IET Generation, Transmission & Distribution

April 2019

Copyright 2019 Institution of Electrical Engineers All Rights Reserved

Section: Pg. 1005 - 1013; Vol. 13; No. 7; 1751-8687

Length: 8068 words

Byline: SangWooPark, spark111@berkeley.edu

QingyuXu

Benjamin F.Hobbs

Body

1

Introduction

The need to transform the electric power sector into a sustainable system that incorporates renewable sources – such as wind, solar, and geothermal – is now a high priority around the globe. Achievement of this goal will, in large part, depend on wisely planned long-distance transmission networks that deliver electricity from distant renewable sources.

Traditionally, transmission has been driven by the need to interconnect planned generation or to facilitate power trade among regions. Lately, however, researchers and some policy makers have argued that transmission planning should shift from a 'reactive' process to an 'anticipative' planning process for two reasons [1, 2]. First, since transmission projects require more time to complete, commitments for transmission investments must be made before generation investments. Second, transmission investment directly influences the siting and type of generation investments by affecting their profitability and environmental impacts. Therefore, a strategically planned network can help steer generation investment towards potentially better economic and environmental outcomes.

However, such an anticipative planning process faces serious challenges. First, the optimisation problem becomes larger because it must consider both generation and transmission investments [3, 4]. To address this, researchers have implemented large-scale single-level co-optimisation models [5, 6] and bi- and tri-level successive transmission-planning models [7]. Second, planning for an infrastructure that is large in scale, costly (with individual lines costing hundreds of million dollars or more), and long-lasting requires careful consideration of the various uncertain conditions that may arise in the future. For transmission planning, this includes multi-decadal uncertainties in, e.g. load growth, technological changes, fuel costs, environmental rules, and renewable mandates. Ignoring these uncertainties can result in stranded investments or missed opportunities [1, 8, 9]. A further complication is the need to model renewable variability and the complicated nature of the transmission network. This has led to proposals for use of large-scale stochastic optimisation to inform transmission planning processes.

A variety of methods have been suggested for planning under long-run uncertainty. These include chanceconstrained programming [10–13], and various scenario planning methods (e.g. least-regret planning [14] used by CAISO [15, 16], and MISO's multi-value planning [17]). A popular approach is multistage stochastic programming,

in which initial decisions must weigh the likelihood of different long-run outcomes, as well as how later recourse decisions can adapt the system as uncertainties unfold [18]. Several recent papers additionally assess the value of flexible network and non-network technologies such as phase-shifting transformers, energy storage, and demand-side management in long-term transmission planning [19, 20].

Here, we use a two-stage stochastic transmission planning model JHSMINE (Johns Hopkins stochastic multi-stage integrated network expansion), which is a mixed-integer linear programming implementation that incorporates cooptimisation (of generation and transmission investments) and long-term uncertainties represented by multiple scenarios [21]. When solved, this model provides an optimal (expected cost-minimising) solution that includes transmission lines to be built in each investment decision stage, the type and capacity of generators that should be built for each location and stage, and how much electricity each power plant should generate for each hour. In the second stage, a separate set of decision variables for each scenario represents 'recourse' or adaptations of the plan.

However, considering multiple scenarios vastly increases the number of variables and parameters considered, making stochastic programs hard to solve. As a result, many simplifications need to be made concerning network flows or operations. Thus, future models that aim to model more realistic conditions (e.g. AC power flow, unit-commitment) will be even more computationally intensive. So there is a need to consider ways to reduce the number of scenarios to accelerate solution times but allow the model to still consider a full range of long-run uncertainties and short-run operating conditions.

In this paper, we analyse the effectiveness of several scenario reduction methods in the context of two-stage stochastic optimisation of multidecadal regional transmission plans. The methods include three existing methods, tailored specifically for our model, and an additional simple but effective heuristic method that we developed. The goal of this work is not to give theoretical guarantees but to use concrete examples to provide useful insight into scenario reduction methods; deriving theory-based general results is left for future work. The performance of these methods is compared by testing them on the Western Electricity Coordinating Council (WECC) 300 bus planning model [21]. The results suggest that two of the methods perform significantly better than the others; furthermore, the results indicate that it is unnecessary in this case to have a large number of scenarios to obtain most of the benefits of stochastic transmission planning, but is still important to carefully choose scenarios and their probabilities.

The rest of the paper is as follows. Section 2 reviews research on scenario reduction. Section 3 summarises the model framework that we use for this research, and then describes the scenario reduction methods compared. Section 4 outlines the experimental design of the WECC case study and compares the results based on various metrics. Section 5 discusses the results, followed by conclusions in Section 6.

2

Review of literature on scenario reduction

Scenario reduction methods have been widely studied in the field of optimisation. Dupačová *et al.* [22] originally laid out the basis for a scenario reduction method based on probabilistic distances between scenarios. The authors of [23] improved that work by proposing more effective forward-selection and backward-reduction algorithms. Hoyland and Wallace [24] suggested a different scenario reduction method that aims to match certain statistical properties of the original scenario set and the reduced scenario set. Recently, Papavasiliou and Oren [25] proposed an importance-sampling inspired approach that aims to select scenarios that best represent the impact of uncertainty on average costs.

Power systems have been a rich source of problems for stochastic optimisation, so researchers in the field have had a keen interest in scenario reduction methods. The distance-based method has been popular. Gröwe-Kuska *et al.* [26] were the first to implement that method in a portfolio management problem for a hydro-thermal power system, while Morales *et al.* [27] proposed a variant that works better for electricity market problems. Carrión *et al.* [28] applied the Kantorovich distance for a consumer's electricity procurement problem. Other methods such as

clustering and importance sampling have also been used. Feng and Ryan [29] proposed to cluster the original scenarios in a generation planning model before performing the forward-selection algorithm in [23]. Papavasiliou and Oren [25] applied their importance-sampling inspired method to multi-area stochastic unit commitment.

However, in spite of the wide range research on scenario reduction in the context of power systems, there has been little application to transmission expansion planning. In one exception, Yu *et al.* [30] proposed a robust transmission expansion planning method with Taguchi's orthogonal array testing, but they consider only a narrow set of uncertainties, and not the full set of economic, technology, and policy risks typically of concern to planners. Also, there has been a lack of research that compares the performance of different scenario reduction techniques for a single power systems problem. The only significant comparison work in power was done by Dvorkin *et al.* [31], who compare multiple scenario reduction techniques in the context of unit commitment. Sun *et al.* [32] present an objective-based scenario selection framework for transmission planning which can potentially be widely used for future research. Here, we will tailor the most promising scenario reduction methods to a continental-scale transmission network planning model, and compare their performance.

3

Methods

In this section, we first summarise the stochastic transmission planning model that we use in this research. Then, we explain how the original (full) set of 20 scenarios were generated. After that we introduce three existing scenario reduction methods and an additional novel method that we develop in this paper.

3.1

Basic stochastic planning framework – JHSMINE [21]

JHSMINE employs a stochastic two-stage optimisation that explicitly models uncertainty using scenarios. For a detailed description, refer to [21]. The model is a mixed-integer linear program, where binary variables represent grid reinforcements. Its objective is to minimise the present worth of total investment and operating costs, subject to various constraints:

Minimise: Probability-weighted present-worth (PW) of investment and operations costs over 40 years

Subject to:

System constraints (supply–demand balance at each node, transmission flow limits, disjunctive constraints to define flow constraints for new lines)

Generator constraints (capacity)

The set of decision variables in this model capture transmission and generation investment decisions, as well as operational decisions. They can be divided into first stage and the second stage decision variables, and because the uncertainty manifests itself between the first and second stage, the second stage decisions are scenario dependent. In other words, there is a same set of second-stage decision variables for each scenario. Furthermore, the choice of scenario set influences the feasible region of the problem because there are constraints in the model that represent the relationships between first- and second-stage decisions. Finally, the scenario set also influences the objective function by assigning different cost coefficients for different scenarios. It is clear that the choice of scenarios will greatly influence the optimal solution of the problem at hand.

One of the expected benefits of including a larger scenario set is that the model is then capable of incorporating a wider range of possibilities for the future, and thus will find a solution whose performance is robust against diverse situations. The disadvantage is that a larger scenario set will make the problem exponentially harder to solve. Scenario reduction tries to lessen the conflict between the objectives of robustness and computational ease. A well-performing scenario reduction method will thus find a scenario set that has relatively fewer scenarios but still results in a solution that is robust against a wide range of possibilities. This paper's goal is to use JHSMINE to compare the performance of reduction methods in a realistic planning context.

The version of JHSMINE used here omits Kirchhoff's voltage law and unit commitment constraints. We have tested the impact of including unit commitment considerations (ramp limits, P_{min} constraints, and startup costs) in long run transmission models [21], and found that they have a noticeable effect only in coal-heavy scenarios. Due to large amounts of announced coal retirements in the western US due to environmental restrictions and low gas prices, such generating units will become a decreasingly important part of the generation mix. We can therefore omit unit commitment constraints, which allow us to solve a significantly larger model with more scenarios. Other versions of JHSMINE include those constraints, however at a cost of including fewer production hours and long-run scenarios. As our focus is on comparing alternative methods for representing long-run uncertainties, we simplify these aspects of the problem so that more long-run scenarios can be considered. For the simulations presented in our paper, we use *k*-means sampling to select 24 h from the available historic hourly data.

In our application to the WECC region, there are 328 buses, 53 backbone transmission lines, and 52 WREZ (Western Renewable Energy Zones) lines. Commitments to investment, both transmission and generation, occur in 2014 (first stage, with an in-service date of 2024) and 2024 (second stage, in-service date of 2034). Planners are assumed to know which scenario has been realised in 2024, just before the second stage investment decisions are made. Therefore, first stage decisions are 'here-and-now' decisions made without knowing what scenario will occur, while second stage decisions are adaptations to the realised scenario (perfect information). Operation decisions are made every year with respect to the current grid situation.

The JHSMINE results that we focus on in this paper include:

The *expected cost of naïve solution* (ECNS), which is the expected cost of the first stage transmission investments from a simplified (fewer scenario) model, but evaluated using all scenarios. This is described further in the next subsection.

First stage transmission investment decisions, which are what planners would focus on because those represent permitting and construction activities that would need to start immediately. These are used in the ECNS calculations.

First stage generation mix, in terms of installed capacity.

3.2

Expected cost of naïve solution

We define ECNS as the actual expected cost of the system across the full set of 20 scenarios if planners implement a 'naïve' first-stage ('here-and-now') transmission solution derived from a model based on a subset of the full scenarios set or the wrong probabilities. By calculating the ECNS, we can assess the performance of investment strategies resulting from the reduced scenario sets. For example, a first-stage solution based on a single 'base case scenario' might have a relatively high cost (when evaluated against the full set of 20 scenarios) compared to a solution developed from a model with all 20 scenarios. The value of the total cost function for the 'ideal solution' (stochastic optimisation considering the full set of scenarios), also called expected cost of the stochastic solution (ECSS) [5], will provide a lower bound for ECNS. The amount by which ECNS exceeds ECSS can be viewed as an expected penalty resulting from using a simplified scenario set and probabilities to solve for the first-stage transmission investments. If the ECNS of an investment plan resulting from a reduced scenario set is insignificantly larger than the ECSS, the scenario reduction is deemed successful.

We calculate ECNS as follows in this paper:

(1) 20 scenarios with equal probabilities (i.e. the original scenario set) are assumed to represent the 'true' reality of the uncertain future.

(2) A solution is obtained from the stochastic optimisation using a reduced scenario set. The resulting first-stage transmission investment decisions are noted.

(3) Those first-stage transmission investments from (2) are fixed (as parameters) in the full stochastic model from (1), which is then re-solved, choosing the optimal solution from the remaining variables against the full scenario set.

The remaining variables include all generation investments, second stage transmission investment, and all generation operations. First-stage generation investment decisions are not fixed because we assume that the generators instead use the 'correct' set of scenarios and probabilities to make investments, accounting for the network design. If instead we fixed the first-stage generation investments at their 'naïve' values, then the expected penalty for naïve decision making would be much larger, and we would be assuming that the transmission planner controls the expectations of generators about the future. In order to avoid overstating the cost of simplifying the scenarios, we only fix first-stage transmission decisions, as in [6].

(4) The objective value (PW of investment and operating costs) of the solution from (3) is called the ECNS.

The process of calculating the ECNS is basically the testing of a subpar investment plan against the 'assumed' true reality.

3.3

Scenario development

The consideration of multiple long-run scenarios and their probabilities are what make our model a stochastic program. We have developed 20 scenarios representing a wide range of possible future technology, policy, and economic developments over the next two decades in collaboration with a project Technical Advisory Group (TAG) that was formed by WECC. TAG included several stakeholders from power companies, public interest groups, and public agencies. Each scenario describes a possible future for 2024–2054 with regards to the uncertain variables, whose values vary among the scenarios. Here we describe the three-step procedure used to define the 20 scenarios that we assume represent the full range of uncertainties; one of the performance metrics for solutions obtained by using fewer scenarios – the ECNS – will be calculated against this complete set. We also note that as opposed to other stochastic problems that can consider hundreds of scenarios, for the transmission planning model that we consider, 20 scenarios is the highest number of scenarios we can use to solve the problem within reasonable time.

Uncertainties included: Our choice of uncertain variables was based upon an initial list generated by a panel of WECC stakeholders and experts – the TAG – which we then trimmed in order to reduce the complexity of the model [21] (e.g. storage was initially identified as an interesting uncertain variable but was judged to be less important than other uncertainties and so was dropped.). The need to limit the number of variables and consequently the number of scenarios emphasises the need for an effective scenario reduction technique. The result was a list of 11 uncertain variables (Table 1).

Table 1 Cluster and uncertain variables modelled in JHSMINE

Cluster	Uncertain variable	Unit	Low	Level medium	High	Cluster-variable relation
1	natural gas price, <i>P</i> ^{gas}	% change in [\$/mmBTU]	-51.2	0	86.2	?+?
	carbon price, Pcarbon	% change in [\$/ton]	33.38	58	112.8 4	?+?
	coal price, P ^{coal}	% change in [\$/mmBTU]	-22	0	22	?+?
	WECC mean load growth, <i>G</i> ^{load}	%/year	-0.91	1.13	3.2	?+?
2	national RPS, RPS ^{fed}	% of WECC-wide demand	0	0	15	?+?
	state RPS, RPS ^{state}	% change of base value	-50	0	50	?+?
	distributed generation, <i>G</i> ^{DG}	% of peak load	3.2	11	20	?+?

Cluster	Uncertain variable	Unit	Low	Level medium	High	Cluster-variable relation
	wind capital cost, C ^{wind}	% change in [\$/installed MW]	-18.3	0	7.49	?
	solar capital cost, C ^{solar}	% change in [\$/installed MW]	- 28.75	0	30	?
	geothermal capital cost, <i>C</i> ^{geo}	% change in [\$/installed MW]	-15	0	10	?
3	WECC peak load growth, <i>G</i> ^{peak}	%/year	-0.37	1.28	2.64	?+?

Ranges of values: Each uncertain variable was assigned high and low values, representing ~90% confidence intervals. The group's low and high values were obtained by averaging the individual responses of the group members. Medium values are based on the WECC 2034 common case values [21].

Clustering uncertain variables: For the scenario development process, uncertain variables were grouped into three clusters. This was done for two reasons. The first was to reduce dimensionality: the 11 uncertain variables in our model are correlated with each other to a certain extent. By grouping variables into subsets whose members are relatively highly correlated, we can reduce model dimensionality and the number of scenarios in the full set.

The second reason was the lack of data on uncertain variables: because there is little relevant historic information for many of the uncertain variables, expert judgment is needed to define the distributions of these variables. It is therefore desirable to have a conceptual framework that relates the variables to each other and makes it easier for us to define a coherent story about how the scenario might arise and influence socio-economic events. For instance, for a future scenario with strong economic growth, breakthroughs in renewable technology, and aggressive clean energy policies, we can define values of variables in one of our clusters (cluster 2) so that it has ambitious portfolio standards and low renewable capital costs.

The uncertain variables were clustered based both on the nature of their economic impact upon the electricity market, as well as their historical correlations. Cluster 1 contains uncertain variables that impact bulk electricity generation. Cluster 2 contains variables that are related to renewable energy, and Cluster 3 contains those that concern system reliability. These clusters can be expanded; e.g. future research could add a fourth cluster related to climate change including variables such as temperature rise and stream flows. Table 1 shows the list of uncertain variables that were included in the model used for this research, and the clusters to which they were assigned. Table 1 also shows the high-medium-low values for the variables, and the direction of their relationship with the cluster, accounting for whether correlations are positive or negative.

Scenario generation: Based on discussions with the TAG and the above-defined clusters and uncertain variables, we expanded the original set of five WECC scenarios used in their 2013 TEPPC process to a total of 20 scenarios in an attempt to capture the full range of possible outcomes [21]. Fig. 1 shows the characteristics of these scenarios. Cells that are coloured dark grey represent high values, light grey represents medium, and white represents low values.

Fig. 1 Levels of uncertain variables in original 20 scenarios

The purpose of the three established sampling methods [random sampling (Rand), importance sampling (IS) and distance-based (DB)] is to select subsets of these 20 scenarios to include in the stochastic planning model. These methods are described in Sections 3.4.1–3.4.3. This paper also proposes a fourth method (stratified scenario sampling, SSS), which instead defines a small set of scenarios that can include one or more new scenarios in addition to the set of 20 above, as explained in Section 3.4.4. In our application to WECC, scenario subsets of size

three or seven are considered by the DB, IS, and SSS methods, while the Rand method considers a range of subset sizes.

Three scenarios represent one possible practical compromise between the computational efficiency of deterministic (single scenario) modelling and the desire to consider the robustness of transmission plans in the face of multiple scenarios. The SSS method aims to characterise the range of possibilities in the 'uncertainty variable space' with a minimum number of scenarios. This method will assign a stratified value for each uncertainty variable for each scenario such that the stratified value for the uncertainty variable will appear only once across all scenarios, similar in philosophy to Latin hypercube sampling [33]. The difference is that the SSS method clusters the variables before performing the Latin hypercube sampling. A specific instance of the algorithm is described at the end of the next section. We can generalise the algorithm by increasing the number of strata (which increases the number of scenarios) or clusters.

3.4

Scenario reduction methods

In this subsection, we provide brief descriptions on the scenario sampling methods that will be used in this paper. Define *S* as the set of original scenarios, *S'* as the reduced scenario set, *C* as the set of clusters, and *V* as the set of selected uncertainty variables used to define a scenario. V(s) will denote the vector of uncertainty variable values for scenario *s*. Also, let |X| denote the cardinality of a set *X*.

3.4.1

Random sampling

Random selection of a small subset scenario represents a baseline, in that more sophisticated reduction methods are worthwhile only if they perform better. Random samples of size 2, 3, 5, 7, 9, and 14 scenarios are repeatedly drawn and tested; ten independent random samples re-chosen for each sample size and the resulting stochastic programs solved, assuming equiprobable scenarios. Further, each of the 20 possible deterministic (1 scenario) models are also solved.

3.4.2

Importance sampling inspired method

Papavasiliou *et al.* use an importance sampling inspired approach to reduce scenarios in a unit commitment problem [25]. We apply this method to sample ten independent sets of three scenarios each, and again to obtain ten sets of seven scenarios. We will name the former IS-3 and the latter IS-7.

Following their method, we first calculate the expected cost $C_D(s)$ of the system (by the model shown in 4.1) for a deterministic transmission investment plan optimised for each scenario *s*

∈

S. Then the probability of selecting each scenario in the sampling procedure, q_s , is calculated by (1) and the weight of the selected scenario, p_s , that goes into the optimisation problem (section 3.1) is calculated by (2) (1)

$$q_{s} = \frac{C_{D}(s)}{\sum_{s} C_{D}(s)};$$

$$p_{s} = \frac{1}{q_{s}}$$

Note that the probability of selecting a scenario is proportional to its deterministic cost. The weight (what we call probability of scenario in other parts of the paper) tries to account for the biasing that occurred in the sampling process when using (1).

3.4.3

Distance-based method

Scenario reduction methods that aim to minimise some distance metric between the original scenario set and the reduced scenario set have often been proposed. One method is to minimise the Kantorovich distance [22, 23, 26], which we use in the fast forward selection DB method. Starting with a set including just the single scenario that is on aggregate closest to the others, that method iteratively adds a single scenario to the set that results in the smallest distance to the non-included scenarios, and continues until the desired number of scenario remains. More specifically, the algorithm from [26] is tailored for our purposes (see below). After finding the reduced scenario set, the probability of each omitted scenario is absorbed into the closest selected scenario. This procedure yields a single set of DB scenarios and probabilities to be tested.

Algorithm 1

Fast forward selection method (based on [26])

Step 0: Compute L-2 norm distances of each scenario pair

$$c_{ij}^{[1]} = \|V(i) - V(j)\|_2, \quad i, j \in S$$

Step 1: Compute

$$z_{i}^{[1]} = \sum_{\substack{j \in S \\ j \neq i}} p_{j} c_{ij}^{[1]}, \quad i = 1, \dots, n$$

Then choose

 $i_1 \in \operatorname*{arg\,min}_{i \in S} z_i^{[1]}$, and set $S^{[1]} = S \setminus \{i_1\}$

Step k: Compute

$$c_{ij}^{[k]} = \min \left\{ c_{ij}^{[k-1]}, c_{i_{k-1}j}^{[k-1]} \right\}, \quad i, j \in S^{[k-1]}$$

$$z_{i}^{[k]} = \sum_{\substack{k \in S^{[k-1]} \\ k \neq i}} p_{j} c_{ij}^{[k]} \quad i \in S^{[k-1]}$$

Then find

 $i_k \in \underset{i \in S^{[k-1]}}{\operatorname{argmin}} z_i^{[k]}$, and set

$$S^{[k]} = S^{[k-1]} \setminus \{i_k\}$$

Step m + 1: The reduced set $S' = \{i_1, \ldots, i_m\} = S - S^{[m]}$

3.4.4

Stratified scenario sampling

Below is a summary of the SSS algorithm for the case in which the reduced scenario set has cardinality 3. The algorithm is readily generalised for higher number of scenarios.

Step 1: For each random variable v

 \in

V, the modeller selects three 'stratified' values from the range of values. For instance, for P_{gas} select three values for P_{gas} Low, P_{gas} Medium, P_{gas} High. The resulting data will look like Table 1.

Step 2: Partition the |V| random variables into |C| clusters. This clustering process groups random variables according to certain criteria. An example of this process is explained in Section 4.3. For our case study, where |C| = 3, the result of the partition is shown in Table 1.

Step 3: Create a 'base case' scenario that has 'medium (M)' stratum assigned for every cluster. For each of the other two scenarios, assign either low (L) or high (H) from each cluster so that low and high each occur in at least one scenario. An example would be: scenario 1 (MMM), scenario 2 (HLH), scenario 3 (LHL).

Step 4: For each scenario, the value of the random variable is decided by the stratum of the cluster that it is part of, and the cluster-variable relationship. For instance, if the cluster is assigned 'high' and the cluster-variable relationship is '(+)', the variable will take on a 'high' value. However, if the cluster-variable relationship is '(-)', the variable will take on a 'high' value.

The above algorithm yields small sets of scenarios (here, 3) to which probabilities must be assigned. For instance, considering all the possible combinations in step 3, there are four possible distinct sets of scenarios: [MMM, HHH, LLL], [MMM, HLL, LHH], [MMM, HHL, LLH]. Two assignment methods are considered. *Equal probabilities* have the advantage of simplicity, and avoid putting the bulk of the weight on one of the scenarios (method SSS-E). In contrast, the advantage of *moment matching* (method SSS-MM) is that the resulting distribution of uncertain variables more closely resembles the original (20 scenario) distribution.

Selecting scenarios and probabilities to match certain statistical properties was previously suggested by Hoyland and Wallace [24]. Moment matching was also adopted by CAISO [15]. We use this method to assign probabilities to the scenarios selected by the SSS method to check if a differentiated set of probabilities will perform better than equal probabilities. Moment matching assigns probabilities to the SSS reduced scenarios so that the resulting mean, standard deviation, and covariance of selected crucial variables come close to those of the original set of 20 scenarios. The chosen variables are 'natural gas price', 'carbon price', 'load growth', 'peak load growth', 'wind capital cost', 'PV capital cost', and 'geothermal capital cost'. The crucial variables were selected based on the result of a workshop where TAG members evaluated each variable's relevance to the economic valuation of transmission additions.

An optimisation model performs the matching, based on normalised values of the uncertain variables (rescaled so that 0 is the low value and 1 is the high value). The first term in the minimisation objective below penalises differences among scenario probabilities in order to encourage the model to assign significant probabilities to each scenario. The second term penalises deviations of variable probability-weighted means from their averages among the 20 scenarios. The other terms similarly penalise deviations of standard deviations and covariances from their values for the full 20 scenarios. The weights on each term can be customised (3)

MIN

$$z = \frac{1}{3} \sum_{s=1}^{3} p_{s}^{2} + \frac{10^{4}}{7} \sum_{i=1}^{7} \left(\widehat{Mean}_{j} - Mean_{j} \right)^{2}$$

$$+ \frac{1}{7} \sum_{i=1}^{7} \left(\widehat{SD}_{j} - SD_{j} \right)^{2} + \frac{1}{21} \sum_{i=1}^{7} \sum_{j>i} \left(\widehat{COV}_{i,j} - COV_{i,j} \right)^{2}$$
Subjectto :

$$\sum_{s=1}^{3} p_{s} = 1 \text{ and } p_{s} > 0 \quad \forall s$$

where

 p_3

is the probability of scenario s,

is the mean value (across the full 20 scenarios) for the uncertain variable j, and

Mean_j

is the probability-weighted average of the uncertain variable for the reduced scenario set. Other terms are defined analogously, where SD and COV stand for standard deviation and covariance, respectively.

4

Case study: WECC 300

The performance of the reduction methods is assessed by solving the JHSMINE model with each reduced scenario set and then comparing the solutions by three criteria. First, the ECNS of the obtained solution is compared with the ECSS (refer to Section 3.2). Reduction methods that yield lower values of ECNS (and are therefore closer to the ideal of ECSS) are preferable. Then in Sections 4.2–4.3, we contrast the transmission and generation solutions for the best reduction methods with the ideal solution, and finally consider how the reduction methods compare in terms of a regret measure (Section 4.4).

4.1

Criterion 1: expected cost of naïve solution

Fig. 2 summarises the results of the different scenario reduction methods as box-whisker plots. The runs shown include the random method Rand (10 samples each for sets of 2, 3, 5, 7, 9, and 14 scenarios, plus the full 20 scenario case), the 10 sets of results for the IS-3, IS-7, and DB, and four runs each for the SSS-E and SSS-MM methods, as described above.

Fig. 2 Box-and-whisker plot showing the ECNS for different scenario sampling methods. (top and bottom whiskers show max/min values; top and bottom box edges are the third and first quartiles. The line in the middle of the box is the median. The number on top is the mean. The numbers on the x-axis are the number of random scenarios used. The grey dotted line is ECSS

For example, when using random sampling (Rand) with 14 scenarios ('Rand-14'), the mean (across the sample of 10 stochastic optimisations) of the value of ECNS (\$881.8B) is just 0.12% above the 'ideal' solution (ECSS = \$880.7B), with the worst Rand-14 ECNS being 0.47% worse (\$884.8B) than the ideal. Although this penalty is not large, neither are the computational savings from using 14 scenarios versus the full 20 scenarios.

As we expected, Rand plan performance degrades as we use fewer scenarios. The Rand results with 5, 7, and 9 scenarios show median performance (across each of their sample of 10 runs) similar to Rand-14, but their worst cases are much costlier, resulting in higher mean costs than for Rand-14 solutions. For instance, the mean ECNS for Rand-3 solutions is \$887.3B; this mean penalty of \$6.6B relative to ECSS is comparable in magnitude to the present worth of the first stage transmission investments themselves and therefore can be viewed as important. Single backbone lines typically cost several hundred million dollars, and the present worth of first-stage backbone transmission investments is between \$1B and \$6B, depending on the solution (0.1–1% of the objective). Note that the bulk of the objective's value consists of fuel and generation investment costs.

Meanwhile, the deterministic runs (Rand results with 1 scenario each) vary greatly in their results, producing ECNS values ranging from 0.15 to 4.17% above the ideal solution. The deterministic run with just using scenario 11 performed the best among the deterministic runs, and even did better than some of the Rand-14 runs. This reinforces the point that more scenarios do not necessarily improve the solution, and intelligent selection of scenarios can be much more beneficial than simply adding a lot more scenarios.

Better performance is obtained with scenarios chosen systematically by the SSS, IS, and DB methods. First, for the four SSS-E cases (equally weighted three scenarios), their average ECNS (\$883.8B) exceeds the ideal solution (ECSS) by 0.35%. The four solutions from the SSS method with moment-matched probabilities (SSS-MM) perform even better (mean ECNS = \$882.7B, 0.22% above the ideal). From among the SSS-MM results, run SSS-MM1 (extreme sample consisting of HHH, MMM, LLL) gives the best ECNS (\$881.4B), just 0.08% above the ideal.

Meanwhile, the IS method with three scenarios gives an ECNS value that is on average 0.37% above ECSS. However, performance is highly inconsistent among the samples. Using importance sampling with seven scenarios gives better results but the computational savings are then modest relative to the full 20 scenario case. Finally, the distance-based method (DB) with three scenarios yields an ECNS (881.5) that is only 0.09% above the ECSS ideal. Thus, with a 85% reduction in the number of scenarios (20 to 3), we are still able to obtain solutions that closely approach the cost of the ideal stochastic solution.

4.2

Criterion 2: transmission investment

Similar expected costs may result from different configurations of the network. Therefore, a careful comparison of two different solutions should contrast not just their total costs but also their transmission investment decisions. From this section on, we focus mainly on the runs that yield low ECNS values. These are SSS-E1, SSS-E2, SSS-MM1, SSS-MM2, and DB. To clarify what we mean by, for instance SSS-E1, this means that it is the first set of scenarios sampled using the SSS-E method. Accordingly, SSS-MM2 means that it is the second set of scenarios sampled using the SSS-MM method. We will occasionally discuss other runs for comparison.

Candidates for transmission line investments include backbone network line additions and WREZ lines. The former is a duplicate of the largest (or sometimes second largest) capacity line in each transmission path. Using binary variables, the model must build a full line or none at all. In contrast, WREZ lines, which connect renewable resources to the existing network, are represented by continuous variables, and can be built in any size between zero and the maximum capacity.

Fig. 3 shows the first-stage backbone line additions. The ideal solution (in the box) suggests that we build eight backbone transmission lines in the first stage. The DB method provides the closest solution because it recommends seven of these eight lines. Meanwhile, the SSS runs shown build only four or five transmission lines. In general, although Figs. *3b–d* show solutions that are not equal to the ideal solution, the lines chosen are all subsets of the ideal solution's set of lines. This is also true for the SSS runs that are not shown in Fig. 3 (SSS-E3, SSS-E4, SSS-MM3, SSS-MM4).

Fig. 3 Plot of 'suggested year 2024 backbone transmission investments' obtained by solving JHSMINE. Blue dotted circles show where the investment decisions are located

(a) Ideal solution, (b) Solution for SSS-E1 and all SSS-MM runs, (c) Solution for SSS-E2, (d) Solution for DB

In Fig. 4, we show the first-stage WREZ line additions. Since WREZ line additions are modelled as continuous variables, the solutions are presented via a colour gradation. Yellow indicates a low investment level and red indicates a high investment. Although Fig. 4 gives a geometric depiction of the solutions, it is difficult to quantitatively compare different solutions. Therefore, as one index of the level of agreement between different solutions, the sum of absolute differences in WREZ line investment decisions in MW with respect to the 'ideal solution' is calculated and graphed in Fig. 5. We can see that the WREZ investment decisions for the SSS-MM1 (extremely stratified sampling with moment matching) are closest to the ideal solution. SSS-E1 and DB also performed relatively well. Just like with the ECNS results, use of moment-matched probabilities (instead of equal probabilities) improves the performance of the SSS method.

Fig. 4 Plot showing the 'suggested year 2024 WREZ investments' obtained by solving JHSMINE (best viewed online in colour)

(a) Ideal solution, (b) Solution for SSS-E1, (c) For SSS-E2, (d) For SSS-MM1, (e) For SSS-MM2, (f) For DB. For each WREZ line, the model is allowed to choose between 0 and 100% of the maximum investable capacity. Line colours vary from yellow to red; yellow is a low amount of capacity and red is high

Fig. 5 Sum of absolute differences in WREZ line investment decisions with respect to the 'ideal solution' for SSS-E, SSS-MM, and DB

4.3

Criterion 3: generation mix

So far, in terms of ECNS and transmission plans, the extreme SSS-MM (i.e. SSS-MM1) and DB methods seem most promising. Another criterion for comparing the performance of different scenario reduction methods is the difference in the mix of first-stage generation investments relative to the ideal. We focus on the first-stage generation investment decisions from the ECNS solutions, which assume that non-naïve generators consider all 20 scenarios, subject to the first-stage grid decisions that naïve grid planners obtain using the reduced scenario set.

The ideal solution projects that generation investment in 10 years will be mostly on-shore wind energy (39.7 GW) and combined cycle capacity (37.2 GW) with a small percentage of gas CT (3.7 GW), distributed solar PV (2.0 GW), and geothermal (2.1 GW). Fig. 6 shows the generation mix by percentage for the SSS-E, SSS-MM and DB solutions.

Fig. 6 First-stage generation investments for the ideal (20 scenarios), DB, SSS-E, and SSS-MM solutions

We can see that the DB method provides a solution that is very close to the 'ideal'. It underestimates gas CT and overestimates solar by a small percentage. Meanwhile, the SSS-E solutions anticipate too much gas and too little wind compared to the ideal solution. The performance improves once we use moment-matched probabilities. The extreme SSS-MM (i.e. SSS-MM1) mix is close to the ideal solution, although not as close as the DB method.

4.4

Criterion 4: maximum regret

Expected cost is not the sole metric by which decisions are made. Risk, which might be defined as the chance of very poor performance, is also a major concern [34, 35]. A plan that does well in expected value terms might do poorly in terms of risk. A recent paper by Konstantelos *et al.* [36] proposes a min–max regret optimisation model to address this issue. Although we do not incorporate risk into our initial optimisation model, it would be interesting to compare how each solution performs in terms of this criteria. As we show below, some reduction methods appear to do as well or better in terms of risk than the full 20 scenario solutions.

In this paper, we consider one possible risk metric, which is the maximum regret (Max-regret) (across the 20 scenarios) of each first-stage transmission solution. The calculation of regret here is based on the standard definition [6]: regret for a particular first-stage plan x is calculated by taking the PW of the cost of that plan for a scenario s (calculated by the ECNS metric in Section 4.2, and designated as CNS(x,s)), and then subtracting the cost of a plan $x^*(s)$ optimised for that same scenario, $C(x^*(s),s)$. The Max-regret is then calculated as (4)

$$\max_{s \in S} \left\{ CNS\left(x, s\right) - C\left(x^{*}\left(s\right), s\right) \right\}$$

Fig. 7 plots the tradeoff between ECNS and Max-regret for all the sampling methods that we tested. The Pareto frontier (red line) of that figure shows that the solutions obtained by using the SSS-E, SSS-MM and DB methods do slightly worse in terms of ECNS compared to the ideal solution (20 scenario stochastic solution) but actually do much better in Max-regret. For instance, the ECNS for SSS-MM1 is \$0.7B higher than the ideal solution but the Max-regret is \$5.9B lower than the ideal solution; the ECNS for DB method is \$0.8B higher than the ideal solution but the Max-regret is \$2.4B lower. The higher maximum regret for the ideal solution results from a relatively poor preparation for the extreme scenarios.

Fig. 7 Maximum regret versus ECNS for all reduction methods. The dotted curve on left shows tradeoffs between the objectives (best viewed online in colour)

These results indicate that preparing for so many future scenarios can indeed put the system in danger of being less robust against extreme cases. Whether this is a general result that applies to other power systems would need to be confirmed by additional case studies.

5

Discussion

5.1

Solution time versus accuracy tradeoffs

All models were run to full optimality (zero MIP gap). Table 2 shows the solving times and ECNS gap (ECNS-ECSS) for all scenario reduction methods. The gaps are calculated as in Section 3.1. As noted before, a 0.5% divergence from the optimal objective value (total costs) is not trivial because it is on the same order of the total transmission investment in the first stage. Table 2 shows that only SSS-MM and DB provide solutions that are consistently within 0.5% of the ideal while also reducing solution times significantly. Among the random sampling solutions, only Rand-14 solutions are within 0.5% of the ideal solution on average, but their solution times improve little upon the full 20 scenario model. Although solution times of several hours are not necessarily unreasonable for a planning model, we should remember that this planning model has many other simplifications.

Table 2 Solution time and ECNS gap

Sampling method	Range of time, s	Range of ECNS (as % above the ECSS, see Fig. 2)
Rand-1	[13, 37]	[0.15, 4.17]
Rand-3	[94, 147]	[0.03, 3.18]
Rand-14	[2486, 8594]	[0.01, 0.47]

Sampling method	Range of time, s	Range of ECNS (as % above the ECSS, see Fig. 2)
SSS-E	[86, 102]	[0.18, 0.53]
SSS-MM	[77, 90]	[0.08, 0.35]
DB	94	0.09
IS-3	[89, 132]	[0.02, 2.09]
IS-7	[631, 1109]	[0.02, 0.49]
20 full scenario	33,600	0 (by definition)

Adding, e.g. unit-commitment, Kirchhoff's voltage law, or a more detailed network will result in much longer solution times, so planners are likely to appreciate the one to three order of magnitude improvements in solution times that some scenario reduction methods yield.

5.2

Additional advantages of SSS-MM: lower Max-regret and less backbone investments to commit

Although presenting a new scenario sampling method is not the focus of this paper, we note two attractive features of the solutions from the SSS-MM method proposed here. This heuristic method worked surprisingly well for this case study, as shown in Section 5. Investigating the generalisability of this result and developing a improved versions of this method is left to future work.

The first feature is that the SSS-MM solutions reduce risk for this case study, as measured by the Max-regret metric. The DB method also has a lower Max-regret compared to the ideal solution (expected cost minimisation), but the reduction is only half that in the case of the SSS-MM solution. Further research is needed to determine whether this is likely to be a general result, and why it occurs.

Another attractive attribute of the SSS-MM method is that it provides a solution in this case whose cost is very close to the ideal but with less investment in backbone transmission (as shown in Fig. 3). Permitting, public hearings, right-of-way acquisition, and aesthetic impacts all stimulate public opposition to new lines. Therefore, planners and regulators may prefer transmission plans that involve less construction. Several of the solutions share similar ECNS values but vary in the number of backbone transmission line investments. In other words, it appears that we can obtain a near-optimal cost solution as long as we invest in a 'core' set of backbone lines. This implies that some of the lines that the ideal solution picks have benefits only slightly in excess of their costs; so if minimising construction is an objective, it should be possible to do so with little loss of economic efficiency.

6

Conclusion

The motivation for this paper is the conflict between the need for larger models with more realistic detail and multiple scenarios, and computer capabilities that greatly limit the size of those models. By clever selection of scenarios, we show that model size can be drastically reduced while preserving the benefits of stochastic programming, which allows the user to either add other features to the model or execute more runs more quickly. This paper explored the performance of four scenario reduction methods within the framework of stochastic transmission planning. The methods analysed include three existing methods: *Rand, IS, DB*, and an additional method, which we call *SSS*. The performance of these methods was measured by comparing values of ECNS, first-stage transmission investment decisions, and reductions in solution time. The following bullet-points summarise the major findings of this paper:

For scenario reduction, an intelligent reduction of the scenario set can be much more beneficial than simply adding a lot more scenarios. To restate one of the examples, a random scenario reduction that reduces the original scenario set from 20 to 14 scenarios performs worse than a DB reduction method that reduces the original scenario set to three scenarios.

The DB method and the stratified scenario sampling/moment-matching (SSS-MM) method both provide solutions with relatively little loss of cost efficiency when compared to the full 20 scenario model solution. Furthermore, both methods greatly reduce solution times.

Compared to other scenario reduction methods, the overall investment decisions obtained from using the DB method better match the optimum from the most complex model that utilises the full 20 scenarios.

The SSS-MM method has slightly less accurate solutions than those obtained from the DB method but has lower worst-case regret, and lower expenditures on backbone investments in the first stage. This may be attractive to policy makers concerned about environmental impacts of transmission.

Although the complexity of the case study gives some confidence that these results will apply to other situations, these conclusions need to be confirmed with further testing in other planning contexts. A future study should also conduct sensitivity analyses on the number of original scenarios and the number of uncertain variables. In addition, if possible, it is desirable to obtain theoretical results that guarantee the performance of different scenario reduction methods when applied to multi-stage stochastic transmission planning, perhaps by taking advantage of the special structure of transmission planning problems. The ultimate goal is to design scenario reduction methods that can be depended on to yield robust solutions with modest computational effort.

Bibliography

REFERENCES

1 Sauma E.E., Oren S.S.: 'Proactive planning and valuation of transmission investments in restructured electricity markets', J. Regul. Econ., 2006, 30, pp. 358–387

2 Pozo D., Contreras J., Sauma E.: 'If you build it, he will come: antici-pative power transmission planning', Energy Econ., 2013, 36, pp. 135–146

3 Li A., Gobbs B.F., Ho J., : 'Co-optimization of transmission and other supply resources', Prepared for NARUC, 2013

4 Awad M., Casey K.E., Geevarghese A.S., : 'Economic assessment of transmission upgrades: application of the California ISO approach', in Zhang X.-P. (Ed.): 'Restructured electric power systems: analysis of electricity markets with equilibrium models' (IEEE Press, Piscataway, 2010)

5 Munoz F.D., Hobbs B.F., Ho J.L., : 'An engineering-economic approach to transmission planning under market and regulatory uncertainties: WECC case study', IEEE Trans. Power. Syst., 2014, 29, pp. 307–317

6 van der Weijde A.H., Hobbs B.F.: 'The economics of planning electricity transmission to accommodate renewables: using two-stage optimisation to evaluate flexibility and the cost of disregarding uncertainty', Energy Econ., 2012, 34, (6), pp. 2089–2101

7 Pozo D., Sauma E., Contreras J.: 'A three-level static MILP model for generation and transmission expansion planning', IEEE Trans. Power Syst., 2013, 28, pp. 202–210

8 Wong W., Chao H., Julian D., : at Transmission planning in a deregulated environment'. IEEE Transmission and Distribution Conf., New Orleans, LA, USA, vol. 1, April 1999, pp. 350–355

9 Moreno R., Street A., Arroyo J.M., : 'Planning low-carbon electricity systems under uncertainty considering operational flexibility and smart grid technologies', Phil. Trans. R. Soc. A, 2017, 375, p. 20160305

10 Hojjat M., Javidi M.H.: 'Chance-constrained programming approach to stochastic congestion management considering system uncertainties', IET Gener., Transm. Distrib., 2015, 9, pp. 1421–1429

11 Yang N., Wen F.: 'A chance constrained programming approach to transmission system expansion planning', Electr. Power Syst. Res., 2005, 75, pp. 171–177

12 Sun Z., Li W., Zhu J., : ' a planning method for siting and sizing of distributed generation based on chanceconstrained programming'. Int. Conf. on Electric Utility Deregulation and Restructuring and Power Technologies, Changsha, China, 2015, pp. 527–531

13 Charnes A., Cooper W.W.: 'Chance-constrained programming', Mgmt. Sci., 1959, 6, (1), pp. 73–79

14 Hannan J.: 'Approximation to Bayes risk in repeated play', in Dresher M. (Eds): 'Contributions to the theory of games, volume III, annals of mathematics Studies' (Princeton University Press, Princeton, NJ, USA, 1957), 39, pp. 97–139

15 '2011–2012 Transmission Plan', California ISO, Technical Report, March 2012

16 'Transmission Economic Assessment Methodology', CAISO, Technical Report, June, 2014

17 'Regional Generation Outlet Study', MW ISO, Technical Report, November 2010

18 Bhattacharya A., Kharoufeh J.P., Zeng B.: 'Managing energy storage in microgrids: a multistage stochastic programming approach', IEEE Trans. Smart Grid, 2016, early-access

19 Konstantelos I., Strbac G.: 'Valuation of flexible transmission investment options under uncertainty', IEEE Trans. Power Syst., 2015, 30, (2), pp. 1047–1055

20 Qiu T., Xu B., Wang Y., : 'Stochastic multistage coplanning of transmission expansion and energy storage', 2017, 30, (1), pp. 643-651

21 Ho J., Hobbs B.F., Donohoo-Vallett P., : 'Planning transmission for uncertainty: applications and lessons for the western interconnection', Report prepared for the (WECC) and Lawrence Berkeley National Laboratory, 2016

22 Dupačová J., Consigli G., Wallace S.W.: 'Scenarios for multistage stochastic programs', Ann. Oper. Res., 2000, 100, pp. 25–53

23 Dupačová J., Gröwe-Kuska N., Römisch W.: 'Scenario reduction in stochastic programming: an approach using probability metrics', Math. Program., 2003, Ser. A 95, pp. 493–511

24 Hoyland K., Wallace S.W.: 'Generating scenario trees for multistage decision problems', Manage. Sci., 2001, 47, (2), pp. 295–307

25 Papavasiliou A., Oren S.S.: 'Multiarea stochastic unit commitment for high wind penetration in a transmission constrained network', Oper. Res., 2013, 61, pp. 578–592

26 Gröwe-Kuska N., Heitsch H., Römisch W.: 'Scenario reduction and scenario tree construction for power management problems'. Proc. IEEE PowerTech, 3, Bologna, Italy, 2003

27 Morales J.M., Pineda S., Conejo A.J.: 'Scenario reduction for futures trading in electricity markets', IEEE Trans. Power. Syst., 2009, 24, pp. 878–888

28 Carrión M., Philpott A.B., Conejo A.J., : 'A stochastic programming approach to electric energy procurment for large consumers', IEEE Trans. Power Syst., 2007, 22, pp. 744–754

29 Feng Y., Ryan S.M.: 'Scenario construction and reduction applied to stochastic power generation expansion planning', Comput. Oper. Res., 2013, 40, pp. 9–23

30 Yu H., Chung C.Y., Wong K.P.: 'Robust transmission network expansion planning with taguchi's orthogonal array testing', IEEE Trans. Power Syst., 2011, 26, (3), pp. 1573–1580

31 Dvorkin Y., Wang Y., Pandzic H., : 'Comparison of scenario reduction techniques for the stochastic unit commitment'. Proc. 2014 IEEE Power & Energy Society General Meeting, National harbor, MD, USA, 2014, pp. 1–5

32 Sun M., Teng F., Konstantelos I., : 'An objective-based scenario selection method for transmission network expansion planning with multivariate stochasticity in load and renewable energy sources', Energy, 2018, 145, pp. 871–885

33 McKay M.D., Beckman R.J., Conover W.J.: 'A comparison of three methods for selecting values of input variables in the analysis of output from a computer code', Technometrics, 1979, 21, pp. 239–245

34 Loomes G., Sugden R.: 'Regret theory: an alternative theory of rational choice under uncertainty', Econ. J., 1982, 92, (4), pp. 805–824

35 Bell D.E.: 'Regret in decision making under uncertainty', Oper. Res., 1982, 30, (5), pp. 961-981

36 Konstantelos I., Moreno R., Strbac G.: 'Coordination and uncertainty in strategic network investment: case on the north seas grid', Energy Econ., 2017, 64, pp. 131–148

Classification

Language: ENGLISH

Publication-Type: Magazine

Subject: ENERGY & UTILITY SECTOR PERFORMANCE (72%); ENERGY & ENVIRONMENT (70%); ENVIRONMENTAL RESEARCH (69%); ENVIRONMENTAL LAW (67%); ENVIRONMENT & NATURAL RESOURCES (62%); POLLUTION & ENVIRONMENTAL IMPACTS (62%); stochastic processes (%); power transmission planning (%); decision making (%); stochastic programming (%); optimisation (%); probability (%); investment (%); sampling methods (%); scenario reduction methods (%); stochastic transmission planning (%); economic uncertainties (%); net benefits (%); grid reinforcements (%); stochastic optimisation (%); transmission plans (%); model size (%); stochastic programming (%); promising scenario sampling methods (%); economic consequences (%); simplifying scenarios (%); expected cost (%); naïve solution (%); first-stage investment decisions (%); maximum regret (%); multidecadal planning (%); Western Electricity Coordinating Council system show (%); distance-based method (%); stratified scenario section method (%); moment-matched probabilities (%); larger model (%); lower worst case regret (%); careful scenario reduction (%); useful models (%)

Industry: ALTERNATIVE & RENEWABLE ENERGY (90%); SOLAR ENERGY (90%); WIND ENERGY (89%); ENERGY & UTILITY POLICY (87%); ELECTRICITY TRANSMISSION & DISTRIBUTION (87%); ENERGY & UTILITY REGULATION & POLICY (72%); ENERGY & UTILITY SECTOR PERFORMANCE (72%); ENERGY & UTILITY TRADE (72%); ELECTRIC POWER PLANTS (72%); ELECTRIC POWER INDUSTRY (72%); ENERGY & ENVIRONMENT (70%)

Load-Date: April 11, 2019