optimization problem itself. The second set of parameters are used by the OptQuest control engine (Figure 1). The total computing budget is the total time allotted to search for an optimal solution. The computing budget per replication is the time allotted for each simulation call by OptQuest – that is, a call to the splitting-allocation engine in Figure 1. Each replication returns one estimate for the probability of a blackout. Since these estimates are stochastic, OptQuest may use multiple replications in order to determine whether or not a candidate solution is better or worse than the best current solution. Once this determination can be made with the specified confidence level, OptQuest moves to a new candidate solution. Otherwise, there is a minimum and maximum number of replications that OptQuest tries per candidate solution.

Figure 5 shows sample best solutions obtained from OptQuest. Note that the solutions obtained by OptQuest are not necessarily the true optimal solutions; they are simply the best solutions found within the allotted time budget. The x-axis shows the existing lines in the grid network, identified by

TABLE I
SAMPLE OPTIMIZATION PARAMETERS

Parameter	Value
Expansion budget T	40
Fixed line cost c_f	0 / distance
Variable expansion cost c_v	1 / MW / distance
Minimum line capacity r	1 MW
Total computing budget	10 hours
Computing budget per replication	30 sec.
Minimum replications per solution	3
Maximum replications per solution	20
Confidence level	95%

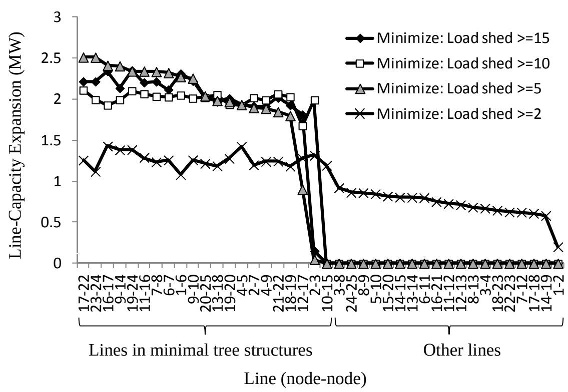


Fig. 5. Best expansion solutions found for grid network.

node indices (see Figure 3). The y-axis shows the additional capacity allocated to each line in the best solution. Four minimization problems are considered with different thresholds for the blackout size: 2, 5, 10, and 15 MW.

There are several observations from this figure. First, the best solutions consist entirely of expansions to *existing* lines – no additional lines are added, even though the fixed cost of adding a new line is zero. Second, the best solutions for minimizing the probabilities of 5-, 10-, and 15-MW blackouts are close to the minimal-tree solution – namely, the solution that adds the same fixed capacity to each line in the minimal trees (Figure 3) while adding no capacity to the other lines. In contrast, the best solution for minimizing the probability of a 2-MW blackout is closer to the solution that adds the same fixed capacity to *all* lines in the network, though it tends to add slightly more capacity to lines in the minimal trees. This is consistent with Figure 4, in which the minimal-tree solution dominates the uniform-expansion solution for all but the smallest blackouts. In summary, even though the minimal-tree solution may not be exactly optimal in all cases, it appears to perform well over a range of different blackout thresholds, and the best solutions found by OptQuest do not deviate too much from this solution. Third, some of the solutions omit the expansion of one or two lines in the minimal tree solution (e.g., line 10-15 or line 2-3). These solutions may be sacrificing the protection of one or two nodes in the network (allowing a higher probability for a small blackout) in order to achieve a lower probability of a larger blackout.

Finally, we note a potential limitation with the structure of the optimization problem, (3)-(4). The optimization attempts to minimize the probability of a blackout *at a specific threshold*.

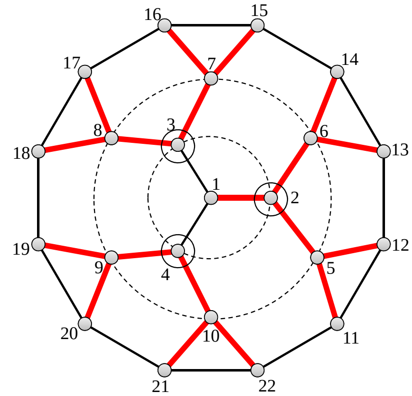


Fig. 6. Tree network; thick lines show a sample set of minimal trees.

This has the potential to favor solutions that "allow" some nodes to fail while still minimizing the probability of a blackout at the threshold. For example, suppose that we wish to minimize $\Pr\{\text{shed load} \geq 10\}$. A potential strategy is to devote all resources to protecting 16 nodes in the network, while "ignoring" the remaining 9 nodes. In other words, we can tolerate a 9-MW blackout, but we do not want a 10th node to fail. It turns out that such solutions tend to be much worse than the best solutions found by OptQuest, sometimes by several orders of magnitude. Thus, while the structure of the optimization problem might favor anomalous behavior, we have not generally found it to do so in the experiments considered here.

B. Tree Network

Now we consider the expansion problem applied to a tree network (Figure 6). Similar tree networks were considered in [24]. Each branch of the tree splits into two sub-branches, except for the root node which splits into three. The outer edge of the network is connected by a ring. Each node of the network is incident to three lines. Three generators are placed at nodes 2, 3, and 4. An example minimal-tree structure is shown by the thick lines. Cartesian coordinates of the nodes are specified as follows: The nth layer of the tree ($n = 0, 1, 2, 3$) is located on a circle whose radius is proportional to n; within a layer, the nodes are evenly spaced around the circle. The network has 22 buses (nodes) and 33 lines. Each node has a constant load of 1 MW. For each line, the mean capacity of the line is set at $4/3$ times the largest flow observed on the line considering all possible single-line failures. In other words, following any single-line failure, no line carries more than 75% of its mean capacity.

For this numerical experiment, the optimization parameters (Table I) are set as follows: The expansion budget T is set at 30 units. The total computing budget is 22 hours (for each of the four minimization problems). The computing budget per replication is 45 seconds. The maximum number of replications per solution is 15. Other parameters are the same as in Table I. The expansion budget $T = 30$ corresponds roughly to a 24% increase in network capacity, if used to increase the capacity of each line by a fixed proportion.

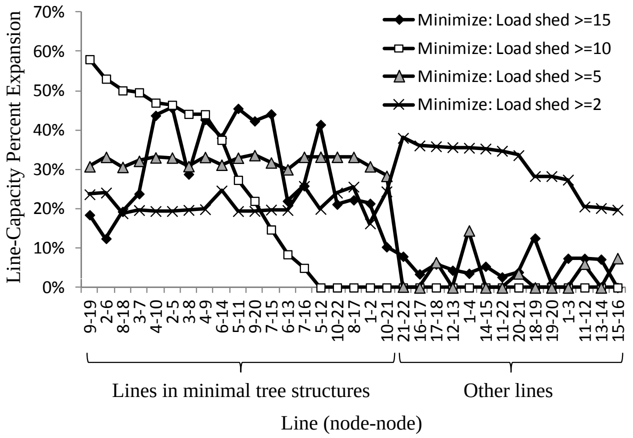


Fig. 7. Best expansion solutions found for tree network.

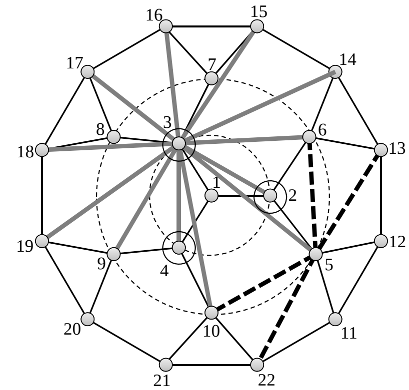


Fig. 8. Partial representation of best solution found, low utilization scenario.

Figure 7 shows sample best solutions obtained from OptQuest for the tree network. Four different minimization problems are considered, each with a different blackout threshold: 2, 5, 10, and 15 MW. The four solutions yield respective blackout probabilities of approximately 1×10^{-2}, 5×10^{-3}, 3×10^{-5}, and 6×10^{-7}. As before, the best solutions consist entirely of expansions to existing lines.

The y-axis in the figure shows the percent increase of capacity for each line in the best solution. Observations from this figure are similar to those from Figure 5. The best solution for minimizing the probability of a *small* (2-MW) blackout is close to the solution that increases *every* line by a fixed proportion – around 20-30% in this case. The best solutions for *larger* (5-, 10-, and 15-MW) blackouts tend to allocate resources to lines within a minimal-tree subset, ignoring lines outside of this subset. This is most clearly seen in the 5-MW blackout case. The best solution proportionally increases each line in the minimal-tree subset by about 30%, leaving the other lines (mostly) unchanged. Another interesting result is the 10-MW case. The solution allocates the budget to a strict subset of lines within the minimal-tree subset – roughly speaking, lines that are directly downstream from the generators (lines 2-5, 2-6, 3-7, 3-8, 4-9, and 4-10) have the highest proportional increase in capacity, while lines that are further downstream have less increase in capacity.

Now we consider a numerical experiment with a much higher expansion budget: $T = 150$ units. Such a budget corresponds to a 121% increase in network capacity (i.e., the budget can be used to more than double the capacity of each line in the network). This experiment corresponds to an over-expanded, low-utilization scenario. The solution was optimized at a blackout threshold of 3 MW (loss of three nodes in the network).

Figure 8 gives a pictorial representation of the best solution found, with an associated blackout probability of about 10^{-7}. The solution consists entirely of *new* lines, shown in the figure by thick gray lines and dashed black lines. In contrast, the previous experiments yielded solutions consisting entirely of expansions to *existing* lines. The figure shows only a partial set of lines in the solution to avoid figure clutter. The lines that are not shown are symmetric permutations of those shown. The solution can be described as follows: Each generator (e.g., node 3) is directly connected to every other generator and to every node on the middle ring and to the six nearest nodes on the outer ring (thick gray lines); each node on the middle ring (e.g., node 5) is directly connected to its two nearest neighbors on the middle ring and its two nearest non-connected neighbors on the outer ring (dashed black lines); there are a few exceptions – some "miscellaneous" lines are added on the edge (lines 12-14, 13-15, 16-18) and line 2-4 is missing. Qualitatively, the solution tries to directly connect each node to a generator (that is, to minimize the number of hops) provided the distance is not prohibitive. The solution passes a simple sanity test in the sense that the best solution is nearly symmetric with respect to the symmetry of the underlying network. The exceptions can be explained by noise in a stochastic rare-event environment.

Figure 9 shows the *entire* blackout distribution for the best solution found, and for the two heuristic solutions used as initial suggestions. The expansion problem was optimized at the 3-MW blackout threshold (3 nodes dropped). Thus the best-found solution provides the lowest probability at this threshold. But the three solutions vary dramatically over the full distribution. The uniform solution (upgrading each line uniformly) is nearly as good as the best-found solution for small blackouts, but it results in much lower probabilities for larger blackouts. The minimal-tree solution results in the worst probabilities (among the three solutions) for small blackouts, but results in, by far, the lowest probabilities for large blackouts. In fact, *zero* blackouts above 8 MW were observed in the one hour of simulation time used to generate each distribution in the figure.

C. IEEE 14-Bus Network

This section considers an IEEE 14-bus test network. The network topology, bus loads, generator locations, and line parameters (conductance / susceptance) are obtained from http://www.ee.washington.edu/research/pstca/. Latitude and longitude of the buses are not given, so line distances are ignored (we assume $d_{ij} \equiv 1$). As in the previous example, we set the mean capacity of each line to $4/3$ times the largest flow observed on the line considering all possible single-line

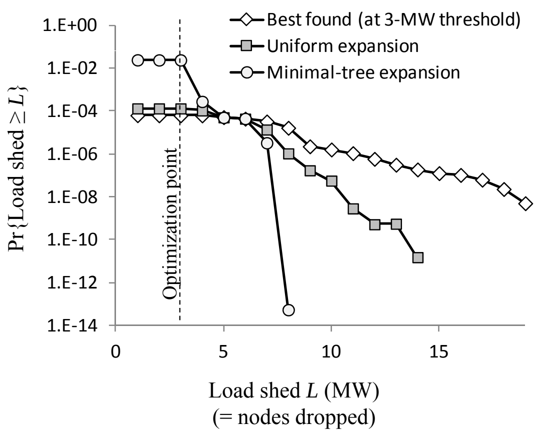


Fig. 9. Comparison of capacity expansion solutions.

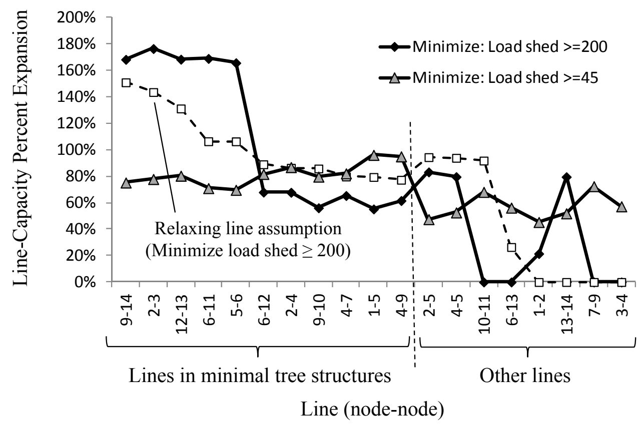


Fig. 10. Best expansion solutions found for IEEE-14 network.

failures.

The optimization parameters are set as follows: The expansion budget T is set to $1,000$ units, corresponding to roughly a 70% increase in network capacity. The computing budget is 12 hours for each minimization problem. The computing budget per replication is 60 seconds. Other parameters are the same as in Table I. Because the network contains two different voltage zones, we add constraints preventing the addition of new lines between the two zones, though lines can be added within a zone.

Figure 10 shows results for two optimization problems – one that minimizes the probability of a 200 MW blackout and one that minimizes the probability of a 45 MW blackout (the total load in the network is 259 MW). These results are consistent with results for the grid and tree networks. For large blackouts, the best solution tends to focus resources on the minimal tree structure, though some lines outside of the minimal tree are also expanded in this case. For small blackouts, the best solution tends to proportionally increase the capacity of every line in the network by a similar amount (roughly $70\% \pm 25\%$ here). Here, the qualitative differences between these two solution types are not as distinct as they are for the two idealized networks (comparing Figures 5, 7, and 10). The solutions from the IEEE 14-bus network seem to closer to those of the tree network than those of the grid network.

The figure also includes an example in which line characteristics change with additional capacity (relaxing one of the assumptions at the end of Section II). Using transmission-line data from [28], we create a curve fit to establish an approximate relationship between line susceptance and surge impedance loading (SIL). Using SIL as a surrogate for capacity then gives an assumed relationship between capacity and line susceptance. The solution using the relaxed assumption is similar to the original in the sense that it focuses resources on the minimal tree structure plus a few other lines.

V. Analytical Model

This section gives an analytical model that explains some of the numerical results from the previous section. Many simplifying assumptions are made, so we do not claim that results are universally true. Rather, the analytical model, as a simplified version of the more complex model, can be useful in understanding the behavior of more complex models. The objective of the analytical model is to address the following high-level question: Given a fixed set of resources, is it better to create one strong path to a node (like a tree network) or to divide resources among multiple redundant paths (like a mesh network)? The analytical model shows that a single strong path is better when the line utilization is high, whereas multiple paths are better when the line utilization is low.

Consider a very simple network consisting of two buses. We wish to determine which is more reliable: (a) One transmission line connecting the two buses, or (b) two redundant lines, but with half the capacity each? We analyze this question with respect to the model in Section III-A.[4] Let $F(x)$ denote the CDF of a non-negative random variable with a mean of 1. For the one-line scenario, assume that the single-line capacity CDF is $F_A(x) \equiv F(x/A)$, which has expected value A. For the two-line scenario, assume that the capacity CDF of each line is a linear scaling of the previous CDF: $F_{A/2}(x) \equiv F(x/(A/2)) = F(2x/A)$, which has expected value $A/2$. Let L be the load carried between the two buses. In the one-line scenario, the load on the line is L; in the two-line scenario, the load on each line is $L/2$. Let $\rho \equiv L/A$ represent the load-to-capacity ratio.

In the one-line scenario, the system fails if the capacity of the single line is less than the load L. This occurs with probability $F_A(L) = F(L/A) = F(\rho)$. In the two-line scenario, the system fails if either of the following two mutually exclusive events occurs:

1) Both lines fail in the initial iteration of the loop – that is, the capacities of both lines are less than $L/2$; this occurs with probability $[F_{A/2}(L/2)]^2 = F^2(L/A) = F^2(\rho)$,
2) One line fails in the initial iteration of the loop and the second line fails in the second iteration of the loop – that is, the capacity of the first line is less than $L/2$ (causing it to fail, resulting in its load being transferred to the other line), and the capacity of the second line is

[4] However, we do not start the model here by deliberately tripping one random line in the network.

between $L/2$ and L; this occurs with probability

$$\begin{aligned} 2F_{A/2}(L/2)&[F_{A/2}(L) - F_{A/2}(L/2)] \\ &= 2F(L/A)[F(2L/A) - F(L/A)] \\ &= 2F(\rho)[F(2\rho) - F(\rho)]. \end{aligned}$$

Thus, the two-line scenario is more reliable than the one-line scenario if and only if:

$$F^2(\rho) + 2F(\rho)[F(2\rho) - F(\rho)] < F(\rho),$$

which can be rearranged to yield

$$F^c(\rho) < 2F^c(2\rho). \tag{9}$$

For example, if the capacity distributions are exponential, $F^c(x) = e^{-x}$, then the two-line option is more reliable than the one-line option if and only if:

$$\begin{aligned} e^{-\rho} &< 2e^{-2\rho} \\ \rho &< \ln(2) \approx .69. \end{aligned}$$

Thus, when the utilization is high (ρ is close to 1), the one-line option is more reliable; when the utilization is low, the two-line option is more reliable.

As a second example, suppose the capacity distributions are normally distributed. Then $F(x) = \Phi((x-1)/\sigma)$, a normal distribution with mean 1 and standard deviation σ, where $\Phi(\cdot)$ is the CDF of a standard normal random variable. By (9), the two-line option is more reliable than the one-line option if and only if:

$$1 - \Phi\left(\frac{\rho - 1}{\sigma}\right) < 2\left(1 - \Phi\left(\frac{2(\rho-1)}{\sigma}\right)\right),$$

A numerical evaluation of this equation with $\sigma = 0.2$ (meaning the standard deviation of line capacity is 20% of its mean) shows that the one-line option is more reliable when ρ is less than about .5 and the two-line option is more reliable otherwise.

This helps to explain why the optimization results tended to favor expanding exactly one path from each node to a nearby generator. As a blackout progresses, the line utilizations tend to increase. To prevent the blackout from progressing further, it becomes more favorable to have a smaller set of stronger paths rather than a larger set of relatively weaker paths. For low-utilization, small-blackout scenarios, multiple redundant paths are favored.

VI. Conclusions

We have introduced a simulation-optimization model to formulate the problem of expanding a power grid in order to minimize the probability of a cascading blackout. The decision space included the addition of new transmission lines and upgrades to existing lines (expansion of generation capacity was not considered). The optimization problem was tackled using the commercial solver OptQuest in combination with a splitting technique to estimate the required rare-event probabilities in a reasonable time. This is a distinguishing aspect of this work compared with the existing literature on capacity expansion.

We conducted a variety of numerical experiments on specific network topologies including a grid network, a tree-based network, and an IEEE test network. The optimization was seeded with two heuristic solutions – a "uniform" solution that proportionally expanded each line in the network, and a "minimal-tree" solution that proportionally expanded a subset of lines forming trees originating at each generator. We evaluated each optimization problem with a computing budget of 1-3 days, and reported the best solutions found.

A key observation is that different expansion strategies lead to different shapes in the tails of the blackout distributions. Our findings show that the strategies which succeed in reducing small-scale blackouts are not necessarily optimal at reducing large-scale blackouts, and vice-versa. In particular, small blackouts are typically minimized via solutions that are qualitatively similar to the heuristic uniform solution, whereas large blackouts are typically minimized via solutions that are similar to the minimal-tree solution. In other words, large-scale blackouts are avoided by focusing all resources on strengthening *one* path to each node in the network, rather than by dividing up resources among multiple paths to each node. A second observation is that the best solutions typically do not add new lines to the network. New lines are only added in a low-utilization, over-expanded scenario. We confirmed this behavior through a theoretic analysis of a 2-node model that could be solved analytically.

Future work may involve increasing the complexity of the underlying blackout model to address issues such as generator failures, transient effects, and human error. In addition, we propose to adapt the optimization framework to incorporate decisions associated with implementation of a smart grid, such as deployment of sensors and self-healing controls.

Acknowledgment

This material is based in part upon work supported by the Department of Energy under Award Number DE-SC0002223. The authors thank the anonymous reviewers for their helpful comments to improve the paper.

References

[1] S. Lee, "For the good of the whole," *IEEE Power & Energy*, vol. 5, no. 5, pp. 24–35, 2007.

[2] D. Morrow and R. Brown, "Future vision: The challenge of effective transmission planning," *IEEE Power & Energy*, vol. 5, no. 5, pp. 36–45, 2007.

[3] W. Li and P. Choudhury, "Probabilistic transmission planning," *IEEE Power & Energy*, vol. 5, no. 5, pp. 46–53, 2007.

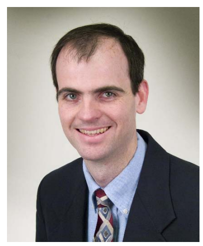
[4] R. Romero, R. Gallego, and A. Monticelli, "Transmission system expansion planning by simulated annealing," *IEEE Transactions on Power Systems*, vol. 11, no. 1, pp. 364–369, 1996.

[5] S. Haffner, A. Monticelli, A. Garcia, J. Mantovani, and R. Romero, "Branch and bound algorithm for transmission system expansion planning using a transportation model," *IEE Proceedings of Generation, Transmission, and Distribution*, vol. 147, no. 3, pp. 149–156, 2000.

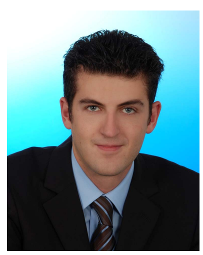
[6] P. Maghouli, S. Hosseini, M. Buygi, and M. Shahidehpour, "A scenario-based multi-objective model for multi-stage transmission expansion planning," *IEEE Transactions on Power Systems*, vol. 26, no. 1, pp. 470–478, 2011.

[7] H. Yu, C. Chung, and K. Wong, "Robust transmission network expansion planning with Taguchi's orthogonal array testing," *IEEE Transactions on Power Systems*, vol. 26, no. 3, pp. 1573–1580, 2011.

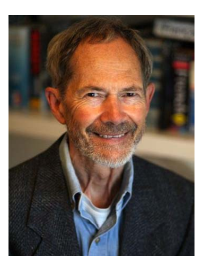
[8] J. Zhao, J. Foster, Z. Dong, and K. Wong, “Flexible transmission network planning considering distributed generation impacts,” *IEEE Transactions on Power Systems*, vol. 26, no. 3, pp. 1434–1443, 2011.
[9] E. da Silva, H. Gil, and J. Areiza, “Transmission network expansion planning under an improved genetic algorithm,” *IEEE Transactions on Power Systems*, vol. 15, no. 3, pp. 1168–1175, 2000.
[10] D. Kirschen and G. Strbac, “Why investments do not prevent blackouts,” *The Electricity Journal*, vol. 17, no. 2, pp. 29–36, 2004.
[11] H. Youssef and R. Hackam, “New transmission planning model,” *IEEE Transactions on Power Systems*, vol. 4, no. 1, pp. 9–18, 1989.
[12] M. Xie, J. Zhong, and F. Wu, “Multiyear transmission expansion planning using ordinal optimization,” *IEEE Transactions on Power Systems*, vol. 22, no. 4, pp. 1420–1428, 2007.
[13] D. Camac, R. Bastidas, R. Nadira, C. Dortolina, and H. Merrill, “Transmission planning criteria and their application under uncertainty,” *IEEE Transactions on Power Systems*, vol. 25, no. 4, pp. 1996–2003, 2010.
[14] M. Carrión, J. Arroyo, and N. Alguacil, “Vulnerability-constrained transmission expansion planning: A stochastic programming approach,” *IEEE Transactions on Power Systems*, vol. 22, no. 4, pp. 1436–1445, 2007.
[15] H. Ren, I. Dobson, and B. Carreras, “Long-term effect of the n-1 criterion on cascading line outages in an evolving power transmission grid,” *IEEE Transactions on Power Systems*, vol. 23, no. 3, pp. 1217–1225, 2008.
[16] H. Ge and S. Asgarpoor, “Reliability evaluationo of equipment and substations with fuzzy Markov processes,” *IEEE Transactions on Power Systems*, vol. 25, no. 3, pp. 1319–1328, 2010.
[17] G. K. Chan and S. Asgarpoor, “Optimum maintenance policy with Markov processes,” *Electric Power Systems Research*, vol. 76, pp. 452–456, 2006.
[18] D. Newman, B. Carreras, V. Lynch, and I. Dobson, “Exploring complex systems aspects of blackout risk and mitigation,” *IEEE Transactions on Reliability*, vol. 60, no. 1, pp. 134–143, 2011.
[19] “Final report on the August 14, 2003 blackout in the United States and Canada: Causes and recommendations,” Tech. Rep., 2004, available: http://www.nerc.com.
[20] J. Shortle, C. H. Chen, B. Crain, A. Brodsky, and D. Brod, “Optimal splitting for rare-event simulation,” *IIE Transactions*, vol. 44, no. 5, pp. 352–367, 2012.
[21] J. Chen, J. Thorp, and I. Dobson, “Cascading dynamics and mitigation assessment in power system disturbances via a hidden failure model,” *Electrical Power and Energy Systems*, vol. 27, no. 4, pp. 318–326, 2005.
[22] K. Bae and J. Thorp, “A stochastic study of hidden failures in power system protection,” *Decision Support Systems*, vol. 24, no. 3-4, pp. 259–268, 1999.
[23] M. Anghel, K. Werley, and A. Motter, “Stochastic model for power grid dynamics,” in *Fortieth Hawaii International Conference on System Sciences*, 2007.
[24] B. Carreras, V. Lynch, I. Dobson, and D. Newman, “Critical points and transitions in an electric power transmission model for cascading failure blackouts,” *Chaos*, vol. 12, no. 4, pp. 985–994, 2002.
[25] M. Better, F. Glover, and M. Laguna, “Advances in analytics: Integrating dynamic data mining with simulation optimization,” *IBM Journal of Research & Development*, vol. 51, no. 3/4, pp. 477–487, 2007.
[26] J. Shortle, “Efficient simulation of blackout probabilities using splitting,” *Electrical Power and Energy Systems*, vol. 44, no. 1, pp. 743–751, 2013.
[27] J. Kim, J. Bucklew, and I. Dobson, “Splitting method for speedy simulation of cascading blackouts,” *IEEE Transactions on Power Systems*, vol. 28, no. 3, pp. 3010–3017, 2013.
[28] A. V. Meier, *Electric Power Systems: A Conceptual Introduction*. Hoboken, NJ: Wiley, 2006.
[29] S. Frank, I. Steponavice, and S. Rebennack, “Optimal Power Flow: A Bibliographic Survey I, Formulations and Deterministic Methods,” *Energy Systems*, vol. tba, p. tba, 2012.
[30] Q. Chen and L. Mili, “Risk-based composite power system vulnerability evaluation to cascading failures using importance sampling,” in *IEEE Power Engineering Society General Meeting*, 2011.
[31] P. L’Ecuyer, V. Demers, and B. Tuffin, “Splitting for rare-event simulation,” in *Proceedings of 2006 Winter Simulation Conference*, L. F. Perrone, F. P. Wieland, J. Liu, B. G. Lawson, D. M. Nicol, and R. M. Fujimoto, Eds. Piscataway, NJ: IEEE, 2006, pp. 137–148.
[32] P. L’Ecuyer, F. L. Gland, P. Lezaud, and B. Tuffin, “Splitting techniques,” in *Rare Event Simulation using Monte Carlo Methods*, G. Rubino and B. Tuffin, Eds. Chichester, U.K.: Wiley, 2009, pp. 39–62.
[33] J. Shortle and P. L’Ecuyer, “Introduction to rare-event simulation,” in *Wiley Encyclopedia of Operations Research and Management Science*, J. Cochran, Ed. Wiley, 2010.
[34] F. Glover, J. P. Kelly, and M. Laguna, “New advances and applications of combining simulation and optimization,” in *Proceedings of the 1996 Winter Simulation Conference*, J. M. Charnes, D. J. Morrice, D. T. Brunner, and J. J. Swain, Eds., 1996, pp. 144–152.
[35] F. Glover and M. Laguna, *Tabu Search*. Boston, MA: Kluwer Academic Publishers, 1997.
[36] F. Glover, M. Laguna, and R. Mart, “Scatter search,” in *Advances in Evolutionary Computation: Theory and Applications*, A. Ghosh and S. Tsutsui, Eds. New York, NY: Springer-Verlag, 2003, pp. 519–537.
[37] ——, “Scatter search and path relinking: Advances and applications,” in *Handbook of Metaheuristics*, F. Glover and G. Kochenberger, Eds. Boston, MA: Kluwer Academic Publishers, 2003, pp. 1–35.
[38] B. Carreras, V. Lynch, M. Sachtjen, I. Dobson, and D. Newman, “Modeling blackout dynamics in power transmission networks with simple structure,” in *34th Hawaii International Conference on System Sciences*, 2001.



John Shortle is an Associate Professor of Systems Engineering and Operations Research at George Mason University. His research interests include simulation, queueing, and analysis of rare events with applications in air transportation, energy, and telecommunications. He received the Daniel H. Wagner Award for Excellence in Operations Research Practice. John received a B.S. in Mathematics from Harvey Mudd College and a Ph.D. and M.S. in Industrial Engineering and Operations Research from UC Berkeley.



Steffen Rebennack is an Assistant Professor in the Division of Economics and Business at the Colorado School of Mines, USA. He received his PhD degree in 2010 from the University of Florida, USA. His research interests are in dimension-reduction techniques for large-scale optimization problems, particularly in stochastic optimization and global optimization. He focuses on power systems applications.



Fred Glover is Chief Technology Officer for OptTek Systems, Inc., and Distinguished Professor affiliated with the School of Engineering and the Leeds School of Business at the University of Colorado, Boulder. He has authored or co-authored more than 400 published articles and eight books in the fields of mathematical optimization, computer science and artificial intelligence. He is the recipient of the von Neumann Theory Prize, and is an elected member of the National Academy of Engineering.