Final Scientific/Technical Report

A Multistage Stochastic Transmission Expansion Algorithm for Wide-Area Planning under Uncertainty

WORK PERFORMED UNDER AGREEMENT

DE-OE0000881

SUBMITTED BY (Recipient Business Contact)

Ms. Jennifer Lear, Director of Research Administration

College of Earth & Mineral Sciences

The Pennsylvania State University

248 Deike Building

University Park, PA 16802

E-mail: qul48@psu.edu

Phone: (814) 865-7650

Award Period of Performance: October 1, 2017 to September 30, 2020

Submitted: {date final scientific technical report submitted}

PRINCIPAL INVESTIGATOR

Mort Webster

Phone: (814) 863-1640

FAX: (814) 865-3248

E-mail: mort@psu.edu

SPONSORING PROGRAM OFFICE

U. S. Department of Energy

Office of Electricity Delivery and Energy Reliability

Via the National Energy Technology Laboratory

TABLE OF CONTENTS

Acknowledgment

This material is based upon work supported by the Department of Energy under Award Number DE-OE0000881.

Disclaimer

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

LIST OF TABLES

LIST OF FIGURES

LIST OF ACRONYMS AND ABBREVIATIONS

AC	Alternating Current
ALFA	Approximate Latent Factor Algorithm
CIGRE	International Council on Large Electric Systems
CVIS	Control Variate Importance Sampling
DC	Direct Current
EPRI	Electric Power Research Institute
EV	Electric Vehicle
FERC	Federal Energy Regulatory Commission
FOA	Funding Opportunity Announcement
FSS	Forward Scenario Selection
IEEE	Institute of Electrical and Electronics Engineers
INFORMS	Institute for Operations Research and the Management Sciences
IS	Importance Sampling
ISO	Independent System Operator
KMC	K-Means Clustering
LUCA	Latent Uncertainty Clustering Algorithm
MAP	Multistage Adaptive Planning
MILP	Mixed Integer Linear Program
MINLP	Mixed Integer Non-Linear Program
NERC	North American Electricity Reliability Council
NREL	National Renewable Energy Laboratory
OPF	Optimal Power Flow
RPS	Renewable Portfolio Standard
RTO	Regional Transmission Operator
SAA	Sample Average Approximation
SVD	Singular Value Decomposition
TNEP	Transmission Network Expansion Planning
WECC	Western Electric Coordinating Council

I. Executive Summary

Wide-area transmission planning for the electric power system is one of the major remaining challenges today. With the likely changes over the next several decades in the location and technologies for power generation, transmission planners need to plan now for improvements to the high-voltage transmissions system to avoid future congestion. However, planning is complicated by the long lead times needed to plan new transmission investments (typically 10 years in the U.S.) and the uncertainty in the location of future generation that exists when transmission investments must be chosen. In addition, uncertainty in other aspects of the power grid must be anticipated in order to avoid making costly investments that fail to provide value or failing to make investments that will be needed later to ensure reliable grid performance. These uncertainties include changes in demand patterns, changes in fuel prices, changes in regulation or electricity markets, and concerns about resilience to natural stressors on the power system. Finally, all transmission planning requires a simulation of the power flow on the network for many different possible hours, for which demand and renewable output varies across the network, to compare the performance of alternative investment plans. It is not practical to simulate every possible hour for all years of a study when there are many investment plans to test and many long-term scenarios.

Mathematically, the selection of which lines to add or reinforce in a large network with uncertainty and several investment decision periods over time is a large optimization problem that is difficult to solve with existing techniques. Increasing the size of the network considered, the number of future scenarios to consider, or the number of distinct decision points increase the size and complexity of the problem. In practice, most power system planners use computational models to inform their decisions by studying a single investment period in which lines will be added and rely on scenario analysis to address uncertainties. Scenario analysis consists of solving the model to find the best transmission plan for one possible future, assuming that future will occur with certainty. This approach does not allow planners to identify investments that provide flexibility when future conditions are not known or to identify investments that should be postponed and only built in some possible futures.

The overall objective for this project was to develop and demonstrate a set of methods for solving the transmission investment problem for a large network considering many possible scenarios of future conditions and multiple decision points when investments can be made. Project sub-objectives achieved this goal through a succession of extending the methods to apply to problems with increasing complexity or additional features, including the number of decision points, whether generation and transmission are co-optimized, and whether AC or DC power flow is used.

The main innovations in the numerical methods developed in this project were in how the very large set of possible futures could be simplified and organized into smaller problems that can be solved and that approximate well the solution to the full problem. Existing methods reduce problems with many possible scenarios to a size that is feasible to solve by grouping together scenarios that are "similar" in terms of the future conditions. Each group is then replaced by one "representative" scenario, resulting in a smaller set of scenarios. In the method developed in this project, we instead use a variety of mathematical techniques to group together scenarios for which a similar decision is made, and the method ensures that the reduced set consists of scenarios that will best distinguish among the alternative investments. A second innovative aspect of the technical approach is that the methods address both long-term uncertainties for conditions that change slowly over years (e.g., natural gas prices or new generation) and short-term variability that changes on a time-scale of minutes to hours (e.g., demand, renewable generation). Our approach uses a nested representation for long-term and short-term uncertainties; this enables a distinct representative set of hours to be selected for each long-term future and improves the accuracy of the approximation.

A transmission model was developed for the Western Electric Coordinating Council (WECC) region, the high-voltage transmission system that serves the western third of the continental U.S. Using a dataset provided by WECC and by researchers from John Hopkins University, we have validated and demonstrated the model and used it to compare the new method for solving multi-stage stochastic transmission planning to several state-of-the-art techniques.

The project has resulted in several key outcomes and achievements:

- The covariance-based method for choosing a small set of hours to represent short-term variability has superior performance in terms of accuracy to existing methods, including K-means clustering and Importance Sampling;

- The combined partitioning method for long-term uncertainty with the nested clustering approach for choosing representative hours for each long-term group has superior accuracy for equivalent computational effort compared with existing methods;

- Using the partitioning/clustering method combined with Sample Average Approximation provides both statistical bounds on the quality of the solution and at the same time, a complete investment plan for all contingencies in the full uncertainty set; no existing methods can provide both at the same time;

- The method is demonstrated to work well for choosing both transmission and generation investments;

- A variant on the method allows for both scenario selection and simultaneous correction for the error from the DC power flow approximation to provide a tractable method for AC power flow-based transmission planning under uncertainty;

- The method applied to the WECC case study demonstrates the additional value to the system operator and the consumer of identifying flexible investment options in the near-term decisions. In particular, the case study exhibits significant option value in postponing some transmission additions that appear useful but in some long-term system states create new congestion problems.

II. Objectives

<u>Technical Background</u>

Transmission Network Expansion Planning Models

One of the primary computational tools to assist in transmission planning is a Transmission Network Expansion Planning (TNEP) model. The simplest and most common version of a TNEP model is an optimization model that finds the minimum cost transmission plan (which new lines to add) that meets some reliability constraint. The subproblem in this optimization is an optimal power flow (OPF) model for one or more time periods, each with a different pattern of demand and renewable output across the network. The OPF subproblem solves for the minimum cost generation from all dispatchable generators to meet the demand at all buses, subject to Kirchhoff's laws of current and voltage and enforcing maximum power flow constraints on transmission lines.

Although the OPF can be formulated as a convex (linear) mathematical program, the discrete decisions for each candidate transmission line (build or do not build) make the TNEP problem non-convex. TNEP models are typically formulated as mixed integer linear programs (MILPs). The computational complexity of solving even a static (one future year) deterministic (no uncertainty) TNEP can be considerable for a network with large numbers of buses and transmission lines.

Before renewable generation became a substantial part of grid operations, it was often considered sufficient to evaluate new transmission lines only for the hour of the year with the highest load. Any transmission network that allowed for adequate operation during this hour was considered likely to operate at least as well during any other demand scenario [1]. This “worst hour” approach to selecting a test scenario is still deceptively appealing, but in modern systems, it is not obvious which hour is most likely to cause reliability issues, and the worst hour may vary by region or change as new lines are added and reliability issues are resolved. The "worst hour" approach also provides no indication of the total operations cost of the proposed network over a full year, so line investments that would significantly reduce operation costs, but are not necessary for the reliability of the system, are likely to be missed. Transmission planning for today's system requires the consideration of many possible operating conditions.

Standard practice is to include a small subset of the possible demand and generation scenarios in the TNEP optimization model. However, it is difficult to determine which scenarios are relevant a priori. There is a need for methods that can accurately determine the performance of a network in the relevant operating conditions without requiring an excessive number of representative hours [2].

One class of problems referred to as stochastic TNEP problems are in fact single period investment models with a second recourse stage where the OPF 9 (i.e., dispatch) is solved for each realized pattern of demand. These one-stage TNEP models are designed to address the short-term (hourly) variability of load and renewable output and are better thought of as approximating the performance of transmission plans over one year. The first objective below is focused on finding a solution for how to best approximate the short-term variability with a small number of representative hours.

Multi-Stage Stochastic Transmission Planning

In addition to the short-term uncertainty discussed above, there is a need for planning to anticipate the needs of the system in the future as technological, economic, and regulatory conditions change. Within the coming decades, these changes are likely to involve dramatically increased renewable generation requirements and/or strict carbon limits, coal generator retirements, possible nuclear retirements, rapidly decreasing costs of energy storage, increased electric vehicle adoption, electrification of industrial and residential energy use, uncertainty in the cost and role of natural gas generation, modifications to electricity market design, and increasing concern about power system resilience to non-traditional external stressors [3-6]. Any additions to the transmission network should provide a net improvement across the possible realizations of these uncertain trends, but each scenario will exhibit distinct spaciotemporal distributions of supply and demand that benefit from different transmission topologies. Thus, different scenarios will benefit from different transmission investments, and an expansion plan must be adaptable to many outcomes. By ignoring uncertainty, planners not only risk overlooking strategic near-term investments that provide future flexibility, they also risk undertaking unnecessary, or even detrimental, projects.

A formal mathematical model to address long-term uncertainties and to represent future transmission investments that are contingent on the system state is a two-stage or a multi-stage (three or more) stochastic program. In a two-stage stochastic TNEP model, one set of transmission investments is made in the first stage of the problem, before uncertainty is realized, and then recourse investments are made in the second stage after the uncertain events have resolved. Modeling transmission investments as recourse decisions is often referred to as adaptive transmission planning, with notable examples [7-10]. Adaptive TNEP problems are even more difficult to solve than the one-stage stochastic problems described above, because each scenario adds another set of binary recourse variables (additional line investments in the second stage) to the analysis and the problem quickly grows too large to solve. A major focus of adaptive TNEP research is to reduce the number of scenarios to a manageable size while still providing a reasonable representation of the long-term uncertainty [10].

The effects of short-term uncertainty are compounded in the adaptive planning paradigm because most of the factors that are uncertain in the long-term will alter the short-term characteristics of the system. Each scenario will have a distinct combination of technological and regulatory parameters that manifest unique short-term behavior, and the specific spatial patterns of demand and renewable generation that cause congestion will occur at different times across scenarios. The short-term uncertainty is nested within the long-term scenarios. An insufficient representation of this interaction will produce inaccurate and/or biased estimates of the expected benefits of candidate investments in each scenario, undercutting the value of using an adaptive framework. A sophisticated and efficient handling of the full joint distribution of long-term and short-term uncertainty is required to properly represent this interaction while remaining computationally tractable.

The challenges described here for adaptive transmission planning motivate the additional goals of this project to develop methods for reducing the number of long-term scenarios to obtain a tractable unbiased approximation of the solution to the original stochastic problem, and to select a distinct set of representative hours to approximate short-term variability for each long-term scenario or group of scenarios.

Additional Challenges to TNEP

Incorporating additional features enable a computational model to capture important aspects of the true problem faced by system planners. In some regions, investments in new generation sources are not independent of transmission investments, and the simultaneous planning or co-optimization of generation and transmission can lead to better decisions. The addition of decision variables for candidate generators in each stage dramatically increases the size of the decision problem.

Accurate representation of the physics of power flow on the network requires solving AC optimal power flow (AC-OPF), which is a non-linear optimization problem which can be challenging to solve for even a single representative hour. AC-OPF is generally too computationally demanding to incorporate into high-level transmission expansion planning analysis for large networks where potential transmission additions are first identified. Instead, AC power flow is typically represented with a linear formulation, such as the DC-OPF approximation. Candidate plans identified with the DC-OPF based TNEP optimizations are then elevated to an AC-OPF based feasibility study where reliability issues are addressed with ad hoc modifications. However, this process has been shown to produce suboptimal plans in practice because the DC-OPF error is non-negligible, especially in the stressed system states that are most relevant to transmission planners.

Finally, two-stage adaptive TNEP models can be useful, but are still limited in representing dynamic investment problems in which there are many potential decision points. Multistage stochastic models are needed to specify the timing of future investments, and the long-term events that should trigger those investments, in more detail. These details are crucial for developing an effective investment strategy and identifying the opportunities for increasing the flexibility of the investment plan.

Additional project goals are motivated by the need to address the challenges above by extending the methods to apply to co-optimization, to approximate AC power flow, and to facilitate tractable solutions to multi-stage formulations.

Existing Approaches for Stochastic Transmission Planning

Traditionally, the representative scenario set was often chosen heuristically, based on the modeler's intuition [1]. Many studies instead randomly select scenarios and give them equal weight in the TNEP model [11]. The performance of these two strategies is erratic, but quickly optimizing scenarios chosen with simple rules, or many small sets of randomly sampled scenarios, will occasionally provide a few acceptable networks that can be evaluated and refined further with other tools. Other studies have addressed volatility by employing variance-reduction techniques like Latin-hypercube sampling or repeated Monte Carlo sampling across iterations of the solution method [12-13].

Another approach is to use a scenario selection technique that only needs information about the operation parameters that vary across scenarios. Most of these techniques share the objective of minimizing the distance between the scenarios omitted from the analysis and their closest counterpart that is included. Forward scenario selection, backward scenario selection, and K-means clustering (KMC) are examples of this approach.

Grouping scenarios together based on distance minimizing criteria; however, will inevitably average extreme events into less extreme clusters so these methods often have trouble representing reliability risks with fewer than 200 scenarios [14], [15]. Some algorithms compensate for this by requiring shorter distances between extreme scenarios and their representative cluster, or by iteratively adding more clusters until the error of the model is sufficiently reduced [16], [17]. This does not fully resolve the issue because a large proportion of the sample is still devoted to scenarios that are unlikely to fail and do not provide useful additional information.

Importance sampling (IS) has been applied to address this limitation by first solving the OPF models for the full set of scenarios of potential interest for a baseline network, and then adjusting the probability of sampling each scenario and its weight in the optimization according to its contribution to the total cost of the baseline run. This ensures that the sampled scenarios are more focused on impactful situations [18], [19].

However, importance sampling has several shortcomings that we aim to address. First, importance sampling focuses on scenarios with a large objective value (i.e., high cost), but the magnitude of this value is often partially due to factors that are not affected by transmission decisions. Some scenarios will be more expensive regardless of the configuration of the transmission network. The representative power of a subset of scenarios would be improved by instead selecting the scenarios that have a large variance across the possible network configurations to ensure that factors affected by the decisions at hand are focused on. Second, implementations of importance sampling reveal that the most important scenarios often come from the same short period of the year and have similar operating conditions [19]. These scenarios all exhibit the same underlying problem and including several of them in the model is redundant. Further, if there are several failure mechanisms that occur, large samples are still required to be sure that the least expensive failure modes are sampled at least once. Third, the use of a baseline network does not account for problems that are unique to the other networks under consideration; new lines can cause new problems and importance sampling is likely to miss problems that did not occur in the baseline network. Finally, it is difficult to estimate the true cost of the system, which is necessary both to justify the investment in new assets as opposed to other risk mitigation techniques and to measure the validity of the sampling and weighting technique.

For long-term uncertainty, adaptive TNEP studies limit the number of scenarios considered, the number of candidate lines considered, and/or the size of the network modeled to be able to solve the problem [7-10]. In addition, most adaptive transmission studies have relied on simple representations of short- term uncertainty in which a small number of representative hours are selected from historical operating conditions and weighted in advance, without considering how the underlying parameters in those hours would change in each long-term scenario or for each candidate investment decision. The selection of representative hours is typically made with random sampling [11], [18], heuristically to represent known seasonal operating regimes [13], or with methods such as k-means clustering [14-15] or moment matching to the historical data. There have been some notable examples of the related generation expansion planning problem that explicitly treated both long-term and short- term uncertainties [9] although none explicitly considered the dependence of nested scenarios.

Proposed Solution

In this section, we briefly describe the main innovations in the methods developed in this project. We focus here on the main ideas, and the technical and mathematical detail is provided in the Technical Approach section.

Our approach for selecting representative hours is called the Approximate Latent Factor Algorithm (ALFA). The key idea that motivates this method is the recognition that if a small number of hours are to be used in selecting the best transmission plan, that the set of hours should be chosen to best distinguish between the alternative candidate plans. Because the focus of transmission is primarily on reliability, and only secondarily on efficiency (lower cost), the hours selected should represent all possible failure modes in the system. Existing methods such as importance sampling focus instead on selecting the hours with the highest cost or worst reliability performance. However, multiple hours with high cost may be examples of the same underlying problem, and the additional transmission that improves reliability in one of those hours may well improve in all the others as well. This approach often includes redundant hours that add no information, and may also omit hours that represent distinct, if less costly, failure modes.

The main approach within ALFA consists of first sampling a small number of feasible transmission plans, and evaluating their performance for the full short-term uncertainty set (e.g., all hours of the year from a historical dataset). The costs from the OPF solution for each hour for each sampled plan can then be organized in a matrix to which singular value decomposition (SVD) is applied. The SVD results identify the distinct failure modes or “latent factors” and indicate the degree of correlation in the costs across hours for plans and across plans for hours. ALFA then applies experimental design methods to this information to select the hours for a specified sample size that maximize the covariance across plans. In other words, ALFA chooses a set of hours for which the different transmission plans have the maximum difference in costs. Finally, a set of weights is calculated associated with the selected hours that best approximate the costs of a plan across all hours based on the cost from the selected sample of hours. These weights can then be included directly in the objective function of the reduced TNEP model.

Our approach to manage the complexity of many long-term scenarios is based on partitioning, rather than selecting a subset of representative scenarios and omitting the rest. Scenarios are still clustered together to increase efficiency. As in ALFA, the clustering is done not on the basis of similarity in parameter space, but rather groups together scenarios that make the same or very similar first stage decisions. Once all scenarios are assigned to a cluster, all are retained in the approximate model. The increase in computational efficiency is obtained by constraining the second stage decisions to be the same within a group, but still allows distinct adaptive decisions for different groups. This dramatically reduces the number of second stage binary variables, allowing for faster computation. At the same time, there are benefits to retaining all long-term scenarios in approximate model. One benefit is that the solution provides a complete decision rule for all future contingencies. This is in contrast to several currently used methods, for example Sample Average Approximation (SAA), which randomly samples a subset of scenarios; as a result, the second-stage strategies for omitted scenarios are unknown.

The third innovation in the method developed in this project is to treat long-term and short-term uncertainties as nested. Specifically, the ALFA method is extended from the one-stage setting to two-stage by applying the same technique to each cluster of long-term scenarios. The nested approach is motivated by the recognition that in different long-term scenarios, the critical hours to capture the underlying failure modes will occur at different times and will consist of distinct load patterns. The combination of clustering long-term scenarios that benefit from similar transmission investments with selecting representative hours for each group that best discriminate between transmission plans produces very good approximate solutions with reasonable computational effort.

One challenge to other methods for approximately solving two-stage adaptive transmission planning is that they do not generally provide any information about the quality of the solution. Each method of simplifying or reducing the original problem from the full uncertainty space to a smaller set of scenarios produces approximate solutions to the full problem, but there is no way to know how far they are from the true optimal solution. Another innovation in the methods developed in the project uses a combination of approximation methods to provide upper and lower bounds on the cost. The partitioning approach to long-term scenarios provides a complete feasible decision strategy. This strategy can then be fixed and simulated for all long and short-term uncertainties to obtain the true cost, which gives an upper bound to the optimal cost. A statistical confidence band for the lower bound on cost can be obtained by applying the SAA method, solving many repetitions of the problem with randomly sampled sets of scenarios. Because the SAA does not include all scenarios in each instance, it provides a lower bound on the cost. The efficiency of SAA is further improved by employing a combination of control variates and importance sampling to reduce the variance in the estimate, both of which are facilitated by re-using the sampled simulation results from the development of the partitioned model with ALFA.

Objectives

To address the challenges identified above to solving multi-stage stochastic transmission planning models with large sets of long-term and short-term uncertainties, this project pursued the following objectives:

1. Develop and demonstrate method for short-term uncertainty for one-stage stochastic TNEP and compare to existing methods;
2. Develop and demonstrate method for nested long-term and short-term uncertainty for two-stage stochastic TNEP and compare to existing methods;
3. Demonstrate method for co-optimizing generation and transmission;

4. Extend and demonstrate method for AC-OPF-based TNEP; and
5. Demonstrate the ability to solve three or more stage stochastic TNEP problems.

Expected Results

This project was intended to address the Transmission Planning objectives in the original FOA from the Dept of Energy. Specifically, this project aimed to develop a rigorous method for performing multi-period transmission planning under uncertainty, and transmission and generation joint planning under uncertainty, applied to large power systems, and with AC optimal power flow. The developed method will be computationally efficient, will produce solutions within a few percent or less of the optimal (minimum cost) plan, and will not require artificial restrictions on the problem as often assumed (reduced list of candidate lines or reduced scenarios of future generation location). This method could be used by RTOs, NERC regions, and other wide-area planning efforts, and would better represent the features of their actual planning problem than do the existing methods.

The primary expected result of the project is a dramatic reduction in the computation time required to determine the best high-voltage transmission lines to add to the existing grid in the near-term when we do not know where future generation, in particular renewable energy sources, will be located. Improved transmission planning would facilitate reduction in the cost of electricity and enable the integration of greater capacity of renewable generation. The method to be developed, which would be placed in the public domain and shared freely, would enable ISO/RTOs and other regional organizations to more effectively plan for changes in the coming decades to the power system, and enable them to consider a broader range of alternative investments and possible future scenarios than current methods.

III. Technical Approach

Project Activities

The project tasks correspond closely to the objectives identified in the previous section. Here we briefly review each task, when it was completed, and the milestones and deliverables that resulted from each:

Task 1: Develop and demonstrate method for short-term uncertainty for one-stage stochastic TNEP and compare to existing methods

Task 1 focused on addressing the short-term variability in demand and renewable generation. The ALFA method was developed in Task 1 and was tested on a one-stage transmission planning problem for the WECC network (see below for case study description). The accuracy of the method was compared to two widely used techniques, K-means clustering and importance sampling. Key outcomes (see Section IV) are quantitative estimates of the accuracy of ALFA relative to other methods for equivalent computational effort. This task resulted in one publication, Bukenberger and Webster (2019a). Task 1 was conducted primarily during Q1-Q4, although the publication of the article took additional time because of the review and editorial process at the journal.

Task 2: Develop and demonstrate method for nested long-term and short-term uncertainty for two-stage stochastic TNEP and compare to existing methods

Task 2 focused on addressing the long-term uncertainty and short-term variability with a nested approach. The Latent Uncertainty Clustering Algorithm (LUCA) method with the extension of ALFA to the multi-stage context was developed in Task 2. The LUCA-ALFA method was tested on two different two-stage transmission planning problems, both using the WECC network. This task also included the development of the bounding approach that combined LUCA-ALFA with SAA to establish upper and lower bound costs for the approximate solution. The accuracy of the method was compared with using only K-means clustering or only Sample Average Approximation. Key outcomes (see Section IV) are quantitative estimates of the accuracy from our method and of the quality of the solutions obtained, relative to other methods for equivalent computational effort. This task resulted in an article currently under review, Bukenberger and Webster (2020a). Task 2 was conducted primarily during Q5-Q12, with many iterations of continuous improvement. The method developed in Task 2 provided the foundation for the subsequent tasks, and those tasks spurred additional improvements in the core method over the performance period.

Task 3: Demonstrate method for co-optimizing generation and transmission

Task 3 focused on extending the partitioning of long-term scenarios to apply to two-stage adaptive models that choose investments in generation and transmission at the same time (the co-optimization problem). The goal of this task was to demonstrate the partitioning method would work for the co-optimization setting. We constructed an alternative two-stage investment problem with a distinct uncertainty set consisting of load growth, carbon price, renewable credit value, capital costs in wind and solar, and early retirements of existing generation. The results from the partitioning method were compared with those from a forward scenario selection approach. This task resulted in one publication, Bukenberger and Webster (2019b), and an associated conference presentation. Task 3 was conducted primarily during Q4-Q8.

Task 4: Extend and demonstrate method for AC-OPF-based TNEP

Task 4 focused on extending the ALFA method for representing short-term uncertainty to be used for transmission planning with AC power flow instead of the DC-OPF approximation. The one-stage stochastic TNEP problem from Bukenberger and Webster (2019a) was augmented with reactive power demand. Rather than simply repeat the ALFA procedure directly on an AC-OPF Model, we demonstrate superior performance from combining the challenge of AC load flow approximation and the scenario reduction problem to manage uncertainty. Specifically, the ACDC-ALFA method solves both AC-OPF and DC-OPF for the same sample plans and all short-term operating conditions, and uses the ALFA method to select the operating conditions to include in the approximate model and, importantly, the weight associated with each hour such that the approximate TNEP model can be solved used DC-OPF, but the weights approximate the solution one would have obtained with AC-OPF. Importantly, this approach enables the DC to AC correction to be adapted to each unique operating condition (i.e., each hourly pattern of demand and renewable generation at all buses). This task resulted in one paper that has been prepared for publication, Bukenberger and Webster (2020b), which will be submitted to *IEEE Transactions on Power Systems* in the near future. Task 4 was conducted primarily during Q6-Q12.

Task 5: Extend and demonstrate method for TNEP with three or more investment stages

Task 5 focused on extending the LUCA method for stochastic multistage adaptive planning (MAP) settings with three or more transmission investment stages. We developed two multi-stage case studies, one with five stages but a relatively simple set of long-term uncertainties, and a three-stage problem with a larger set of long-term scenarios representing uncertainty in more dimensions. The nested long-term and short-term scenario methodology from Bukenberger and Webster (2020a) was extended to the multistage context by adapting the notion of clustering long-term scenarios in the two-stage case to clustering subtrees and using dynamic programming to solve the successive subproblems for each stage. This task resulted in one paper that has been prepared for publication, Bukenberger and Webster (2020c), which will be submitted to *INFORMS Journal on Computing* in the near future. Task 5 was conducted primarily during Q8-Q12.

Methods

WECC Network Data and Assumptions

For all tasks and experiments in the project, we use the Western Electricity Coordinating Council (WECC) as the context for the transmission planning problems. Specifically, all models in this project use a 312-bus aggregation of the WECC network extended from the work done by the Johns Hopkins University [7], [15], [20], and using the methodology from Shi [21]. In this model there are 654 existing transmission lines and 51 candidate lines; this results in over 2.25×10^{15} possible candidate networks that must be chosen from. Our model includes 980 generation units which are aggregated from the full set of actual generation units by bus and fuel type. In our model, 438 of the generators are dispatchable and the remaining 542 are intermittent renewable generators.

Fuel prices, regional load, and renewable resource data is taken from the WECC 2024 Common Case Dataset [22]. The fuel prices vary by location and natural gas prices also vary by month. This results in 52 different fuel price profiles, 25 of which change over the year. There are 261 distinct renewable generation profiles developed from site specific data to preserve the geospatial and temporal relationships between different generation units. The hourly wind, solar, and hydroelectric data are modeled as intermittent resources with hourly profiles detailed in [22]. The hourly load shapes are extrapolated from historical data for 40 balancing areas and proportionally distributed to each bus according to the balancing areas that are represented at that bus. Most buses represent multiple balancing areas so the load profiles at these buses are proportionally constructed from the relevant load shapes as described in [20]. The system load is increased by 30% relative to the reference data. Unserved load incurs a penalty of $1,000 per MWh and curtailed renewable energy incurs a penalty of $300 per MWh.

Detailed assumptions and methods for each specific task are described in the associated publications: Bukenberger and Webster (2019a, 2019b, 2020a, 2020b, 2020c).

One-Stage TNEP Experimental Design and Assumptions

ALFA is designed to select the set of hours where the impact of transmission investments is most apparent while avoiding redundant hours that represent the same underlying problem. Beyond identifying individual hours, the cost-covariance between different hours can be used to accurately approximate the costs in unobserved hours. ALFA should therefore provide an accurate measure of the total expected cost of an expansion plan with fewer representative hours, which in turn should result in better transmission recommendations when compared to models developed with other uncertainty representation techniques.

The performance of ALFA in terms of accuracy and in solution quality will be compared to k-means clustering, which is the preferred method for transmission planning, and importance sampling, which is a popular method for transmission planning and the broader stochastic optimization community.

The validity and usefulness of this hypothesis is tested on 52 weeks of hourly data, a total of 8,736 hours, taken from the WECC common cases data set. The experiment has variable demand, renewable generation, and natural gas prices.

Two-Stage TNEP Experimental Design and Assumptions

LUCA, the covariance-based scenario partitioning method, should be compared to both scenario reduction methods and other partitioning methods. The performance of these methods with and without the added consideration of short-term uncertainty is also relevant. Finally, LUCA should be tested both with and without generation co-optimization, since there are advocates for both paradigms in multistage TNEP.

Partitioning methods may outperform scenario reduction methods by giving better resolution into the short-term events of the full scenario set while reducing the number of binary transmission planning variables. For example, in generation co-optimization studies, the partitioning approach will give better resolution into the decisions of generation planners in every scenario, therefore leading to better first-stage investments from both generation and transmission planners. In models with extensive short-term uncertainty, partitioned models will be able to represent the important extreme events from many scenarios within a single block whereas scenario reduction methods will need to identify both the important scenarios and the important extreme events from within those scenarios. However, grouping long-term scenarios into blocks, as opposed to selecting representative scenarios, artificially reduces the adaptability of the modeled decision process, which may degrade the solution overall.

Several experiments are investigated with a two-stage model. First, comparisons between all methods on the quality of first stage decisions both with and without short term uncertainty. Next partitioning methods provide a complete investment policy with a single optimization, whereas scenario reduction methods only provide the first stage investment decisions, so both partitioning and scenario reduction methods will be compared on the quality of their first stage decisions, and partitioning methods will also be compared on the basis of their full decision policy (first and second stage solutions). Finally, generation co-optimization will be performed on a realistic data set to compare partitioning methods.

Three different test cases are developed to conduct these tests.

A small version of the WECC problem is developed with only the 10 candidate lines that are frequently found to be beneficial. This model has 20 scenarios of demand growth as the only dimension of long-term uncertainty. A summer and a winter day are used to give 48 hours of short-term uncertainty. The benefit of this model is that, with just 10 candidate lines, there are 1,024 possible expansion plans; this is few enough that the cost of every expansion plan can be calculated exactly and stored. Then, not only can the problem be solved exactly with dynamic programming, but approximating models can be generated and solved very quickly with the same data.

A larger and more realistic model uses the full set of 51 candidate transmission lines and the full year of hourly short-term data. Additionally, several dimensions of long-term uncertainty are included in this case study. The sources of long-term uncertainty include load growth, RPS levels, water availability for hydropower, generation expansion, generation retirements, and regulation to mitigate the impact of a natural gas pipeline disruption as described in [23]. The long-term uncertainty scenarios are summarized in Table 1. This model highlights the difficulty of realistically representing long-term uncertainty; the long-term scenario data is an aggregation of much more complex short-term behavior. Simple scenario reduction methods, and methods for selecting hours from within those scenarios, are likely to show degraded performance on the more realistic uncertainty sets.

A generation co-optimization model uses the first 15 scenarios, again with demand growth, hydro availability, generation retirements, natural gas disruptions, and RPS levels, but now 759 generating units are treated as variables rather than uncertain scenario realizations. Rather than a full year of short-term data, only two hours from the year are used on this model so the effects of long-term uncertainty can be separated from short-term uncertainty and LUCA can be compared directly to other partitioning methods on a multi-dimensional long-term uncertainty set.

Table 1: Scenario Assumptions for Large Two-Stage TNEP Experiment

State/Scenario	Demand Growth	Hydro Availability	Rapid Retirements	Natural Gas Disruption	RPS penalty	New NG	New Solar	New Wind
1	20%	High	Y	Y	$100	0	0	0
2	20%	High	Y	Y	$900	0	0	0
3	10%	Low	Y	Y	$300	0	0	0
4	20%	Low	Y	Y	$100	9614	7484	2958
5	20%	Low	Y	Y	$900	0	0	0
6	10%	High	Y	Y	$300	9614	7484	2958
7	20%	Low	N	N	$900	9614	7484	2958
8	20%	Low	N	Y	$100	0	0	0
9	0%	Low	Y	Y	$900	9614	7484	2958
10	20%	High	N	N	$900	0	0	0
11	10%	Low	Y	N	$300	5864	8490	1935
12	20%	High	Y	N	$100	9614	7484	2958
13	10%	Low	Y	N	$900	0	0	0
14	10%	High	Y	N	$900	0	0	0
15	10%	High	N	N	$100	7559	5783	1089
16	10%	Low	N	Y	$300	5864	8490	1935
17	10%	Low	N	Y	$900	5864	8490	1935
18	20%	Low	Y	N	$900	0	0	0
19	20%	Low	N	Y	$100	9614	7484	2958
20	10%	Low	N	Y	$100	9614	7484	2958
21	20%	High	N	Y	$300	9614	7484	2958
22	0%	High	N	Y	$100	0	0	0
23	10%	High	N	Y	$300	0	0	0
24	0%	Low	N	N	$300	0	0	0

One-Stage AC-OPF TNEP Experimental Design and Assumptions

By simulating both the AC-OPF and the linearized DC-OPF for the same hours and transmission topologies, the error in the DC-OPF can be characterized and corrected in the transmission optimization. This can improve the transmission optimization in two ways: first, the objective function will more accurately reflect the AC operation cost of the topology, so the DC optimization model should be more capable of finding lines that are marginally not worth their cost in the DC system but in fact are worthwhile in the AC system. Second, the error characterization will correlate certain DC behavior with AC reliability penalties, even when the DC costs are low; the resulting DC-based optimization can then maintain AC feasibility and reduce reliability penalties by avoiding the DC behavior that is indicative of AC reliability failures.

The ACDC extension to ALFA will be compared to ALFA applied to the DC simplification of the problem and to ALFA applied to the AC problem. We would expect the AC-based TNEP model to provide the best overall solutions because the most accurate OPF is used; however, the resulting optimization is a MINLP and is much more difficult to solve than the MILP models from the other two approaches.

We test these three approaches with 8,760 hours of operating conditions on the WECC network. We compare the methods based on the accuracy and quality of the final solutions when simulated with AC power flows.

Multi-stage TNEP Experimental Design and Assumptions

Multistage adaptive planning models allow for the timing of investments, and the sequence of long-term events that might trigger investments, to be represented in more detail. Realistic multistage TNEP models have not been successfully demonstrated with other methods, so comparisons will not be possible for large-scale multistage tests. Of course, finding a high quality TNEP investment strategy for a problem with thousands of scenarios and unique short-term events is itself a significant result.

Comparisons between LUCA-MAP, sampling-based scenario reduction methods, and other partitioning techniques will be performed on a smaller WECC model that can be solved exactly with dynamic programming. All methods will be compared on the quality of the first stage investments. Scenario reduction approaches do not generally provide a method for recovering the full investment policy, so alternative partitioning methods will be used here to evaluate the quality of the complete investment policy of LUCA-MAP for the full investment horizon.

The large multi-stage experiment uses the full set of 51 candidate lines and considers 2,025 long-term scenarios with uncertainty from load-growth, RPS level, EV adoption, and several policies that determine the charging behavior of EV. Four future EV possibilities are considered: negligible EV adoption, EV adoption with peak charging aligned with the existing demand peak, EV adoption with charging delayed 6 hours, and EV adoption with charging moved forward 6 hours to align better with solar generation. States without EV adoption have equal probability of transitioning to any other EV state; states with either delayed charging or early charging remain in that EV charge state; states with EV charging aligned with peak demand can transition to the three EV charge states but cannot backtrack to the state without significant EV adoption. Short-term uncertainty is represented with 4 weeks of hourly data, one week from each season to capture the seasonal operation patterns, this gives a total of 672 hours in each node of the decision tree.

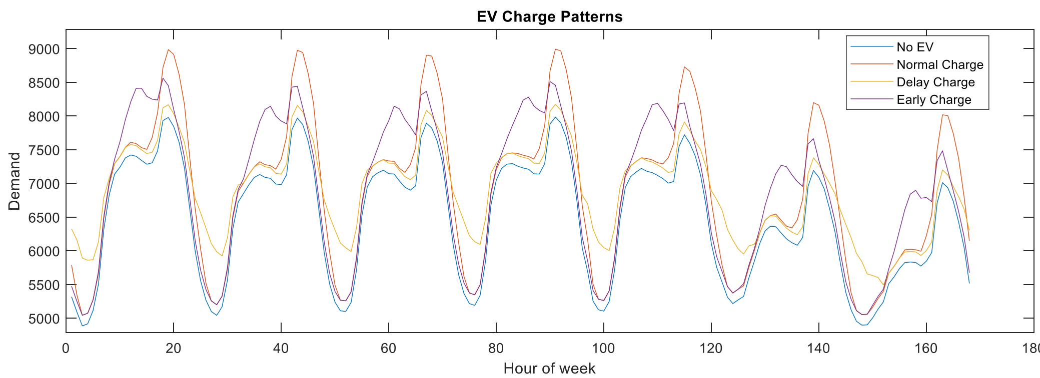

Figure 1: Short-term electricity demand patterns in long-term scenarios with differing EV adoption and charge behavior.

The small WECC model with 10 candidate lines is extended to a multistage model with using the same data as in the two-stage case. A total of 5 decision stages are considered with a branching factor of 5 in each stage. The model has 3,125 total scenarios.

Table 2: ALFA Algorithm

Step	Approximated Latent Factor Algorithm (ALFA)
1	Take sample of candidate transmission topologies
2	Calculate OPF costs for each topology and each short-term event. Store values in data table Z
3	Center the OPF cost data around the mean of each row (hour)
4	Decompose the mean centered OPF cost data with SVD
5	Select D-optimal subset of rows (hours) with row exchange algorithm, i.e. candexch, detmax, Fedorov exchange algorithm
6	Calculate weights for each selected hour with OLS regression equations; alternatively, calculate strictly positive weights with minimum mean squared error
7	Form approximating optimization problem with selected hours and appropriate weights

Table 3: Nested LUCA/ALFA Algorithm

Step	Latent Uncertainty Clustering Algorithm (LUCA)
1	Take sample of candidate transmission topologies
2	Calculate expected annual costs (operation and fixed costs) for each topology in each scenario by solving the OPF cost for every short-term event in that scenario. Store in data table Z
3	Filter for delta-optimal plans in each scenario so the best sampled topology has a score of delta and any sub-optimal plans have a floor of zero.
4	Calculate Q, the covariance matrix (or alternatively the correlation matrix) of delta scores between each pair of scenarios
5	Cluster scenarios into blocks of related scenarios (blocks of scenarios that benefit from the same transmission expansion plans) with spectra of Q.
6	Combine short-term OPF cost data for each scenario block; approximate the short-term costs for each block with ALFA.
7	Form approximating optimization problem with expansion decisions modeled for each scenario block and short-term costs modeled approximated with ALFA.

Table 4: LUCA Algorithm for Multistage Adaptive Planning

Step	LUCA for Multistage Adaptive Planning (LUCA-MAP)
1	Generate scenario tree with nodes and unique information states
2	Take sample of candidate transmission topologies
3	Calculate expected annual costs (operation and fixed costs) for each topology in each information state by solving the OPF cost for every short-term event in that scenario. Store in data table Z
4	Solve the dynamic program over full scenario tree with the sample of transmission topologies. Store optimal expected value of state action pair at stage t in matrix V
5	From the root block, gather descendant nodes in stage 2 and group them into node blocks with the LUCA algorithm (steps 3 to 5) using the dynamic program result V as the matrix of costs. Step forward into stage t = 2.
6	For each block in stage t; gather the nodes in stage t+1 that are descended from the current block and group them with the LUCA algorithm using Vt+1 as the matrix of costs. Step forward into stage t+1 and repeat until every node in the scenario tree is assigned to a block.
7	Combine short-term OPF cost data for each node block; approximate the short-term costs for each block with ALFA.
8	Form approximating optimization problem with expansion decisions modeled for each node block and selection of representative hours.

IV. Accomplishments and Conclusions

Major Activities Completed

All tasks as described in Section III and as outlined in the original proposal have been completed. We have developed the methods for approximating short-term uncertainty for transmission planning, nested long-term and short-uncertainty approximation, and extended the algorithms to apply to generation and transmission co-optimization, AC power flow-based transmission planning, and multi-stage planning problems. Each of these methods were tested extensively on transmission planning problems for the WECC network with one, two, three, and five transmission investment stages, depending on the method being tested.

The methods have been tested and validated in several ways. Two of the papers have been published in peer-review outlets, one is under review, and two more will be submitted shortly. In each experiment, solutions from our method are rigorously compared to one or more state-of-the-art solution methods, both to validate the solution obtained as well as to establish the performance metrics described below. Finally, all experiments have been presented in numerous fora to experts in the field in academic conferences, seminars, and presentations to industry.

Five papers have resulted from this project, as planned in the original SOPO. Two have already been published in peer-reviewed journals (Bukenberger and Webster, 2019a;2019b), one is currently under review (Bukenberger and Webster, 2020a), and the final two papers are ready for submission (Bukenberger and Webster, 2020b;2020c). The complete references are provided in Appendix A.

The methods developed in this project have been presented to several industry, government, and academic organizations. Specifically, we have given seminars or presentations and explored applications of these methods in visits to PJM Interconnection, New York ISO, a webinar to U.S. Dept of Energy, Office of Electricity, presentation at two INFORMS annual conferences, the CIGRE annual meeting in 2020, and a seminar at Johns Hopkins University. Discussions continue with several collaborators for future application and extensions to these methods, including with PJM, NYISO, EPRI, and NREL.

Project Goals Achieved

All project goals and objectives identified have been achieved:

1. Developed and demonstrated method for short-term uncertainty for one-stage stochastic TNEP and compare to existing methods;

2. Developed and demonstrated method for nested long-term and short-term uncertainty for two-stage stochastic TNEP and compare to existing methods;

3. Demonstrated method for co-optimizing generation and transmission;

4. Extended and demonstrated method for AC-OPF-based TNEP; and

5. Demonstrated the ability to solve three or more stage stochastic TNEP problems.

The detailed outcomes and metrics are provided in the next section.

Noteworthy Results

Performance of ALFA on 1-stage

ALFA is significantly more accurate than approximations based on importance sampling and K-means clustering. Figure 2 compares performance across these methods for one key metric: the error in the estimated expected cost of out-of-sample transmission plans relative to the true expected cost. The accuracy translates into better solutions than the other two methods, especially with very small sample sizes. A second key metric is the expected cost of the recommended best transmission investment plan relative to the cost of the true optimal investment plan; Figure 3 compares this metric from ALFA to other methods. ALFA also finds more of the optimal lines and identifies the most critical lines more frequently than the other methods. When compared to importance sampling ALFA finds fewer non-optimal lines. K-means tends to recommend plans with fewer new transmission lines, and therefore it frequently misses important lines but includes fewer non-optimal lines. Key outcomes for the new lines identified by each method for the one-stage case study are provided in Table 5.

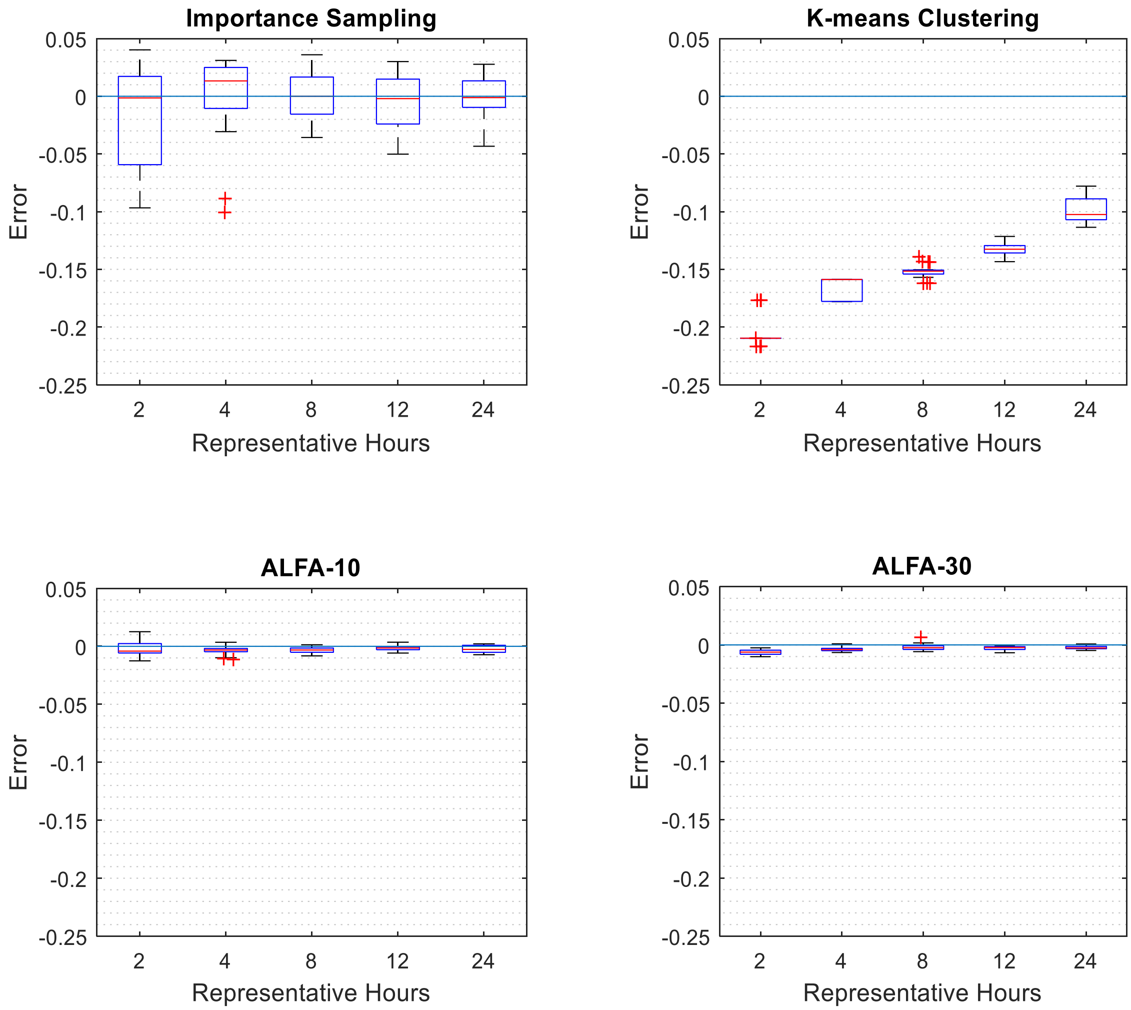

Figure 2: Error in the estimated expected total cost of transmission plans relative to the true expected cost of those plans, based on estimated costs from approximate models with representative hours chosen by Importance Sampling, K-Means Clustering, and two variants of ALFA.

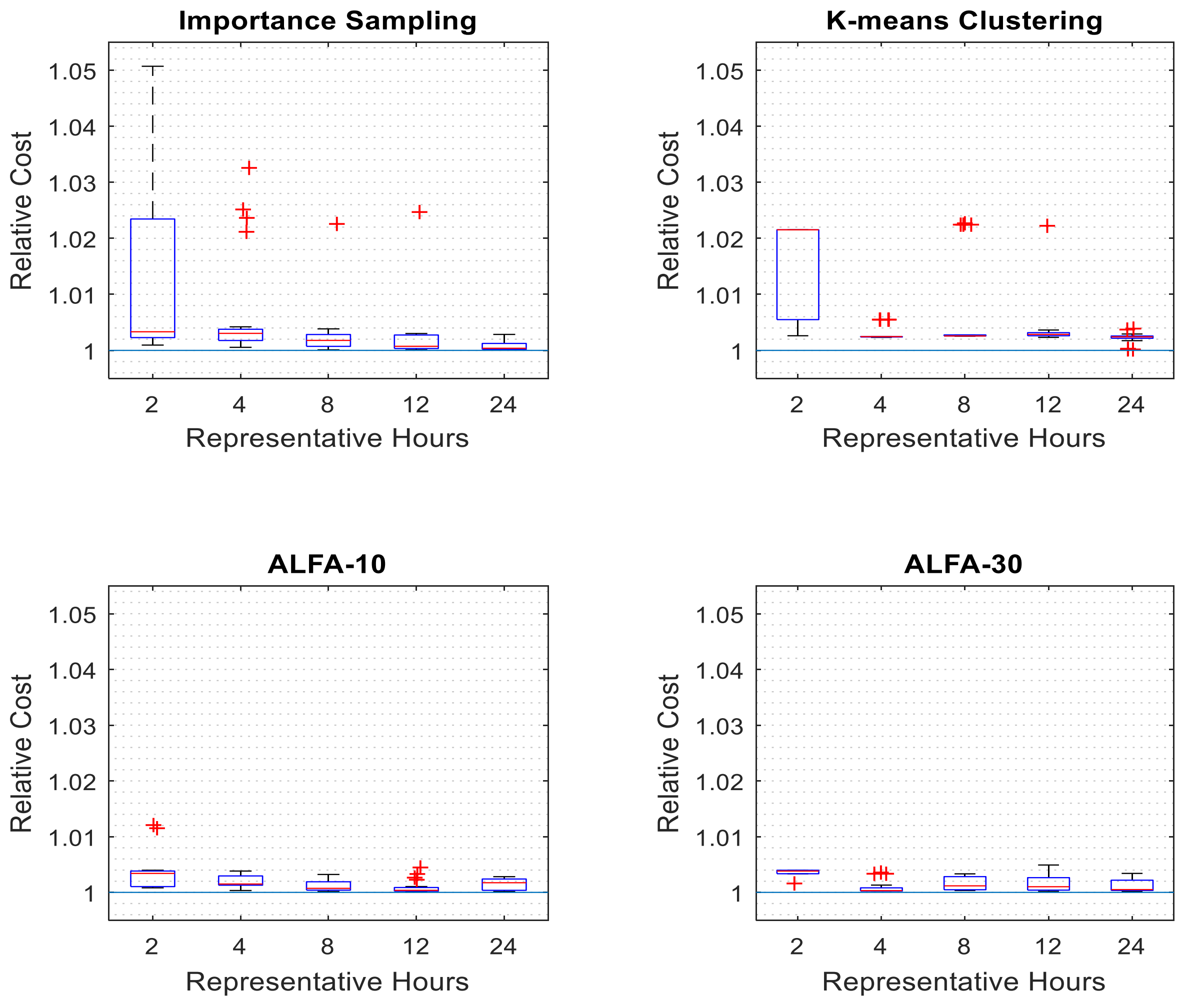

Figure 3: Relative expected cost of the recommended best transmission investment plan from chosen by Importance Sampling, K-Means, and ALFA, compared to the cost of the independently solved true optimal transmission plan.

Table 5: Frequency of Identifying Critical Transmission Lines in the Recommended Optimal Plan from each Method

Method	Sample Size	Freq. of Including Critical Line					Opt. lines	Nonopt.
		1	29	43	19	13		
ALFA-10	2	1	1	0.9	1	0.95	0.63	0.13
	4	1	1	1	1	1	0.71	0.15
	8	1	1	1	1	1	0.81	0.14
	12	1	1	1	1	1	0.82	0.10
	24	1	1	1	1	1	0.81	0.12
ALFA-30	2	1	1	1	1	1	0.67	0.17
	4	1	1	1	1	1	0.78	0.11
	8	1	1	1	1	1	0.76	0.14
	12	1	1	1	1	1	0.81	0.14
	24	1	1	1	1	1	0.85	0.15
Imp. Samp.	2	0.95	0.65	0.9	0.75	0.85	0.57	0.14
	4	0.95	0.8	1	0.95	0.55	0.65	0.13
	8	1	0.95	1	0.9	0.8	0.70	0.12
	12	1	0.95	1	1	0.85	0.76	0.14
	24	1	1	1	0.95	0.95	0.80	0.15
K-means	2	1	0.3	1	1	0	0.54	0.04
	4	1	1	1	1	0.8	0.68	0.06
	8	1	0.8	1	1	0.9	0.69	0.09
	12	1	0.95	1	1	0.9	0.67	0.11
	24	1	1	1	0.9	0.85	0.67	0.10

Performance of LUCA/ALFA on 2-stage

For two-stage TNEP problems, a key performance metric is the excess first-stage cost of the recommended investment plan from an approximation method relative to the true cost of the optimal plan.

Without considering short-term uncertainty, scenario reduction methods make worse first stage decisions when compared to partitioning with either LUCA or the Forward Scenario Selection (FSS) algorithm. Forming random partitions does the worst of all the methods, and the CVIS variance reduction method is an improvement over simple random sampling. Both LUCA and FSS always find the optimal first stage decision when used to form two or more partitions of long-term uncertainty. Figure 4 compares the excess cost metric for several methods of partitioning or sampling long-term scenarios of various sizes.

When short-term uncertainty is included in the model, the LUCA-ALFA combination makes the best first stage decisions, especially with small samples of short-term hours. The FSS and K-means combination method has inconsistent performance. All methods make substantially more costly first stage decisions without perfect short-term information. Figure 5 compares the excess cost metric for several methods for selecting representative hours to approximate short-term uncertainty.

A second key metric for two-stage problems is the excess expected cost over both stages from the recommended plan relative to the expected cost of the true optimal plan. Figure 6 compares the excess total expected costs for both stages across LUCA-ALFA and several alternatives. The full decision policy (first and second stage decisions) found with LUCA are better than those found with other partitioning method. While both LUCA and K-means partitioning methods often find the optimal first stage decision, the second stage decisions found with LUCA are better than the other partitioning methods. LUCA should have better performance on multistage problems for this reason.

A third key metric for two-stage TNEP problems is the optimality gap identified, the difference between the best upper and lower bounds on the true cost. The smaller this gap, the more confidence in the recommended investment plan from the approximate model. Figure 7 shows the optimality gap size from four variants of the sample average approximation method to find the lower bound on the optimal cost: pure random sampling, using importance sampling to reduce the variance, using control variates to reduce the variance, and using the CVIS method developed in this project which combines importance sampling and control variates. Although scenario reduction via sampling has poor performance overall, it provides good bounds. The bounds can be improved by about one order of magnitude with control variates or importance sampling. The CVIS method improves the bounds by about two orders of magnitude. Because CVIS uses the same data as LUCA, these two methods work well in conjunction. For finding the upper bounds on the optimal cost, the LUCA-ALFA method's advantage scales to larger problem where it outperforms the FSS and K-means method (Figure 8). While the LUCA-ALFA method and the FSS and k-means combination have similar performance on the small test case, LUCA-ALFA outperforms on the larger test case, indicating that the performance of the distance-based method FSS degrades when faced with more complex long-term uncertainty from a more realistic problem.

Finally, a key outcome is the demonstration of the value of these methods for long-term transmission planning under uncertainty. LUCA-ALFA models tend to postpone investments until they are needed later, when compared to the best solutions from FSS-KMC. But some alternative strategic investments are also made. Illustrative examples from the two-stage WECC case study are shown in Figure 9. Numerical results are provided in Tables 6 and 7.

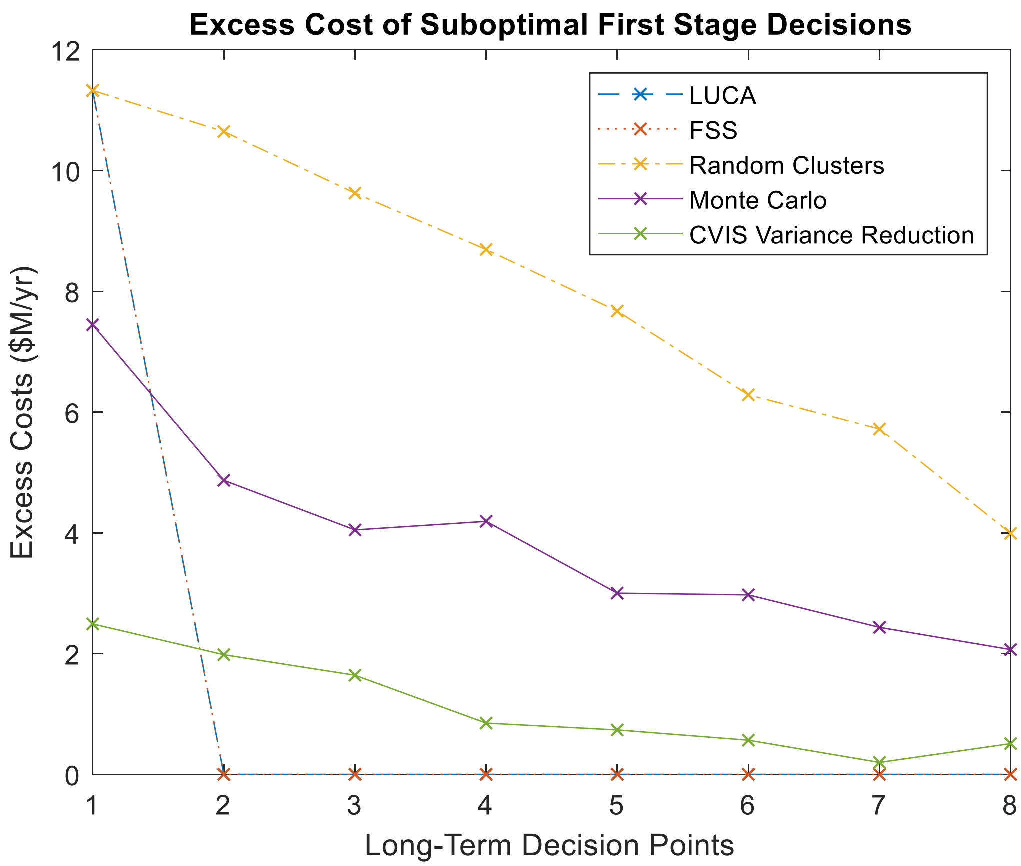

Figure 4: Excess first-stage costs above the cost of the true optimal plan for the two-stage TNEP case study. The excess cost of the recommended plan from LUCA (developed in this project), forward scenario sampling, randomly sampled clusters, randomly sampled scenarios, and scenarios sampled with variance reduction. The x-axis shows the number of clusters of long-term scenarios formed with each method.

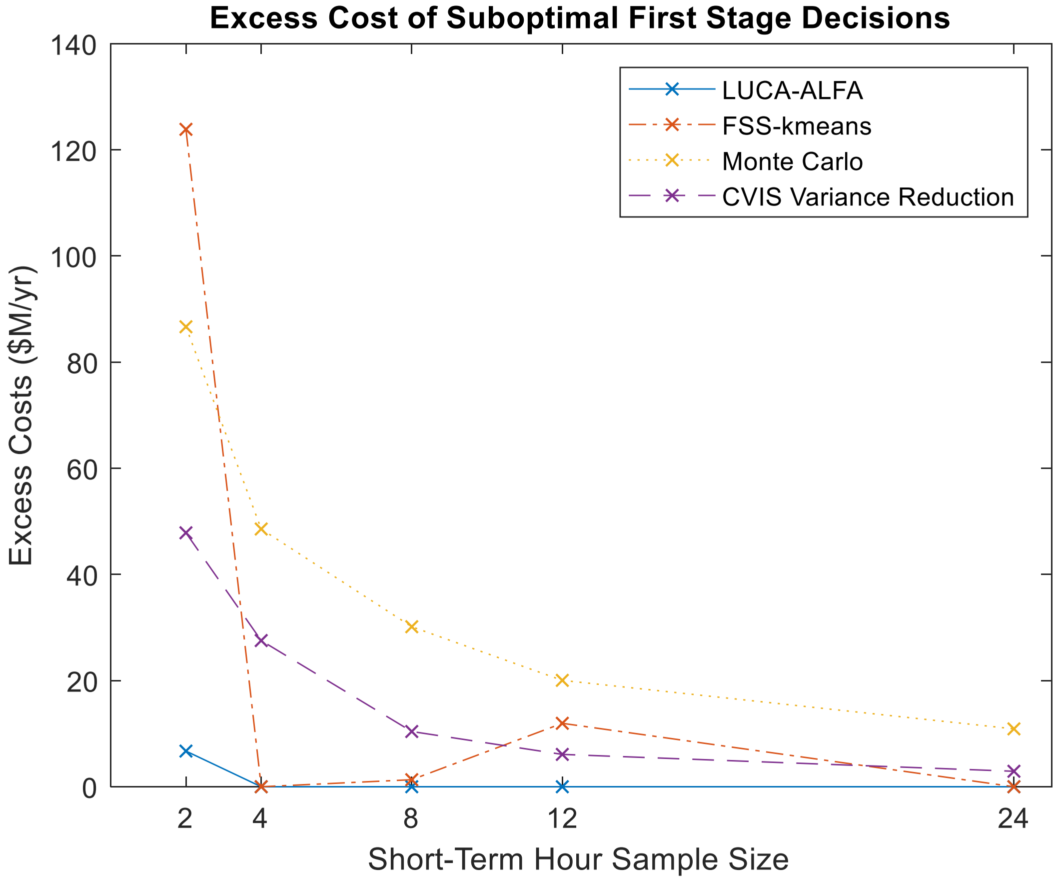

Figure 5: Excess first-stage costs above the cost of the true optimal plan for the two-stage TNEP case study. The excess cost of the recommended plan from approximate models with short-term uncertainty represented by LUCA-ALFA (developed in this project), forward scenario sampling with K-means clustering, randomly sampled scenarios, and scenarios sampled with variance reduction. Two long-term decision points are used for this example, and the x-axis shows the number of hours included.

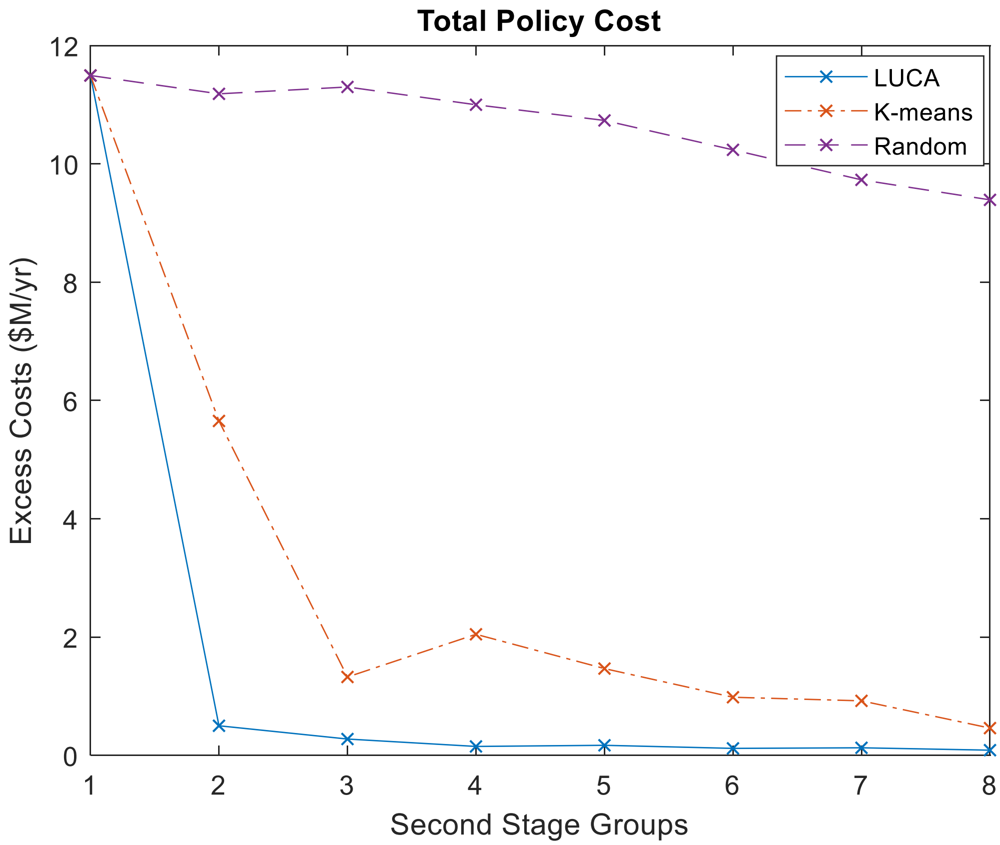

Figure 6: Excess total expected costs (first and second stage) above the total cost of the true optimal plan for the two-stage TNEP case study. The x-axis shows the number of second-stage groups formed by each method.

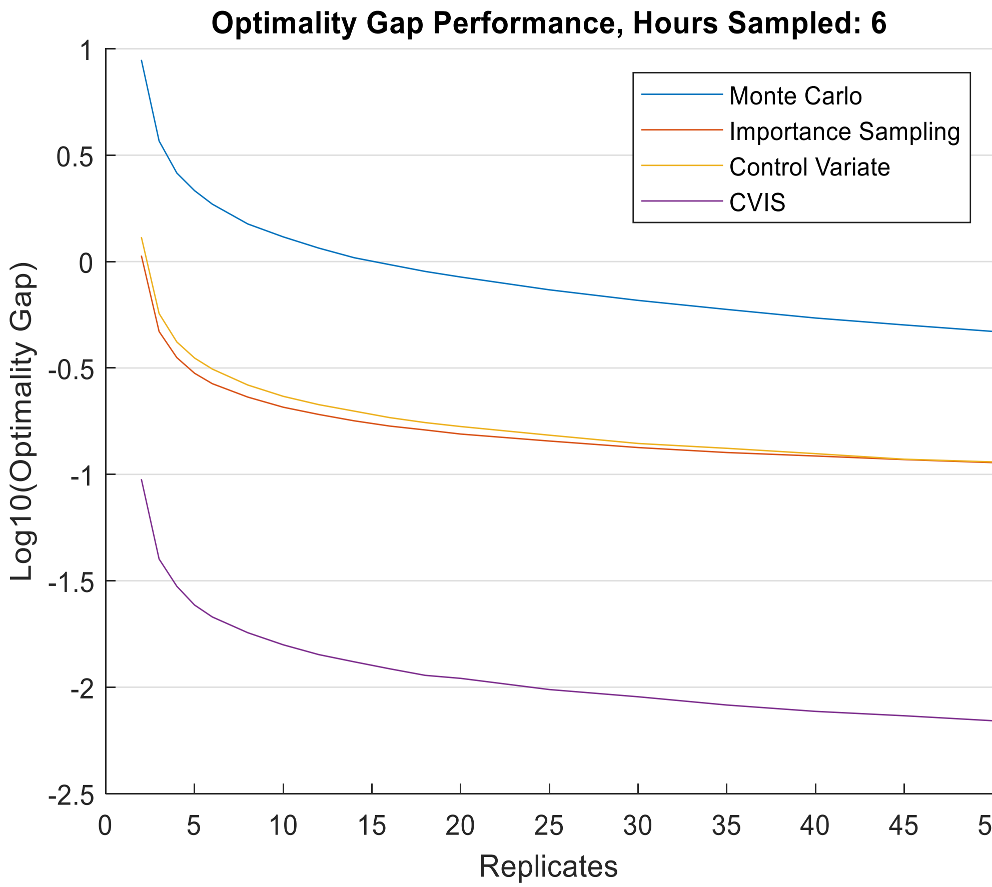

Figure 7: Log of the optimality gap (distance between upper and lower bound estimate of optimal cost) for four different variants of SAA to estimate the lower bound. The x-axis shows the number of replicants of the sampled approximate models for lower bound estimation.

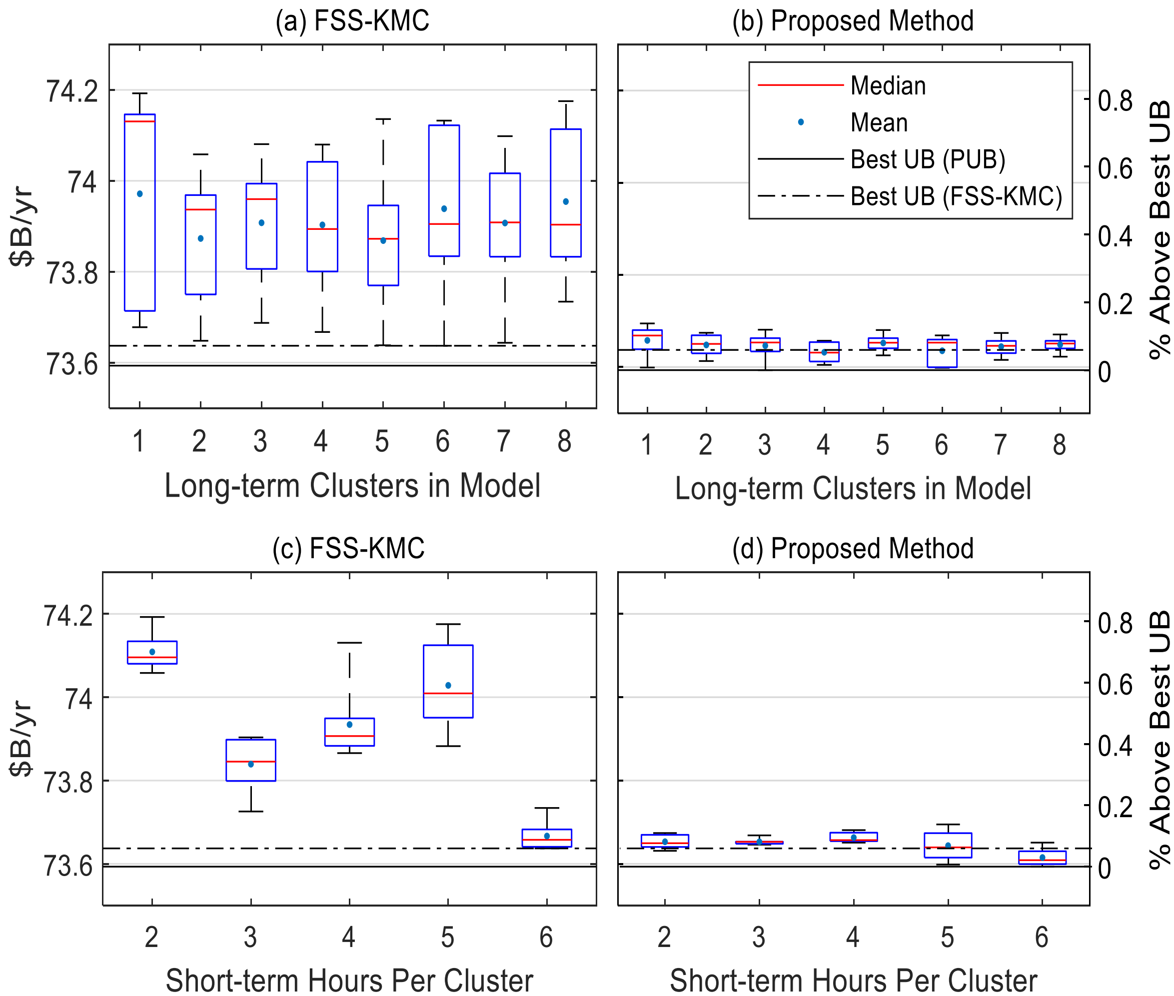

Figure 8: The estimated upper bounds on optimal cost from Forward Scenario Selection – K-Means Clustering (a) and (c) and from the LUCA-ALFA method developed in this project (b) and (d). Both absolute magnitude of cost and the excess cost above the best upper bounds estimate from any method are both shown. Panels (a) and (b) show results for different numbers of long-term scenario clusters, and panels (c) and (d) shown results for different numbers of hours included in approximate models. Dashed horizontal line in all panels indicates the best found upper bound as a reference.

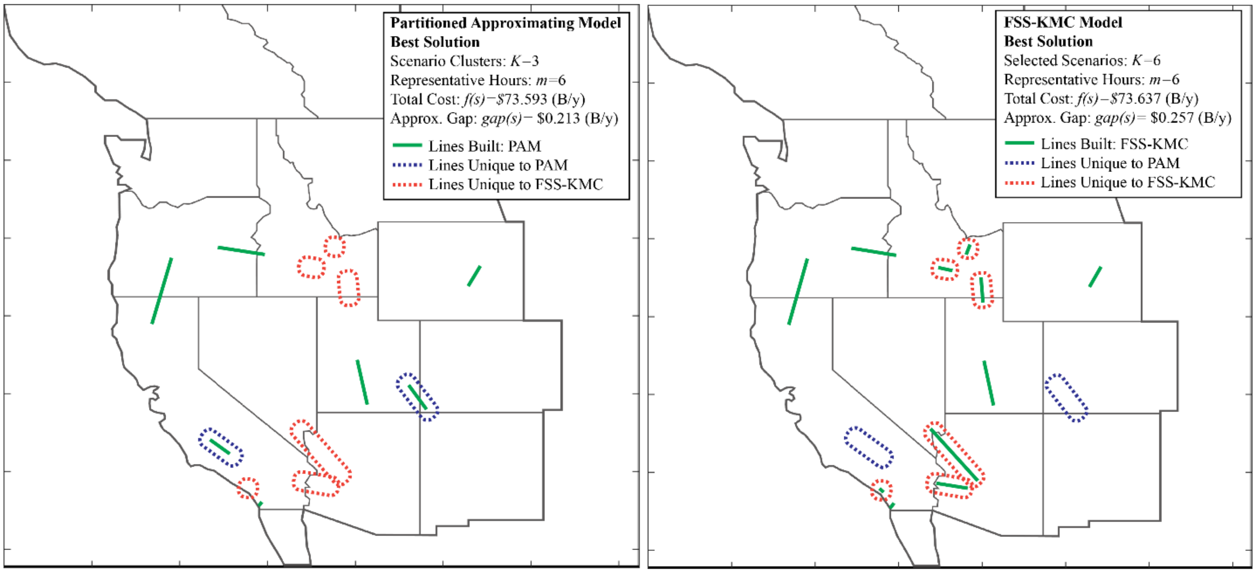

Figure 9: First stage investments for the large case study from LUCA-ALFA (left) and FSS-KMC (right). Recommended new lines from each method shown in green. Blue circles indicate new lines recommended in plans from LUCA-ALFA but not FSS-KMC, and red circles indicate new lines recommended by FSS-KMC but not LUCA-ALFA.

Table 6: Upper Bound Results for Two-Stage Experiment

		Partitioned Approximating Model				FFS-KMC Model			
n	m	$\hat{v}^{\theta*}_{Km}$	$f(\hat{s})$	Approx. Error %	$\widehat{\text{gap}}(\hat{s})$	$\hat{v}^{*}_{Km}$	$f(\hat{s})$	Approx. Error %	$\widehat{\text{gap}}(\hat{s})$
3	2	73.625	73.647	−0.030	0.267	69.528	74.081	−6.146	0.701
	3	73.647	73.653	−0.009	0.273	70.474	73.846	−4.567	0.466
	4	73.648	73.681	−0.045	0.301	71.126	73.960	−3.831	0.580
	5	73.679	73.657	0.030	0.276	70.890	73.965	−4.157	0.585
	6	73.680	73.593	0.117	0.213	71.414	73.687	−3.086	0.307
6	2	73.636	73.669	−0.045	0.289	68.445	74.133	−7.672	0.752
	3	73.651	73.653	−0.002	0.273	69.343	73.900	−6.167	0.520
	4	73.645	73.657	−0.016	0.277	69.890	73.905	−5.433	0.525
	5	73.680	73.598	0.110	0.218	69.802	74.119	−5.824	0.739
	6	73.676	73.600	0.103	0.220	70.024	73.637	−4.907	0.257

Table 7: Lower Bound Results for Two-Stage Experiment

		Pure Monte Carlo				CVIS			
n	m	$\overline{v}^J_{nm}$	$\hat{\sigma}$	$100\% \times \frac{\hat{\sigma}}{\mu}$	L^J_{nm}	$\hat{\mu}^{q,z}$	$\hat{\sigma}$	$100\% \times \frac{\hat{\sigma}}{\mu}$	L^J_{nm}
3	2	74.790	4.670	6.250	72.080	73.270	0.100	0.130	73.220
	3	72.180	5.240	7.260	69.140	73.250	0.170	0.230	73.150
	4	74.030	4.760	6.430	71.270	73.370	0.120	0.160	73.310
	5	70.730	3.660	5.180	68.600	73.160	0.380	0.520	72.940
	6	71.960	3.930	5.460	69.680	73.360	0.250	0.340	73.220
6	2	72.230	5.530	7.660	69.030	73.290	0.220	0.300	73.160
	3	74.430	4.390	5.900	71.890	73.260	0.360	0.490	73.050
	4	71.920	2.580	3.580	70.430	73.030	0.260	0.360	72.870
	5	73.430	3.160	4.310	71.600	73.220	0.330	0.450	73.030
	6	74.460	4.510	6.050	71.850	72.980	0.340	0.470	72.780

Performance of LUCA for Co-optimization partitioning vs simple partitioning

With generation co-optimization, LUCA with perfect short-term information has better performance than distance based partitioning methods. The difference between LUCA and FSS is larger here than in the small case studies, which demonstrates how simple clustering heuristics (FSS, K-means) degrade with more complex scenario spaces (Figure 10). This improvement shows fewer line investments and fewer generation investments indicating that the improvement comes from a more efficient use of capital as opposed to making more investments that are ultimately valuable (Figure 11). The first stage decisions for capacity investments show similar generation investments for the two methods while fewer line investments are made with LUCA than with FSS. Differences in first-stage investments in transmission and generation between LUCA and FSS are illustrated in Figure 12.

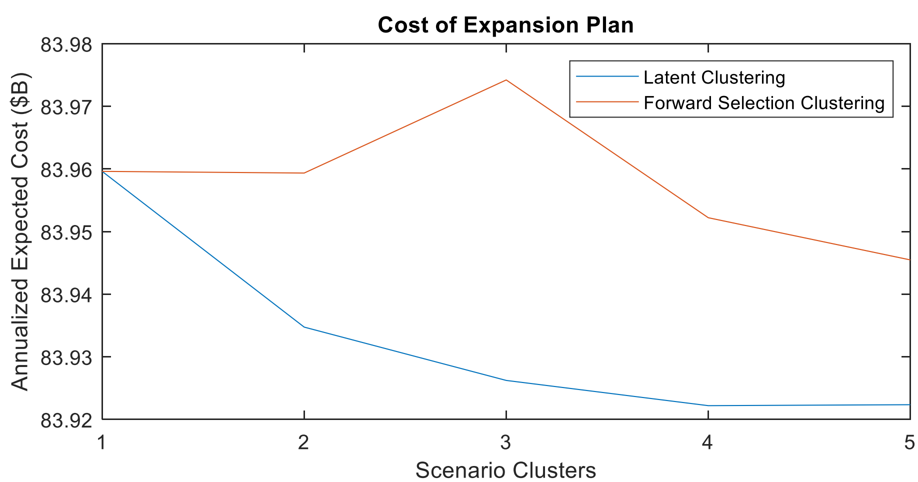

Figure 10: Total expected cost of recommended plan for generation and transmission co-optimization for two methods of clustering long-term scenarios in a two-stage problem.

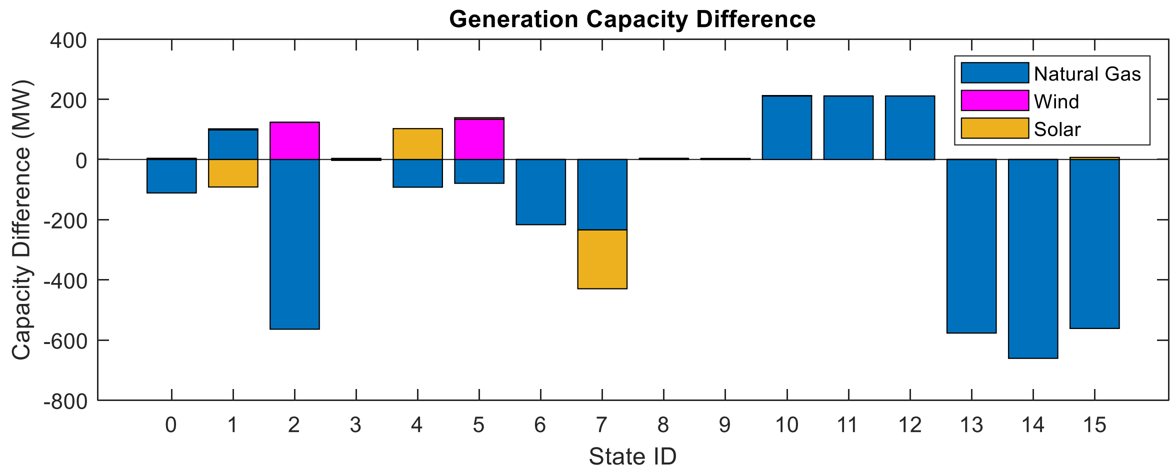

Figure 11: Difference in new generation capacity built in Stage 2, shown by generation type and realized long-term scenario. Difference shown is the new capacity recommended in the approximate solution from LUCA minus the new capacity recommended by the FSS method.

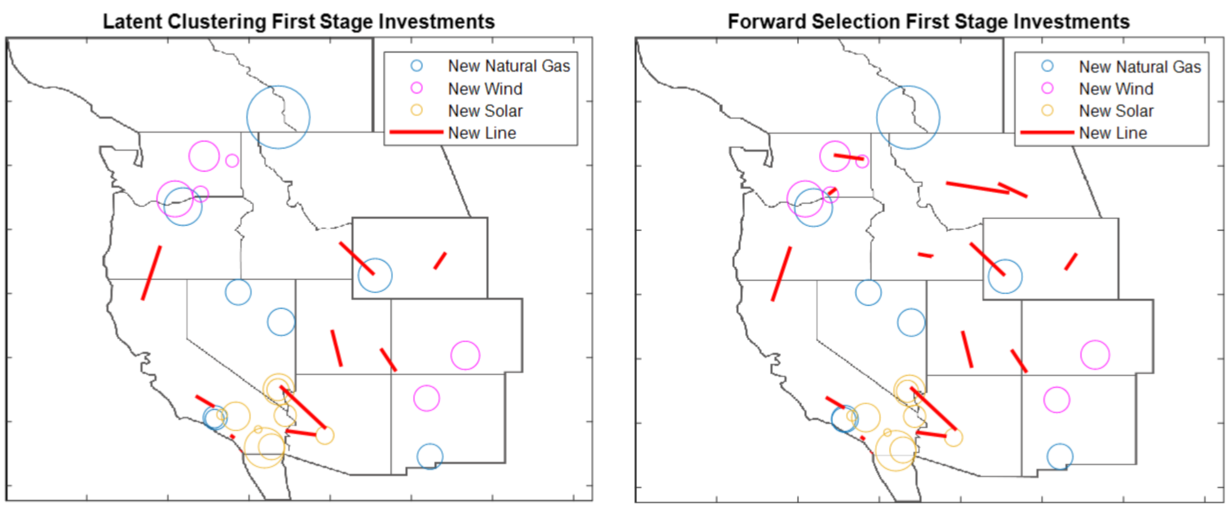

Figure 12: First stage generation and transmission investments with LUCA (left) and FSS (right) with perfect short-term information. Bubble sizes indicate the quantity of new generation capacity installed at a location.

Performance of AC-OPF ALFA

The metrics for one-stage problems, error (%) of the estimated cost of plans from the approximation models relative to the true cost and the total expected cost, are used to compare the performance of alternative methods to choose transmission lines considering AC power flow. A key outcome of this work is to show that ACDC-OPF approximation is as accurate as fully AC-Approximation, and more so than the DC-approximation. Figures 13 and 14 show the error in approximate investment plans and the expected cost of plans from each method, respectively. Because AC-TNEP is often not feasible with generic solvers, and DC-OPF based TNEP models are sometimes infeasible when validated with AC-OPF, a significant advantage to the ACDC method is that it is solvable but stills better produces investment plans that are feasible for AC-OPF.

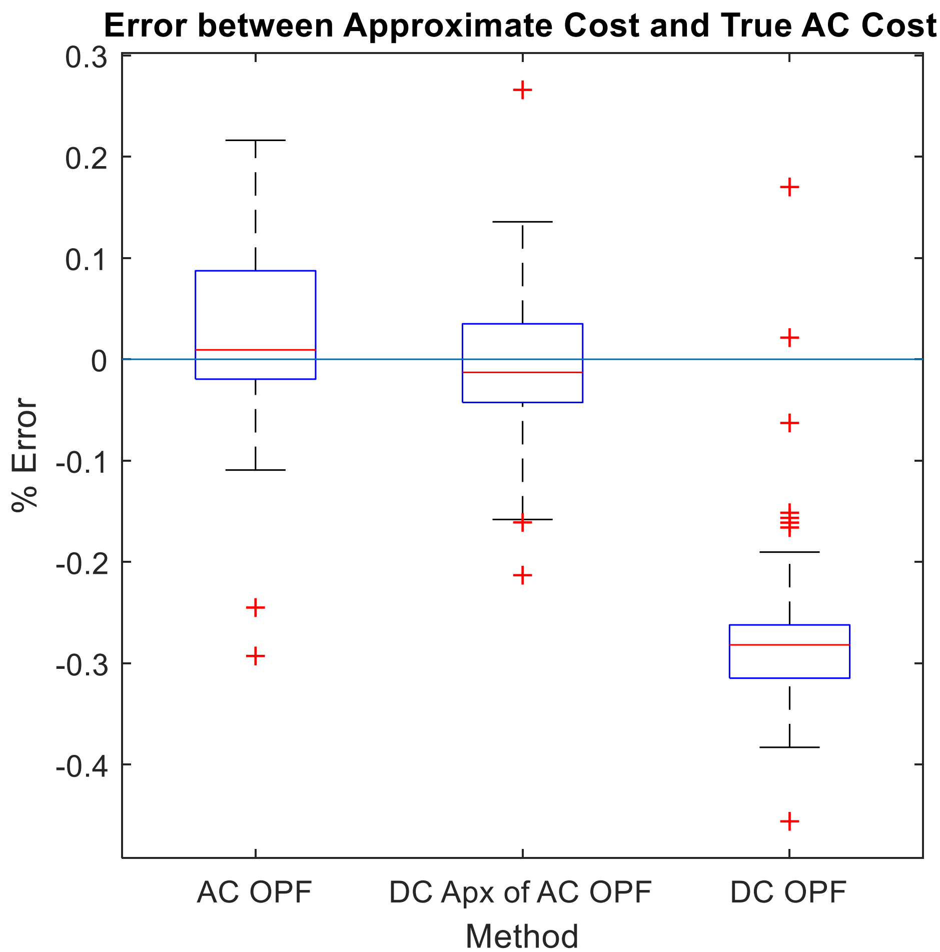

Figure 13: Error in estimated cost of transmission plans evaluated with AC-OPF relative to the true cost of each plan.

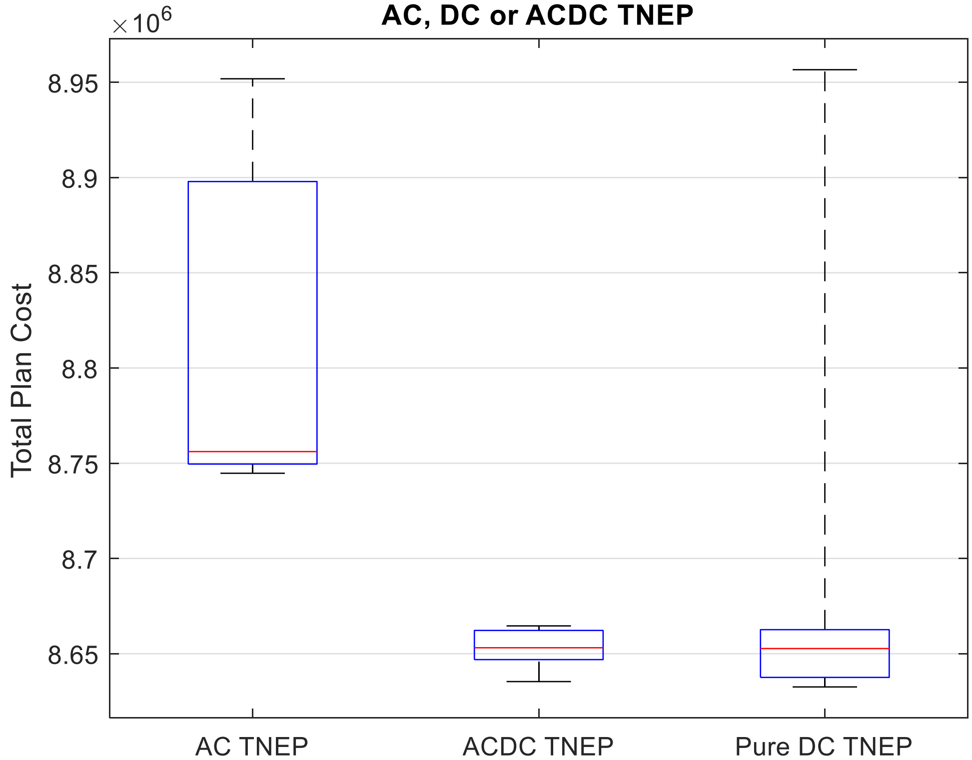

Figure 14: Total expected cost of recommended plan from each method, based on evaluation with AC-OPF.

Performance on Multi-stage TNEP Problems

Key metrics were assessed and outcomes were developed for the multi-stage problems using two case studies: a five-stage problem with fewer candidate transmission lines and uncertainty only in demand growth, and a three-stage problem with all 51 candidate transmission lines considered and multiple uncertain factors varied across long-term scenarios. The key performance metrics from the two-stage experiments were used in these cases.

The five-stage TNEP problem consists of 10 candidate transmission lines, and uncertainty in demand growth. Different instances of this case were generated with different numbers of total long-term scenarios but choosing different branching factors: in each stage, how many different demand growth scenarios branch off for the next stage. This problem is small enough to solve for the true optimal decision policy with dynamic programming. We compare the excess first-stage costs in Figure 15, and the excess total expected costs from all 5 stages across methods in Figure 16. The key outcome from this experiment is to show that while random sampling scenario reduction outperforms random partitioning, it does not typically find the optimal first-stage decisions whereas LUCA and K-means partitioning do (Figure 15). We also demonstrate that constructing a high-quality policy with LUCA-MAP is possible and exhibits better performance than those from other partitioning methods (Figure 16).

The three-stage experiment is a larger problem in its absolute size and represents several relevant uncertainties identified in WECC's public planning documents. In this case study, we compare the investment strategies from our method applied to the multi-stage problem to a non-adaptive or static strategy that does not make stage 2 and 3 investment decisions contingent on the realized uncertainty. A key outcome of this experiment is to demonstrate that adaptive strategies for investments will incur higher upfront (stage 1) cost but will have lower expected long-term cost and reduced tail end risk (Figure 17).

A second noteworthy outcome of this experiment is the demonstrated ability to generate a solution to a problem of this size. Existing methods, including the comparison methods in the previous experiments, are not able to generate a solution to this problem. The developed multi-stage version of LUCA-ALFA is able to solve large multi-stage TNEP problems that were previously too large to consider. The results from the three-stage example are consistent with the findings in the two-stage experiments that adaptive transmission planning delays investments that are only beneficial in some scenarios (Figure 18). The higher first stage costs in the adaptive plan come from more costly operations rather than from investments in transmission; the savings in later stages comes from increased adaptability to extreme scenarios and from a general efficiency gain in operations overall.

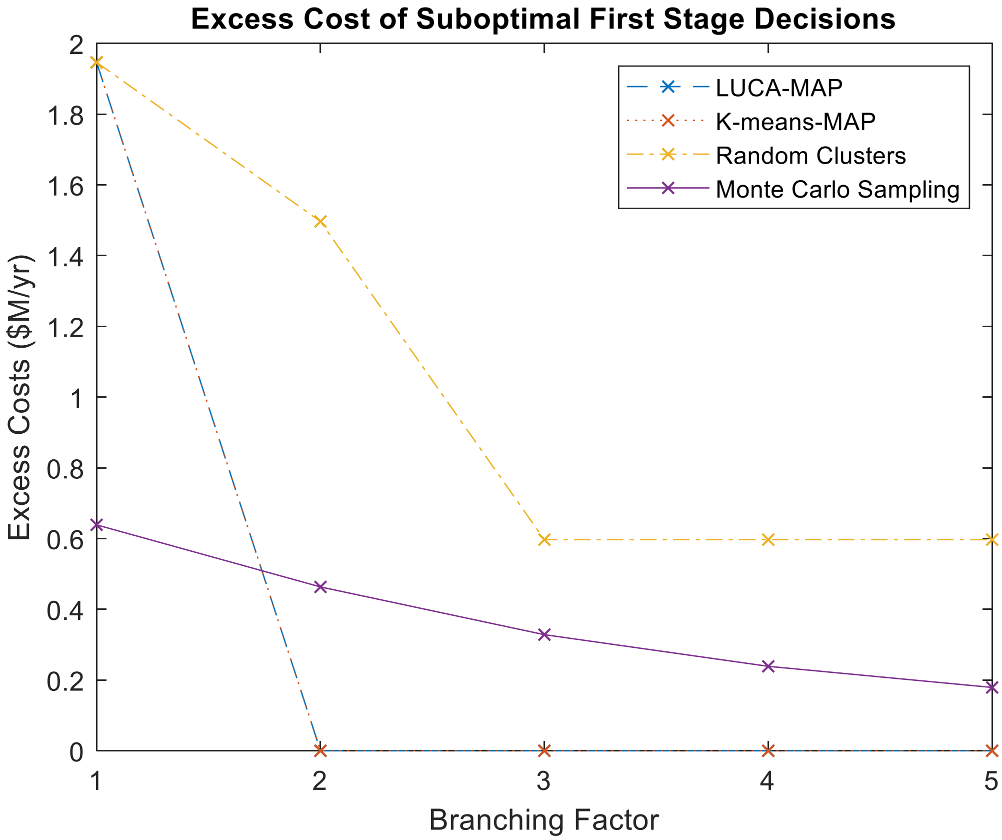

Figure 15: Excess first-stage cost of recommended investment plans relative to best known first-stage plan from any method for the five-stage case study.

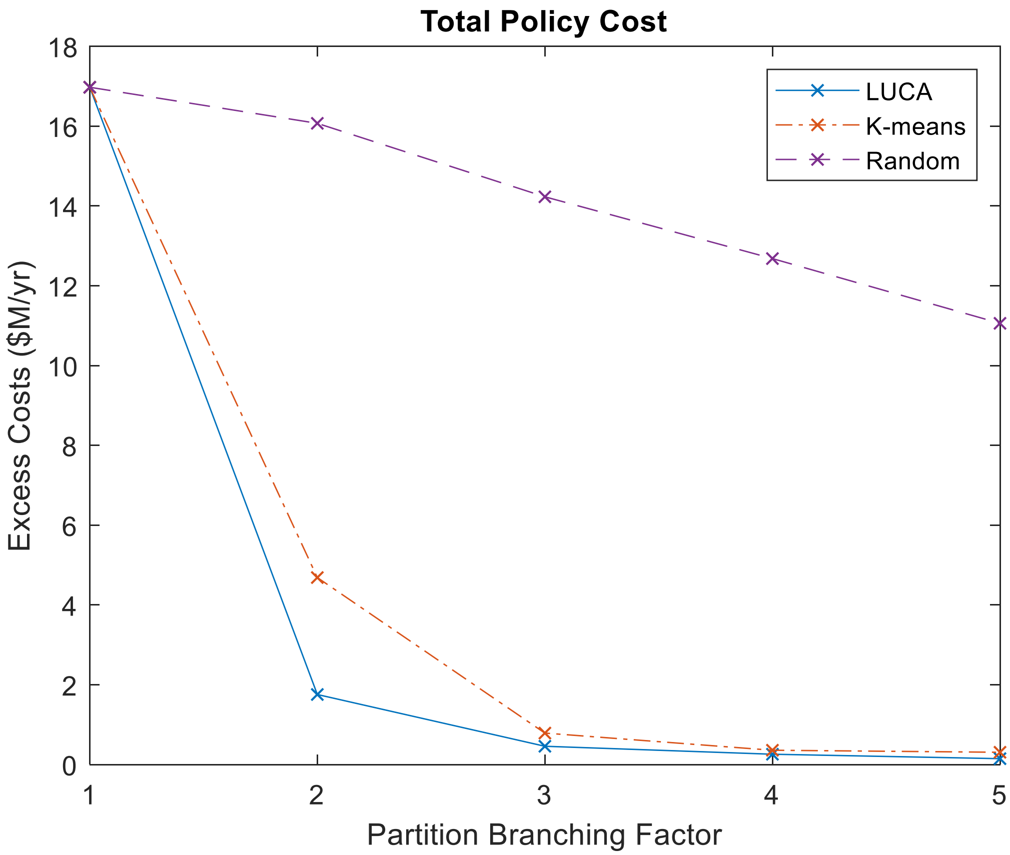

Figure 16: Excess total (first-stage plus second-stage) expected cost of recommended investment plans relative to best known investment policies from any method for the five-stage case study.

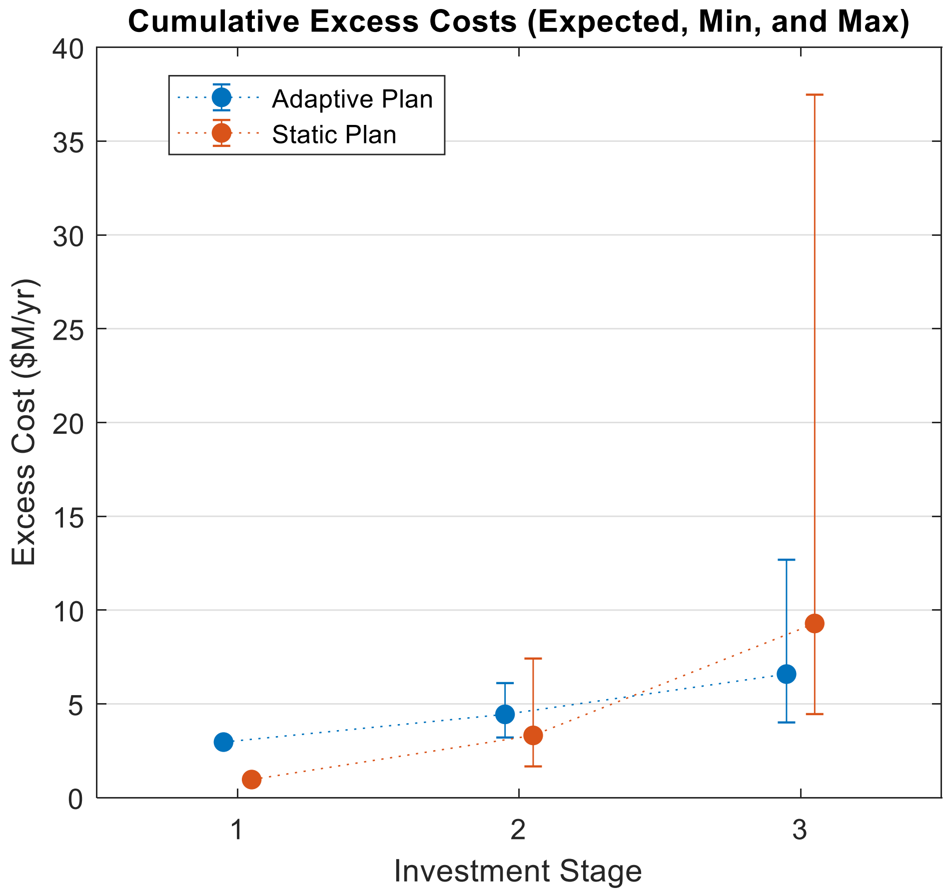

Figure 17: Excess costs in each stage relative to best known investment policy for three-stage case study. Blue ranges show the costs in each stage of the adaptive plan (different investments in Stages 2 and 3 conditional on realized scenario) and red ranges show the costs from the static plan (distinct investment decisions in each stage but Stages 2 and 3 make the same investments regardless of the scenario realized).

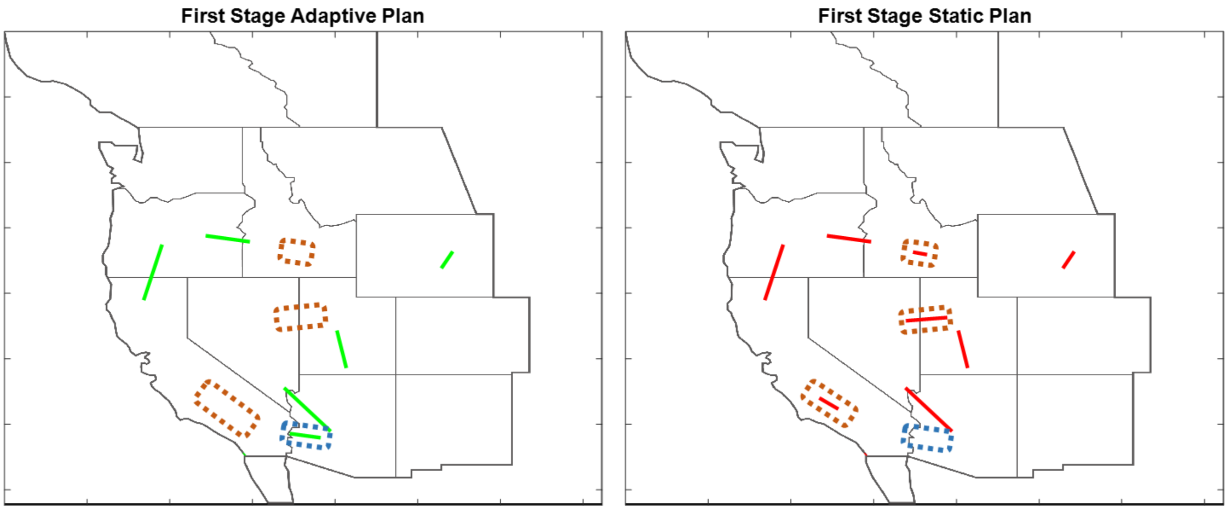

Figure 18: First stage decisions in the adaptive plan with a branching factor of 2 (left) and the static plan (right). The adaptive plan delays some investments until it is clear they are needed in later scenarios (brown circles), while making one additional strategic investment (blue circles).

Conclusions

This project was successful in achieving the original goals of developing and demonstrating methods to facilitate the solution of multi-stage transmission investment under uncertainty for very large problems that are not solvable by existing methods. The methods are computationally efficient, produce high-quality transmission plans that approximate the optimal plans of the original large decision problem, and provide tight confidence bounds to quantify the quality of the solution provided.

The experiments and case studies solved show that the method developed here has lower errors and more consistently produces solutions with lower costs than alternative methods for solving these problems. Importantly, the largest example problems tested were able to be solved using our method, but alternative methods cannot generate solutions.

The longer-term goal beyond this project is for the methods developed to be adopted in practice by organizations that conduct transmission planning. This goal will take longer to achieve, in part because of the complexity of the decision-making processes for transmission planning in the U.S. context (e.g., stakeholder processes and voting procedures in ISOs). However, outreach activities performed during this project has generated interest in the methods in several organizations, and that dialog will continue beyond the formal end to this project.

In addition, the methods developed have application to other relevant problems for the electric power sector beyond transmission planning. For example, the short-term uncertainty selection of ALFA could be adapted to aid ISOs or utilities in revisiting and prioritizing among all possible contingencies used in reliability planning. Other important innovations in electricity market design and market clearing in the coming decades may include a shift to two-stage stochastic unit commitment for day-ahead or intra-day markets to manage the risk of increased generation from renewables. This would require stochastic unit commitment models to be solved very quickly for large electricity markets. The methods developed here may provide a useful tool for enabling the implementation of these new market designs.

APPENDIX A: Product or Technology Production

Published Peer-Reviewed Articles

Bukenberger, J.P., and Webster, M. (2019a). Approximate Latent Factor Algorithm for Scenario Selection and Weighting in Transmission Expansion Planning. *IEEE Transactions on Power Systems* 35 (2) 1099-1108. DOI 10.1109/TPWRS.2019.2942925.

https://ieeexplore-ieee-org.ezaccess.libraries.psu.edu/stamp/stamp.jsp?tp=&arnumber=8846058

Bukenberger, J.P., and Webster, M. (2019b). Latent Clustering Model for Co-optimization of Transmission and Generation Investments Under Uncertainty. CIGRE 2020 Session.

https://www.cigre.org/GB/publications/papers-and-proceedings

Currently Under Review

Bukenberger, J.P. and Webster, M.D. (2020a). A Partitioning and Bounding Method for Adaptive Transmission Planning. *Operations Research*. (in review).

In Preparation:

Bukenberger, J.P. and Webster, M.D. (2020b). Stochastic Transmission Planning: Latent Factor Approximation of AC Costs with DC Subproblems. *IEEE Transactions on Power Systems.* (to be submitted).

Bukenberger, J.P. and Webster, M.D. (2020c). Multistage Transmission Planning with Correlation Clustered Scenario Trees. *INFORMS Journal on Computing*. (to be submitted).

2. For all presentations given during the past year to external audiences (in person or via webinar) that are related to this Assistance Agreement, provide the presentation title, date of presentation, event title, an electronic copy of the presentation, and a hyperlink to the presentation (if available.) Indicate if the presentation was peer-reviewed.

Bukenberger, J. and Webster, M. (2019). Stochastic Transmission Expansion Planning: Approximating the Annual AC Operating Cost with the DC Subproblem. INFORMS Annual Meeting, Seattle WA, October 24, 2019. [not peer reviewed].

Bukenberger, J. and Webster, M. (2020). A Multistage Stochastic Transmission Expansion Algorithm for Wide-Area Planning Under Uncertainty. U.S. Department of Energy, Office of Electricity, Webinar (Virtual), September 24, 2020. [Invited; not peer reviewed].

Webster, M. (2020). Technology Portfolio Planning for an Uncertain Future: Overview of Decision Concepts. Electric Power Research Institute, Program 201-B, Webinar (Virtual), August 31, 2020. [Invited; not peer reviewed].

Bukenberger, J. and Webster, M. (2020). Latent Clustering Model for Co-optimization of Transmission and Generation Investments Under Uncertainty CIGRE, Webinar (Virtual), September 3, 2020. [Peer reviewed].

REFERENCES

[1] J. Quintero, H. Zhang, Y. Chakhchoukh, V. Vittal, and G. T. Heydt, “Next generation transmission expansion planning framework: models, tools, and educational opportunities,” *IEEE Trans. Power Syst*., vol. 29, no. 4, pp. 1911–1918, Jul. 2014.

[2] B. F. Hobbs, “Optimization methods for electric utility resource planning,” *Eur. J. Oper. Res*., vol. 83, no. 1, pp. 1–20, May 1995.

[3] FERC (2007) Order no. 890: Preventing undue discrimination and preference in transmission service. Technical report, Federal Electric Regulatory Commission.

[4] FERC (2011) Order no. 1000: Transmission planning and cost allocation by transmission owning and operating public utilities. Technical report, Federal Electric Regulatory Commission.

[5] WECC (2019) 2019 integrated study program. Technical report, WECC Reliability Assessment Committee, URL https://www.wecc.org/Reliability/2019%20Integrated%20Study%20Program.pdf, [Online; accessed 2019-10-17].

[6] Eto JH (2017) Planning electric transmission lines: A review of recent regional transmission plans. Technical report, Lawrence Berkeley National Laboratory.

[7] Munoz FD, Hobbs BF, Ho JL, Kasina S (2014) An engineering-economic approach to transmission planning under market and regulatory uncertainties: Wecc case study. *IEEE Trans. Power Syst*. 29(1):307–317.

[8] Ho JL, Hobbs BF, Donohoo P, Xu Q, Kasina S, Park SW, Ouyang Y (2016) Planning transmission for uncertainty: Applications and lessons for the western interconnection. Technical report, The Johns Hopkins University.

[9] Liu Y, Sioshansi R, Conejo AJ (2018) Multistage stochastic investment planning with multiscale representation of uncertainties and decisions. *IEEE Trans. Power Syst*. 33(1):781–791.

[10] Park S, Xu Q, Hobbs BF (2019) Comparing scenario reduction methods for stochastic transmission planning. *Transmission Distribution IET Generation* 13(7):1005–1013.

[11] H. Yu, C. Y. Chung, K. P. Wong, and J. H. Zhang, “A chance constrained transmission network expansion planning method with consideration of load and wind farm uncertainties,” *IEEE Trans. Power Syst*., vol. 24, no. 3, pp. 1568–1576, Aug. 2009.

[12] H. Yu, C. Y. Chung, K. P. Wong, H. W. Lee, and J. H. Zhang, “Probabilistic load flow evaluation with hybrid Latin hypercube sampling and Cholesky decomposition,” *IEEE Trans. Power Syst*., vol. 24, no. 2, pp. 661–667, May 2009.

[13] K. Poncelet, H. Höschle, E. Delarue, A. Virag, and W. D’haeseleer, “Selecting representative days for capturing the implications of integrating intermittent renewables in generation expansion planning problems,” *IEEE Trans. Power Syst*., vol. 32, no. 3, pp. 1936–1948, May 2017.

[14] E. Shayesteh, B. F. Hobbs, and M. Amelin, “Scenario reduction, network aggregation, and DC linearisation: Which simplifications matter most in operations and planning optimisation?,” *IET Gener. Transm. & Distrib*., vol. 10, no. 11, pp. 2748–2755, 2016.

[15] F. D. Munoz, B. F. Hobbs, and J.-P. Watson, “New bounding and decomposition approaches for MILP investment problems: Multi-area transmission and generation planning under policy constraints,” *Eur. J. Oper. Res*., vol. 248, no. 3, pp. 888–898, Feb. 2016.

[16] R. Alvarez, A. Moser, and C. A. Rahmann, “Novel methodology for selecting representative operating points for the TNEP,” *IEEE Trans. Power Syst*., vol. 32, no. 3, pp. 2234–2242, May 2017.

[17] H. Park and R. Baldick, “Transmission planning under uncertainties of wind and load: Sequential approximation approach,” *IEEE Trans. Power Syst*., vol. 28, no. 3, pp. 2395–2402, Aug. 2013.

[18] A. Papavasiliou and S. S. Oren, “Multiarea stochastic unit commitment for high wind penetration in a transmission constrained network,” *Oper. Res*., vol. 61, no. 3, pp. 578–592, May 2013.

[19] R. Chatthaworn and S. Chaitusaney, “Improving method of robust transmission network expansion planning considering intermittent renewable energy generation and loads,” *IET Gener. Transm. Distrib*., vol. 9, no. 13, pp. 1621–1627, 2015.

[20] J. L. Ho, B. F. Hobbs, P. Donohoo, Q. Xu, S. Kasina, S. W. Park, and Y. Ouyang, “Planning transmission for uncertainty: Applications and lessons for the western interconnection,” The Johns Hopkins University, Jan. 2016.

[21] D. Shi, D. J. Tylavsky, V. Vittal, K. Hedman, R. Ayyanar, and Arizona State University, “Power system network reduction for engineering and economic analysis,” in ASU Electronic Theses and Dissertations, Arizona State University, 2012.

[22] “System Adequacy Planning Datasets.” [Online]. Available: https://www.wecc.biz:443/SystemAdequacyPlanning/Pages/Datasets.aspx. [Accessed: 02-Jul-2018].

[23] Mackenzie W (2018) Western interconnection gas-electric interface study. Technical report, WECC.