

**Design of Wide-Area Electric Transmission Networks under
Uncertainty: Methods for Dimensionality Reduction**

by

Pearl Elizabeth Donohoo-Vallett

S.M. Technology and Policy, Massachusetts Institute of Technology, 2010
B.S. Mechanical Engineering, Franklin W. Olin College of Engineering, 2007

Submitted to the Engineering Systems Division in Partial Fulfillment of the Requirements for
the Degree of

Doctor of Philosophy in Engineering Systems: Technology, Management and Policy
at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 2014

© 2014 Massachusetts Institute of Technology. All rights reserved.

Signature of Author.....
Engineering Systems Division
January 6, 2014

Certified by.....
Mort D. Webster
Associate Professor of Engineering Systems
Committee Co-Chair

Certified by.....
Ignacio Pérez-Arriaga
Visiting Professor, Engineering Systems Division
Professor & Director of the BP Chair on Energy & Sustainability, Pontifical Comillas University
Committee Co-Chair

Certified by.....
Yossi Sheffi
Elisha Gray II Professor of Engineering Systems
Professor of Civil and Environmental Engineering
Committee Member

Accepted by.....
Richard C. Larson
Mitsui Professor of Engineering Systems
Chair, Engineering Systems Division Education Committee

THIS PAGE INTENTIONALLY BLANK

Design of Wide-Area Electric Transmission Networks under Uncertainty: Methods for Dimensionality Reduction

by

Pearl Elizabeth Donohoo-Vallett

Submitted to the Engineering Systems Division
on January 6, 2014 in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Engineering Systems: Technology, Management and Policy

Abstract

The growth of location-constrained renewable generators and the integration of electricity markets in the United States and Europe are forcing transmission planners to consider the design of interconnection-wide systems. In this context, planners are analyzing major topological changes to the electric transmission system rather than more traditional questions of system reinforcement. Unlike a regional reinforcement problem where a planner may study tens of investments, the wide-area planning problem may consider thousands of investments. Complicating this already challenging problem is uncertainty with respect to future renewable-generation location. Transmission access, however, is imperative for these resources, which are often located distant from electrical demand. This dissertation frames the strategic planning problem and develops dimensionality reduction methods to solve this otherwise computationally intractable problem.

This work demonstrates three complementary methods to tractably solve multi-stage stochastic transmission network expansion planning. The first method, the St. Clair Screening Model, limits the number of investments which must be. The model iteratively uses a linear relaxation of the multi-period deterministic transmission expansion planning model to identify transmission corridors and specific investments of interest. The second approach is to develop a reduced-order model of the problem. Creating a reduced order transformation of the problem is difficult due to the binary investment variables, categorical data, and networked nature of the problem. The approach presented here explores two alternative techniques from image recognition, the Method of Moments and Principal Component Analysis, to reduce the dimensionality. Interpolation is then performed in the lower dimensional space. Finally, the third method embeds the reduced order representation within an Approximate Dynamic Programming framework. Approximate Dynamic Programming is a heuristic methodology which combines Monte Carlo methods with a reduced order model of the value function to solve high dimensionality optimization problems. All three approaches are demonstrated on an illustrative interconnection-wide case study problem considering the Western Electric Coordinating Council.

Thesis Co-Supervisor: Mort Webster

Title: Associate Professor of Engineering Systems

Thesis Co-Supervisor: Ignacio Pérez-Arriaga

Title: Professor and BP Chair on Energy & Sustainability, Pontifical Comillas University

THIS PAGE INTENTIONALLY BLANK

*This work is dedicated to the memory of
Professor Michael Moody, a beloved teacher and mentor.*

THIS PAGE INTENTIONALLY BLANK

Acknowledgements

While it takes a village to raise a child, it took an international force to create this dissertation. I have been blessed to draw on support from friends, colleagues and researchers across both the United States and Europe. Without their help, this research would never have made it off the ground.

First and foremost, this research was shaped and supported by my advisors Ignacio Pérez-Arriaga and Mort Webster. I started this research with an interest in the power sector but without power systems or stochastic optimization experience. As mentors, teachers and role models Mort and Ignacio have patiently guided me through the past four years.

At the Institute, I owe great thanks to the faculty, staff and my student cohorts within the Engineering Systems Division. Yossi Sheffi generously took time from his busy schedule to serve on my committee, kept us all on track, and pushed me forward. Beth Milnes helped keep me sane and made sure I ticked the right boxes; she is the glue that keeps the ESD doctoral program together. I also owe thanks to my Technology and Policy Program, ESD Doctoral Program and Electricity Student Research Group cohorts, especially Nidhi, Bryan, Rhonda, Berk, Maite, Fernando, Olivia, and Don.

Throughout my time at MIT I have held two intellectual homes, MIT and the National Renewable Energy Laboratory. Within NREL, the folks in the Transmission and Grid Integration Group have taught me about power system operation and the on-the-ground issues utilities are facing today. For the past several years, Michael Milligan has been my enthusiastic mentor at the lab; I aspire to simultaneously be as expert and humble as Michael. Other folks at the lab who have helped me with my research and kept me sane include Debbie Lew, Bri-Mathias Hodge, Aaron Bloom, Vahan Grevorgian, Mohit Singh, and YC Zhang.

I also owe great thanks to the community of researchers at IIT at the Pontifical University of Comillas in Spain. Without prompting, everyone at the institute has been incredibly warm and welcoming. Andrés Ramos and Fernando de Cuadra especially have gone out of their way to aid this work. Andrés is an absolute GAMS wizard and also taught me about traditional optimization approaches. Fernando cheerfully donated his time and innovative thought to the problem, suggesting the Method of Moments for dimensionality reduction and interpolation.

Last, but certainly not least, I must acknowledge my family. First, Mom, Dad, Rose and Claire, who have offered endless support and listened to me talk for years about the trials and tribulations research that I never fully explained. Second, Dave, Carol, Aaron and Laura Vallett who have opened their doors and hearts to me as Paul's partner in life. I must also acknowledge my family of friends, especially Andrew and Russell, who have followed the highs and evened-out the lows of the past six years. Most importantly is husband, Paul, who has never wavered in support of my work or our young family of two. Paul, I love you more than words, especially these, could ever express. Now and always, IBIP.

Contents

1. Introduction.....	17
1.1 Transmission Expansion Planning	27
1.1.1 Motivations for Transmission Expansion Planning	28
1.1.2 Timescales for Transmission Expansion Planning.....	31
1.2 Optimal Transmission Network Expansion Planning Model Formulations	33
1.2.1 Multi-Stage Planning.....	38
1.2.2 Planning Under Uncertainty.....	43
1.2.3 Stochastic Multi-Stage Planning with Recourse	46
1.3 Dissertation Overview.....	50
2. St. Clair Screening Model: Dimensionality Reduction through Algorithmic Screening	53
2.1 St. Clair Screening Model Formulation	54
2.1.1 Characterization of Transmission Investments	55
2.1.2 Linearization of the TNEP Problem.....	58
2.1.3 Transformation from Corridor Capacity to Investments.....	63
2.1.4 St. Clair Screening Model Flow	64
2.2 Screening Model Demonstration Problem	65
2.2.1 WECC System Model	66
2.2.2 Generation Sampling.....	68
2.3 Results	70
2.3.1 Reduction in Transmission Investments	71
2.3.2 Frequency of Corridor Use.....	72
2.3.3 Insights into Linear Relaxation	78
2.4 Conclusions and Future Work.....	82
3. Dimensionality Reduction and Interpolation for Multi-stage Stochastic Transmission Expansion Planning Algorithms	85

3.1	Approximating the Multi-Stage Stochastic Transmission Expansion Network Planning.....	86
3.2	Dimensionality Reduction.....	89
3.2.1	Method of Moments	90
3.2.2	Principal Component Analysis	92
3.3	Interpolation	95
3.4	Evaluation of Dimensionality Reduction and Interpolation Techniques	96
3.4.1	Cost Prediction	98
3.4.2	Ranking Prediction	101
3.5	Conclusions	103
3.6	Future Work	104
4.	STEP Model: An Approximate Dynamic Programming Approach for Multi-Stage Stochastic Transmission Network Expansion Planning	107
4.1	Approximate Dynamic Programming	107
4.1.1	Approximate Dynamic Programming Framework.....	109
4.1.2	Exploration and Exploitation Phases of ADP	111
4.2	ADP for the MS-TNEP Problem.....	112
4.2.1	Problem Structure.....	113
4.2.2	Approximate Value Function	115
4.2.3	Exploration Phase.....	115
4.2.4	Exploitation Phase.....	118
4.3	Evaluation of the STEP Model.....	124
4.3.1	Test System	125
4.3.2	Cost of Plans.....	127
4.3.3	Composition of First-Stage Plans.....	128
4.3.4	Convergence.....	132

4.4	Conclusions and Future Work.....	132
5.	Conclusions.....	135
5.1	Dissertation Summary.....	135
5.2	Policy Implications.....	137
5.3	Future Work	138
6.	Works Cited	141
Appendix A.	Map Generation for Method of Moments	149
Appendix B.	Depth First Search Transmission Plan Generation.....	151
B.1	Disjunctive Formulation.....	154
B.2	Calculation of M Values	158

Table of Figures

Figure 1-1 Western Planning Areas [69]	19
Figure 1-2 Wind Resource Map of Europe [22]	19
Figure 1-3 Solar Resource Map of Europe [23].....	19
Figure 1-4 Wind Resource Map of the United States [43]	20
Figure 1-5 Solar Resource Map of the United States [44].....	20
Figure 1-6 Three Electrical Interconnections in the United States	21
Figure 1-7 ReEds Transmission Expansion Planning Results from the Renewable Electricity Futures Study.....	24
Figure 1-8 AEP Transmission Overlay Plan [3].....	24
Figure 1-9 Decision Tree Representation of the Multi-Stage Stochastic Transmission Expansion Problem with Uncertainty in Future Generation.....	26
Figure 1-10 Steps in the Transmission Planning Process	28
Figure 1-11 Illustration of Alternating Current Power	29
Figure 1-12 Comparison of Static, Static Sequential and Multi-Stage Modeling	39
Figure 1-13 Test System for Time Horizon Simplification	40
Figure 1-14 Spanish Case Study Static Results for 2015 (L) and 2035 (R)	41
Figure 1-15 Spanish 2015 Case Study Dynamic Programming Results (L) and Static Planning (R).....	42
Figure 1-16 Expansion of the Decision Space Including Multiple Time Horizons.....	43
Figure 1-17 Test System for Uncertainty Simplification.....	45
Figure 1-18 Single Investment, Two Decision-Stage, Two-Uncertainty Stage Problem Illustration	48
Figure 1-19 Network Effects Illustration	50
Figure 2-1 Schematic of St. Clair Screening Model.....	55
Figure 2-2 St. Clair Curve [62].....	57
Figure 2-3 Piecewise St. Clair Curve Used	57
Figure 2-4 Distortions in Per Unit Cost Due to St. Clair Assumptions and Linearization	61
Figure 2-5 St. Clair Screening Model.....	65

Figure 2-6 North American Interconnections	66
Figure 2-7 Existing Nodes and Susceptance Lines in the Modified WECC System.....	67
Figure 2-8 Existing Nodes (Blue), WREZs (Green) and lines in the Modified WECC System.....	69
Figure 2-9 Number of New Unique Corridors with Additional Generation Samples	73
Figure 2-10 Frequency of Corridor Development by Percentage of Developed Corridors	74
Figure 2-11 Frequency of Corridor Development by Percentage of Developed Corridors in First-Stage Scenarios	75
Figure 2-12 Map of Corridors Developed in At Least 90% of Scenarios	76
Figure 2-13 Illustration of Short-line Bias in Demonstration Case	77
Figure 2-14 Artificial Network Constraint Example	80
Figure 2-15 Histogram of Investment Variables by Percentage of 345kV Capacity	82
Figure 3-1 Aerial of the Grand Canyon [35]	87
Figure 3-2 Illustration of Costs	87
Figure 3-3 Feature Identification Test System.....	88
Figure 3-4 Relative Costs for Feature Identification in This Test System	88
Figure 3-5 Hamming Distance Example.....	89
Figure 3-6 Test System for Value Function Approximation Testing	98
Figure 4-1 Illustration of Forward Pass in ADP Algorithm	110
Figure 4-2 Decision Tree for Sample Problem.....	111
Figure 4-3 Exploration Phase Sample Paths Through the Tree.....	111
Figure 4-4 Estimated Expected Values for ADP Example Problem	112
Figure 4-5 Optimal Actions and Uncertainty Outcomes for the ADP Example Problem	112
Figure 4-6 Optimal Solution Structure and State Variables	114
Figure 4-7 Four Phases of the Explorative Phase	116
Figure 4-8 One Iteration of the Second Step of the Exploration Phase.....	117
Figure 4-9 Process for the Generation of Additional Valid Plans	118

Figure 4-10 Overview of forward pass	119
Figure 4-11 Selecting a Set of Candidate Lines.....	120
Figure 4-12 Backward Pass	123
Figure 4-13 Process for Creation of Bias in ADP Algorithm.....	124
Figure 4-14 Optimal Solution Structure for STEP Test Problem.....	126
Figure 4-15 Cost Comparison of STEM and Branch and Bound Solution	127
Figure 4-16 Distribution of Expected First-Stage Costs.....	128
Figure 4-17 Reduction in the Number of Potential Transmission Investments.....	129
Figure 4-18 Distribution of All Transmission Investments	130
Figure 4-19 Distribution of the Most Frequent Transmission Investments	130
Figure 4-20 All Lines in the First-Stage Top 30 Plans	131
Figure 4-21 Lines in at least 50% of Top 30 First-Stage Plans	131
Figure 4-22 Lines in at least 75% of Top 30 First-Stage Plans	131
Figure 4-23 Lines in All Top 30 First-Stage Plans.....	131
Figure 4-24 Action Convergence.....	132
Figure 6-1 Outline of Depth-First Search for Non-Served Energy.....	152
Figure 6-2 Binary Search Tree Composed of Randomly Partitioned Sets of Lines	153

Table of Tables

Table 1-1 Timescales for Transmission Expansion Planning.....	31
Table 1-2 Scenario Costs for Uncertainty Test System.....	45
Table 1-3 Summary of Tradeoffs Considered in Stochastic Multi-Stage Models.....	47
Table 2-1 Surge Impedance Loading and Thermal Ratings [2].....	57
Table 2-2 Characteristic Transmission Line [2]	58
Table 2-3 Linearized Transmission Costs Assuming Thermal Limit Constraint (Non-Annualized Costs).....	59
Table 2-4 Illustration of Linear Model Results.....	64
Table 2-5 Capacity for 500 Mile Lines by Investment Type.....	64
Table 2-6 Size Comparison of Original and Simplified WECC Models.....	67
Table 2-7 Summary of Transmission Line Properties	68
Table 2-8 Number of Generation Scenarios by Expected Percentage of WREZs.....	70
Table 2-9 X-Axis Ordering Given in Figure 2-9	73
Table 2-10 Percentage of Corridors Developed in At Least 90% of Scenarios by Type .	76
Table 2-11 WREZ Wind and Solar Resources (GW) Accessed by Lines in the First Stage	78
Table 2-12 Number and Percentage of Lines By a Single Voltage Level St. Clair Screening Model	81
Table 2-13 Summary of Dimensionality Reduction Through St. Clair Filter	82
Table 3-1 Binary Matrix Representation	93
Table 3-2 Capacity Matrix Representation.....	93
Table 3-3 Covariance Matrix for Table 3-1 <i>Data Centered Before Calculation</i>	94
Table 3-4 Principal Components for PCA Test System	94
Table 3-5 Eigenvalues for PCA Test System	94
Table 3-6 Reconstructed Test PCA Data Using One Principal Component.....	95
Table 3-7 Reconstructed Test PCA Data Using Two Principal Components	95
Table 3-8 Cost Prediction: Moment Results.....	99
Table 3-9 Moments Used in Six Moment Set.....	99

Table 3-10 Cost Prediction: PCA Results.....	101
Table 3-11 Moment Results.....	102
Table 3-12 PCA Results Trials	102
Table 4-1 Generic Description of ADP Double-Pass Algorithm.....	110
Table 4-2 STEP Model Parameters.....	125
Table 4-3 Comparison of Stochastic and Scenario Stage-One Results	126
Table 6-1 Transformation of Latitude and Longitude to Cartesian Coordinates.....	149
Table 6-2 Calculation of Zeroth Moment for Arbitrary Line shown in.....	150

1. Introduction

The transmission investments made today will shape the power system for decades to come, determining what types of new generation will be available to meet policy goals from climate change to local air quality and water consumption. This dissertation addresses the problem of planning this next generation electric transmission network. This is an immense task, given the overwhelming dimensionality and large uncertainty that characterizes this complex problem. Planning the next generation grid is significantly complicated by the anticipated large penetration of wind and solar generation and the vast scope of present electricity markets. This work proposes a comprehensive framework to tackle capacity expansion of the transmission network and develops the building blocks to complete the approach. In isolation, these tools are designed to aid transmission planners in identifying new robust transmission investment patterns.

The transmission grid is the backbone of the electric system, transporting power from generators to distribution centers to our homes, businesses and schools. The extra high voltage transmission network (EHV), defined to be all transmission lines rated 345kV and above, is the bulk power transportation system from large generators to demand centers, the federal highway system of the road network. Residential, commercial and industrial customers all rely on the electric grid for reliable and economic service. When the transmission system fails, the power system fails. As was dramatically demonstrated during the Northeast Blackout of 2003 in the United States, the successful operation of this network is vital. Without it, commerce grinds to a stop and the other networks, such as transportation and communication that rely on the power system, also fail. However, as electricity demand grows, population centers change and new generation replaces the old, the demands on the transmission network change. To meet the demands of the evolving system, new transmission investments must be made to maintain the reliability and economic viability of the network.

Climate change and other policies which induce high penetrations of renewable generation are forcing both the transmission network and its planning to evolve quickly.

Transmission planning has historically been done on a regional (United States¹) or country-wide basis (in Europe); the planning areas in the western United States, for example, are shown in Figure 1-1. In the past, these regions were able to plan independently as their systems were not tightly interconnected and fossil generation could be sited within the regions to meet electric demands. Unlike fossil generation, however, renewable energy generators are location constrained. That is, renewable energy generators can only economically be sited in areas with strong natural resources. As seen in Figure 1-2 through Figure 1-5, these resources are not distributed evenly across regions or nations and are not correlated with areas of high load (along the coasts in the United States and in northern Europe). Accessing these resources requires significant transmission investment as detailed in reports such as the Eastern Wind Integration and Transmission Study [42], and dramatically displayed in 2011 when a \$12 billion wind farm proposed in the Texas Panhandle was cancelled due to lack of transmission capacity [26].

¹ “To call the U.S. grid balkanized would insult the Macedonian.” Mark Spitzer, former FERC Chairman. [66]



Figure 1-1 Western Planning Areas [70]

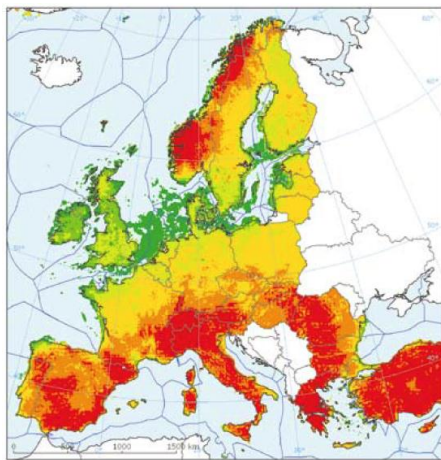


Figure 1-2 Wind Resource Map of Europe [23]

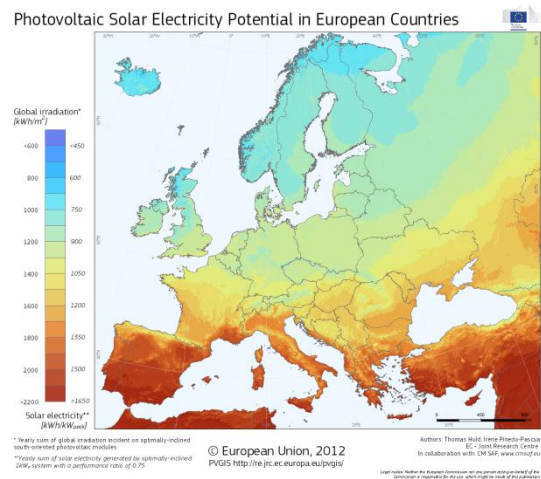


Figure 1-3 Solar Resource Map of Europe [24]

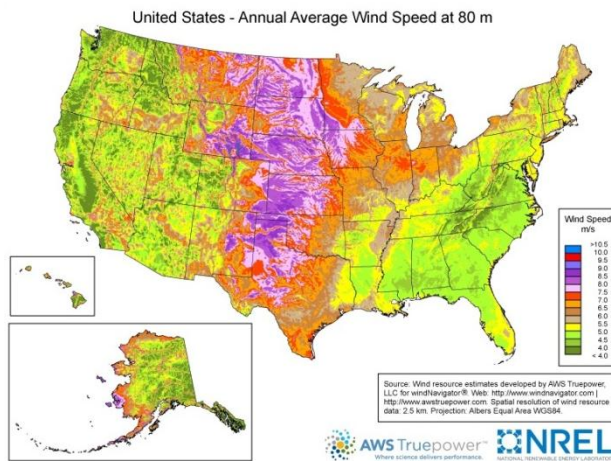


Figure 1-4 Wind Resource Map of the United States [44]

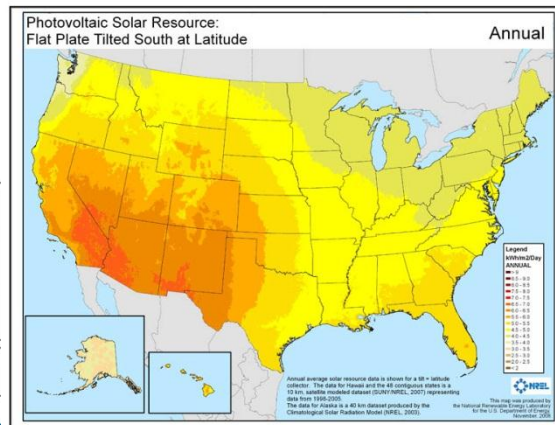


Figure 1-5 Solar Resource Map of the United States [45]

The historic conditions which allowed electric areas to be planned separately no longer hold true. Location-constrained renewables are being considered to replace location-unconstrained fossil generation. The once independent systems are becoming more interconnected and operational issues (termed seams issues) resulting from planning and operating interconnected systems independently are growing. To cope with these and other issues, larger areas are being considered for transmission planning. In Europe, the European Network of Transmission System Operators (electricity) has been tasked with developing European-wide transmission plans. In the United States, the Department of Energy has sponsored transmission planning across the three electrical interconnects, shown in Figure 1-6 [19], [20], [68], and the Federal Electricity Regulatory Commission has mandated wider regional planning [25].

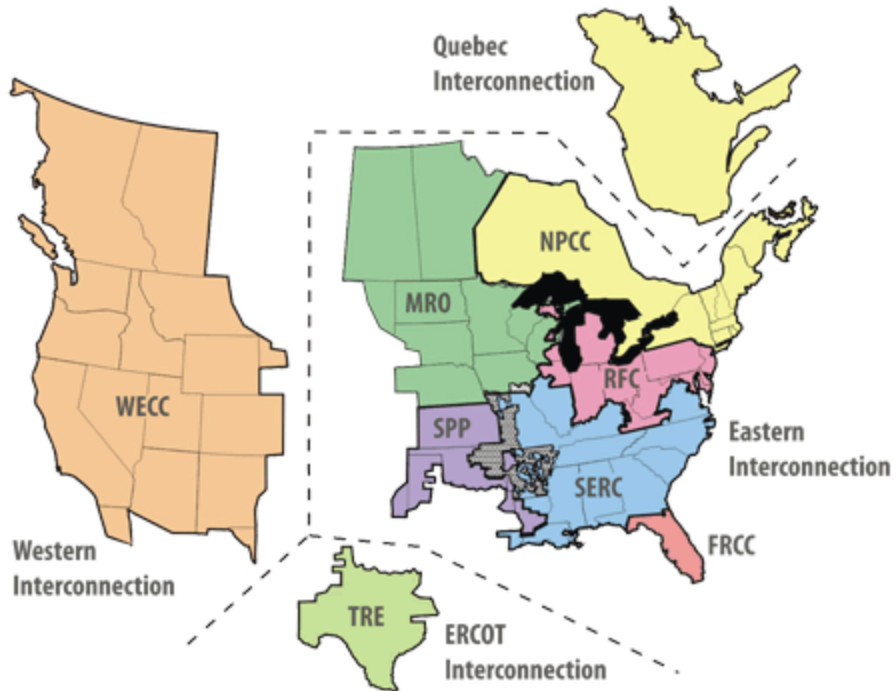


Figure 1-6 Three Electrical Interconnections in the United States

Under the most favorable conditions, planning the transmission network is a non-trivial problem. The transmission network is an infrastructure system with investments characterized by their high capital costs and long lifetimes. Each transmission line is a lumpy investment with costs ranging from \$1.1 million to \$4 million per mile depending on the voltage rating of the line [3]. These cost estimates, however, are more realistically viewed as lower bounds. For example, aesthetic and environmental concerns elevated the cost of a 69 mile line in Connecticut to \$1.3 billion compared to the \$76 million predicted using the costs above [47]. Transmission investments must also be planned prospectively. The time from a transmission line’s selection for construction to energization is five to ten years. In that time, the need for the investment may be eliminated, for example a new power plant may have been constructed elsewhere in the system, or exacerbated, for example a boom in the economy prompting higher electricity demands. Once construct-

ed, transmission lines have operating lifespans of 40 to 50 years². As part of a networked system, the effects of adding capacity are often non-intuitive. In some cases, adding capacity to the system can actually exacerbate rather than relieve problems.

Planning the transmission network to incorporate high renewable generation penetrations is complicated both by the geographic scope required and the uncertainty in the development of future generation. The growth of the location constrained generation is dominantly policy driven. These policies, however, are unstable due to changes in political will. For example, in the United States the development of wind resources has been tightly coupled and fluctuated with the Production Tax Credit (PTC) which subsidizes wind power production. This policy uncertainty is compounded by the mismatch in generator build time and transmission build time [52]. Location-constrained generators such as wind and photovoltaics require a two to five year construction timeline, while transmission lines require five to ten years to plan and construct. With generator build times significantly lower than those for transmission, planners are forced to either anticipate new generation and build potentially unnecessary infrastructure or be reactive in building new transmission, potentially discouraging new generation investment.

Tools to support this new type of wide-area stochastic planning do not yet exist. There are models which demonstrate that new capacity is needed, for example the ReEDS model [60]; however, these models lack the technical detail required to examine specific plans. An example of the plans produced by the ReEDS model, shown in Figure 1-7, identifies capacities between grossly aggregated regions, but does not have the detail required to specify investments (eg a 500kV double circuit line between two specific system buses). There are also indicative transmission plans, the most well-known of these developed by American Electric Power (AEP) and shown in Figure 1-8. These indicative plans, also sometimes referred to as crayon plans, have not been economically or technically evaluated; these crayon plans are also not created using underlying models or other scientific rationale. Plans such as the AEP proposal are also often viewed with skepticism

² In reality, transmission lines are rarely if ever decommissioned. The wires, pylons and insulators are simply replaced or upgraded as necessary.

by regulators and other stakeholders when they are proposed by the transmission companies which would profit from their construction.

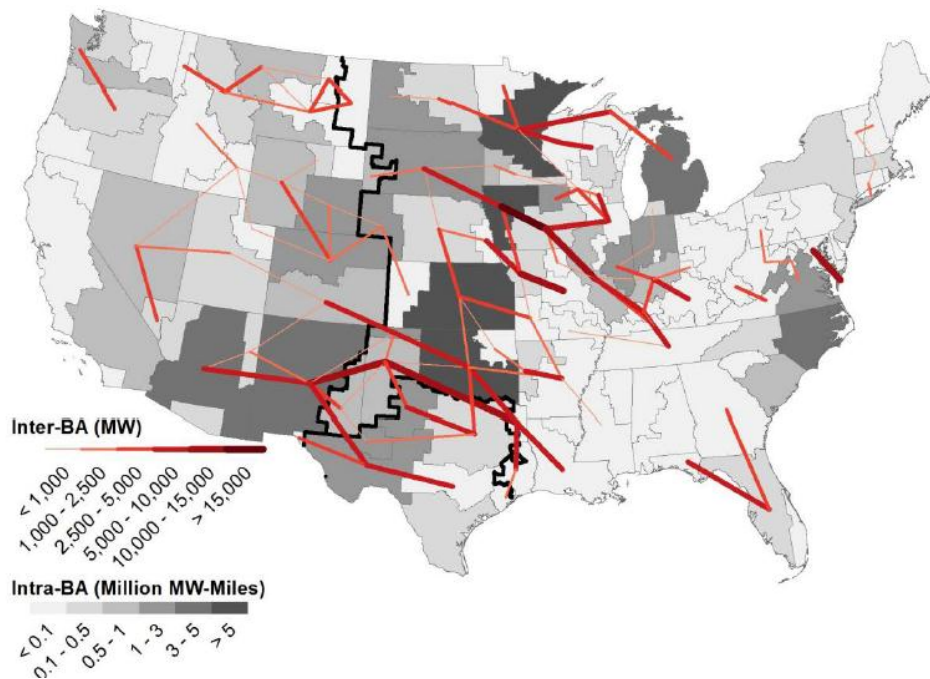


Figure 1-7 ReEds Transmission Expansion Planning Results from the Renewable Electricity Futures Study

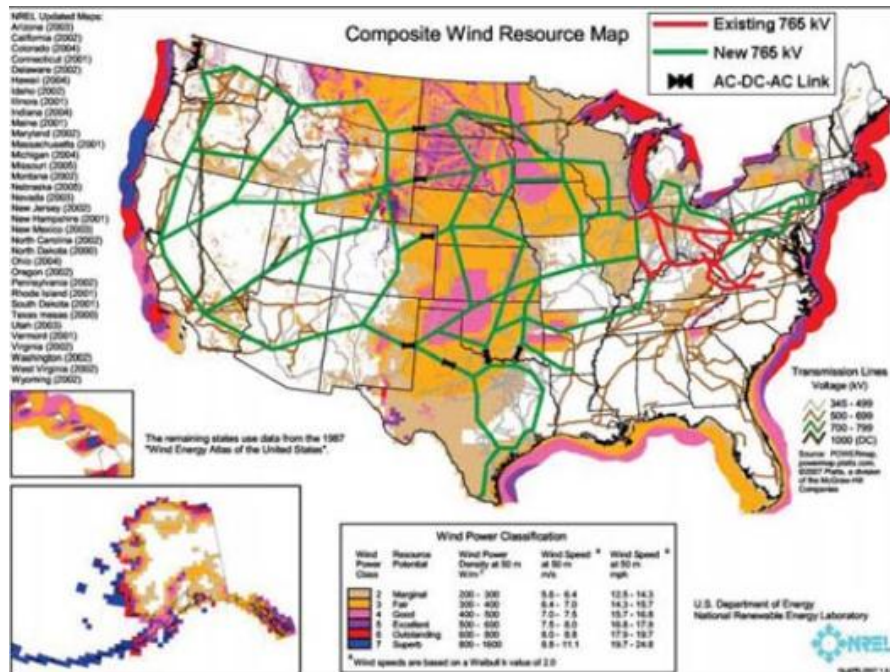


Figure 1-8 AEP Transmission Overlay Plan [4]

Tools to guide transmission planners toward high-value investment options must capture the temporal and technical details inherent to the problem. While the technical models are well defined, the temporal characteristics are less well framed. Given the stochastic and temporal characteristics, the planning problem can very naturally be represented as a decision tree as shown in Figure 1-9. Following a single path (darkened) through the decision tree, the present day decision ($T=0$) is what new transmission lines should be added. These lines are planned without knowing where the new generation will be added to the system. While these new transmission lines are being sited and constructed, new generation is being added to the system. In the two-stage model presented in Figure 1-9, both the new generation and new transmission lines come online in year ten. In year 10, the system planner can observe the new generation investments and with this new knowledge plan new lines to come online in year 20. When the full problem is enumerated and solved in this framework, discussed further in 1.2.3, the value of any transmission investments today is dependent on how well it performs across a variety of future scenarios.

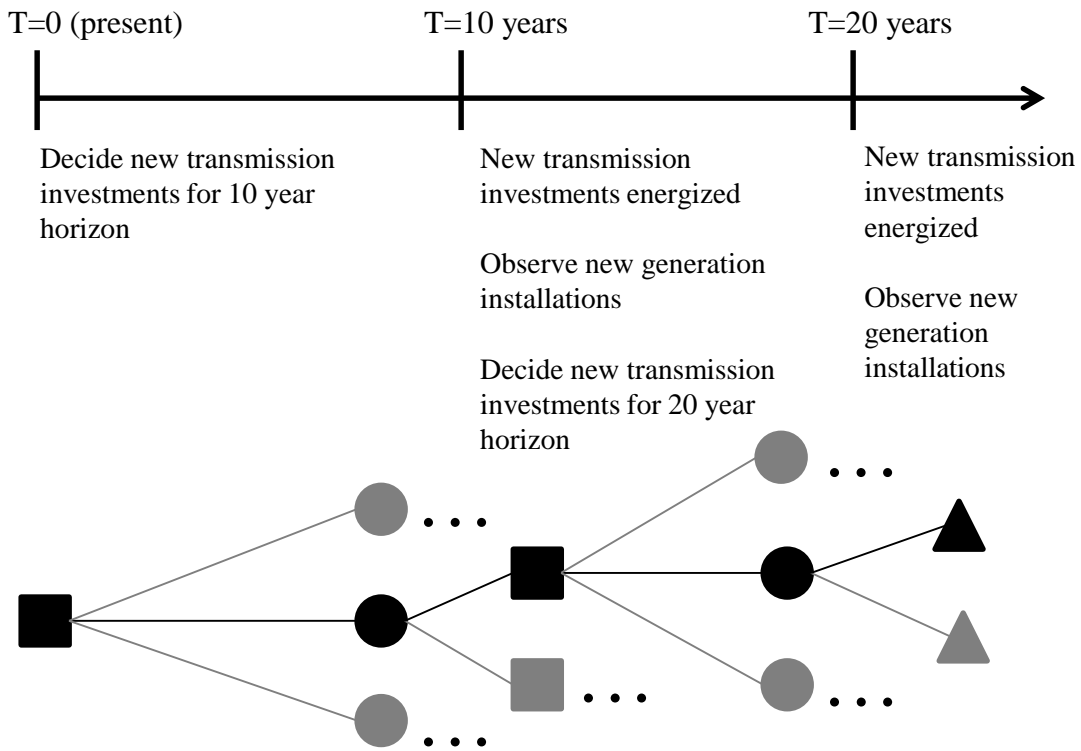


Figure 1-9 Decision Tree Representation of the Multi-Stage Stochastic Transmission Expansion Problem with Uncertainty in Future Generation

In a standard decision tree representation, squares represent decisions points and circles represent the realization of an uncertain parameter or variable. In this case, the squares represent the selection of new transmission lines (each spoke coming from the square represents a distinct investment option) and the circles represent the resolution of both the quantity and location of new generation (each spoke coming from the circle represents a specific future scenario of generation expansion).

1.1 Transmission Expansion Planning

Transmission planning is often used colloquially as an omnibus description for the entire process of adding a new transmission investment to the existing system. This process involves deciding on the new assets to be added, reliability analyses, cost-allocation and siting and routing for new lines. Traditionally, cost allocation and siting have been the most contentious issues in transmission expansion planning. Cost allocation is the determination of who pays for the new multi-million dollar investment and siting is both the determination of the geographic route and permitting of that route. Siting is a contentious issue because transmission towers and lines belong to the class of locally unwanted land uses (LULUs) which reduce property value. It is also one of the few pathways that local authorities and stakeholders have to block lines for environmental, aesthetic or economic reasons. As shown in Figure 1-10, these processes are not independent but instead inform one another. For example, the expansion plan decided upon may fail the required reliability analysis or stakeholder input may require a transmission line to be rerouted. With the new rerouting, the transmission line may no longer be an economic solution. For the purposes of this dissertation, however, transmission expansion planning will be more narrowly defined to refer to the selection of new transmission assets.

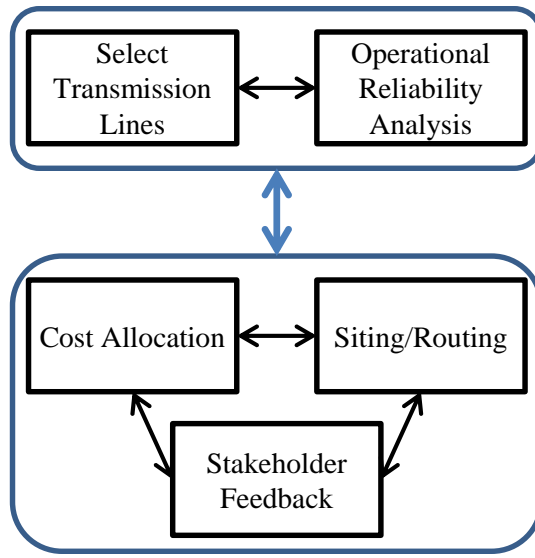


Figure 1-10 Steps in the Transmission Planning Process

1.1.1 Motivations for Transmission Expansion Planning

Transmission investments are made for two major reasons: improving the reliability of the power system and lowering the cost to operate the power system. These investments can be transmission lines (overhead lines or underground cables) or additional support equipment such as protection systems, transformers or reactive power controls which can improve the operation of existing transmission lines. The traditional treatment of transmission expansion planning frames the problem as trading-off between economic benefits through lower system operating costs and the investment costs of new transmission investments. In this view, the reliability of the system is treated as a constraint in the planning process. There are, however, streams of research which consider reliability probabilistically and explicitly within a risk context [14].

The reliability requirements of the power system can be represented both with economic costs and engineering constraints. From an economic perspective, the power system should be able to meet the electric demand all hours of the year with very low probability of non-served demand, more generally called electricity non-served (EENS) or power non-served (PNS)³. The engineering constraints on the system are more compli-

³ In the United States, the a common loss of load expectation used is one day in ten years.

cated. The majority of power flows through the transmission network are alternating current (AC) as shown in Figure 1-11. With alternating current, the power flowing through each line on the transmission network fluctuates with time at a set frequency (60 hertz in the US and 50 hertz in Europe) and a set voltage level (for example 345kV). If the frequency and voltage are not maintained within tight bounds, some demand may not be met (brownouts) and/or the power system may totally collapse (blackouts). These bounds must be maintained even if there are large disturbances on the power system, such as a generator unexpectedly failing or a storm disabling a transmission line. These failures of individual pieces of equipment are called contingencies, and the most common contingency scenarios to consider contain the failure of any one single component, referred to in the field as n-1 contingency analysis. Investments in the transmission network can alleviate these threats by, for example, providing alternative routes for power to flow or providing voltage or frequency support directly.

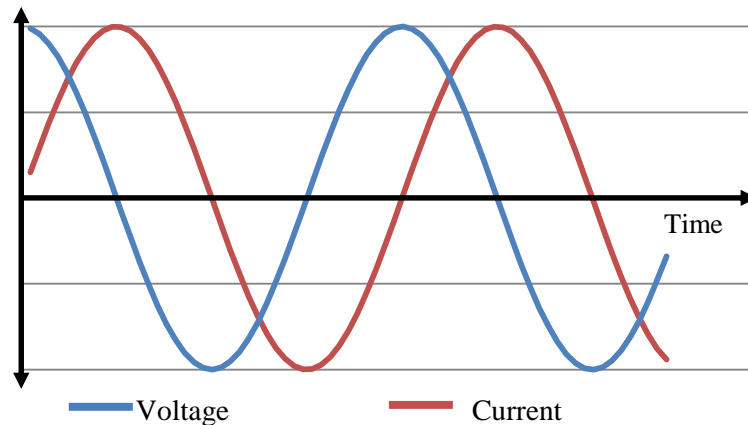


Figure 1-11 Illustration of Alternating Current Power

Investments in the transmission network can lower the operating costs of the power system by facilitating access to lower cost generation sources. There are areas of the power system which experience congestion. Like traffic congestion, congestion in the power system means that the flow into or out of a specific area is limited. For example, a hydro power plant may have a capacity of 1,000 MW but may be connected to the system with a transmission line capable of carrying only 500 MW to the demand center. As a result of the congestion at the hydro power plant, a more expensive generator closer to

the load must be dispatched. If a new transmission line is added to the system which connects the hydro power plant to the demand center with an additional 500 MW, the congestion in the system will be reduced. The more expensive plant will no longer be dispatched, and the cost of operating the system will be lessened.

New transmission investments can also reduce the losses in the system. Losses in the power system reflect electricity that is produced by generators but does not make it to the demand in order to do useful work. Instead, this useful energy is lost to the system, for example through production of heat in transmission lines. In the United States, for example, approximately 7% of all generation, or 262,000 MWh is consumed annually by losses across the transmission and distribution networks [21]. Using a back-of-the-envelope calculation, this implies that in the United States \$15.7 billion is spent annually on non-productive generation⁴. Losses in transmission lines can be reduced either by investment in new equipment to increase the operating voltage transmission of lines or by adding new transmission lines such that less power flows on each line.

The addition of transmission capacity can also lower costs by easing operational and market restrictions within the power system. For example, when bordering areas are well connected, the generation reserves burden can be shared at a lower cost rather than borne individually [36]. Transmission capacity can also help reduce the effects of variability inherent to wind and solar generators by providing access to flexible generators and capitalizing on the reduction of variability as the geographic scope considered increases [31]. By eliminating congestion in the network, transmission capacity investments can also help mitigate market power in electricity markets [65].

More recently, it has also been explicitly recognized that transmission may be constructed for a third reason, the meeting of public policy goals. These public policy lines may not be justified under either the economic or reliability criteria above (for example, see Section X in FERC Order 1000 [25]). The dominant public policy goal considered in today's transmission expansion planning studies is the inclusion of renewable resources.

⁴ Assume an \$60/MWh for coal plant generation

Because these resources are location-constrained, they cannot be effectively accessed without new transmission investments.

1.1.2 Timescales for Transmission Expansion Planning

Transmission expansion planning takes place on multiple timescales ranging from 5 to 30 years. These timescales, roughly outlined in Table 1-1, have different foci and require varying degrees of technical accuracy.

Table 1-1 Timescales for Transmission Expansion Planning

	(Very) Near Term	Mid-Term (Tactical)	Long-Term (Strategic)	Very Long Term
Horizon	0-5 years	5-10 years	15-20 years	25-30 years
Major Focus	Reliability	Reliability and Economics	Economics and Scenario Analysis	Scenario Analysis
Load Flow Type	AC Load Flow	DC Load Flow	DC Load Flow	Transportation Load Flow
Representative Models	Dynamic Simulation	Optimal Transmission Expansion Planning	Optimal Transmission Expansion Planning	Optimal Transmission Expansion Planning
	AC Optimal Power Flow	DC Optimal Power Flow	DC Optimal Power Flow	Pipes and Bubbles

The shortest time scale considers investments to be made in the next five years. Economic issues may be considered, but the major focus of this timescale is reliability. In order to capture near-term reliability issues, the types of models used have very high technical fidelity. The investments considered on this timescale are more likely to be transformer upgrades, reconductoring, or other projects related to existing transmission lines rather than new transmission lines in new rights of way. Realistically, this is because it is unlikely that a new line could be permitted and constructed within a five year period.

Mid-term or tactical planning includes both reliability and economic foci. Looking five to ten years out, the investments identified in this horizon may present longer-term

fixes to the problems identified in the near-term planning or address structural rather than temporary economic and reliability issues. These investments may be identified using expert judgment or an optimal Transmission Network Expansion Planning model (TNEP). Different types of optimal TNEP models will be explored in 1.2, but briefly, the goal of such models is to algorithmically minimize the total system cost trading off generation, non-served energy and transmission investment costs. The models used to identify and evaluate transmission solutions in mid-term planning are of lower technical fidelity than those in the near-term horizon. These models typically use an approximation of the AC load flow called the DC load flow. The DC load flow simplifies the AC load flow by assuming the magnitude of the voltage sine waves (shown in Figure 1-11) at each node remains constant. The most common form of the DC load flow is also linearized, allowing it to be integrated into traditional linear or mixed-integer linear optimization models. Mid-term planning is designed to identify new transmission lines (either green-field development or re-enforcing an existing transmission line) that would then iterate with the more technical reliability analyses, siting, and cost allocation procedures identified in Figure 1-10.

Long-term or strategic planning is used to guide mid-term planning toward higher-value transmission solutions and for scenario analysis. The most common use is scenario analysis; for example, most high renewable penetration studies are strategic planning studies. Like mid-term planning, the models used in strategic-planning studies are based on DC load flows and both expert analysis and optimal transmission expansion algorithms are used. The models used for long-term studies are also most likely to be abstractions of the real system with aggregated generation and demand buses. Unlike mid-term planning, the transmission lines identified in a long-term planning study are unlikely to be formally considered as potential investments.

Finally, the furthest time horizon typically considered for transmission planning is 25-30 years. On this timescale, the power system is very abstracted. If power flows are modeled in a system, they are often assumed to be directable, simplifying the physics driven flows but continuing to respect capacity restraints of the lines. More often, sys-

tems are aggregated to a pipes and bubbles model. In this approach, large areas are aggregated to a single load/generator, or bubble, and the transmission lines between areas aggregated into a single transfer capacity, or pipe. For example, a pipes and bubble approach might consider each country in Europe as a single bubble and connect a single pipe to each adjacent country. This type of modeling examines how much transfer capacity might be useful between regions, but does not have the technical detail to identify specific transmission investments.

The work presented in this dissertation focuses on the mid-term and long-term planning horizons. The emphasis is on using information from the long-term planning horizon to inform mid-term planning investments. With this, the planning emphasis is also on the major economics tradeoff rather than operational reliability.

1.2 Optimal Transmission Network Expansion Planning Model Formulations

Optimal TNEP models algorithmically balances the investment costs of new lines against reductions in generation and non-served energy costs. Models for the optimal TNEP problem are used to guide system planners toward promising investment choices which are then subject to further operational reliability analysis. The balance of costs is achieved by minimizing annualized systems costs, the sum of annualized investment, non-served energy and generation costs. From an optimization perspective, single-horizon transmission planning falls into one of the hardest problem classes. Transmission network expansion planning is a network problem with non-directable flows and integer investments. It is a combinatorial optimization problem and classified as an NP-Complete problem for which no solution method exists in polynomial time. The number of potential plans for a single investment horizon grows exponentially with respect to the number of potential investments (2^n where n is the number of investment options). A very small problem with only 10 possible transmission investments has 1,024 different plans possible, while a problem with 100 possible transmission investments has over 10^{30} plans possible. A more realistically sized problem with 1,000 possible investments has over 10^{300} possible plans.

The full transmission expansion planning problem is a mixed-integer non-linear problem (MINLP). The integer constraints arise from the transmission investment variables which represent the lumpy investments. The non-linearity arises from representing the physics-driven flows of electricity within the system. The complete DC load flow formulation of the MINLP formulation is given equations Eq. 1 through Eq. 8. The objective function, given in Eq. 1, reflects the tradeoffs typical to mid-term and long-term planning horizons. In Eq. 1, the total cost of the system including new transmission costs, generation costs and non-served energy is minimized. Kirchoff's first law, that demand at each bus must be equal to the sum of flow in to the bus, plus generation at the bus, and non-served energy at the bus minus flow out of the bus is given in Eq. 2. The non-linearities of the optimal transmission expansion problem are shown in Eq. 3 and Eq. 4 which represent flows on the networks. The first non-linearity results from the fact that flows on transmission lines are inversely proportional to physical properties of the line, captured in the reactance of the line X_k and directly proportional to the sine of the difference in voltage angles between buses. The second non-linearity is given in Eq. 4, which describes flow equation for new lines and the non-linearity results from the multiplication of the integer investment variable x_k by the flow.

$$\text{Eq. 1} \quad \text{Min } ct \cdot x_k + \sum_{h=1}^H \left(\sum_{n=1}^N (cg \cdot g_{n,h}) + \sum_i^I (c\mu \cdot \mu_{i,h}) \right)$$

s.t.

$$\text{Eq. 2} \quad d_{i,h} = \mu_{i,h} + \sum_{k \in \Omega_i} f_{k,h} + \sum_{n \in \sigma_i} g_{n,h} \quad i=1, \dots, I$$

$$\text{Eq. 3} \quad f_{k,h} - 1/X_k \sin(\theta_{i(k),h} - \theta_{j(k),h}) \cdot b = 0 \quad k \in K^0$$

$$\text{Eq. 4} \quad f_{k,h} - 1/X_k \sin(\theta_{i(k),h} - \theta_{j(k),h}) \cdot b \cdot x_k = 0 \quad k \in K^+$$

$$\text{Eq. 5} \quad -f_k^{max} \cdot x_k \leq f_{k,h} \leq f_k^{max} \cdot x_k \quad k \in K^+$$

$$\text{Eq. 6} \quad d_{i,h} \geq \mu_{i,h} \quad i=1, \dots, I$$

$$\text{Eq. 7} \quad g_{n,h} \leq g_{n,h}^{\max} \quad n=1, \dots, N$$

$$\text{Eq. 8} \quad f_{k,h} \leq f_k^{\max} \quad k=1, \dots, K^0$$

Indices and Sets:

i	index of buses
k	index of circuits
$i(k), j(k)$	index of terminal buses of circuit k
h	index of load hours
n	index of generators
K_+	set of candidate circuits
K^0	set of existing circuits
Ω_i	set of all circuits connected to bus i
σ_i	set of all generators located at bus i
H	number of load hours
I	number of buses
M	number of candidate circuits
N	number of generators

Parameters/Constants

c_g	generator costs [\$/MW]
c_t	annualized cost of candidate circuits [\$]
c_μ	cost of non-served energy [\$/MW]
d	bus demands [MW]
f^{\max}	circuit capacities [MW]
g^{\max}	generator capacities [MW]
X	circuit reactances [pu]
b	per unit base

Free Variables

$\theta_{i,h}$ bus voltage angle

Positive Variables

$f_{k,h}$ circuit flow [MW]

$\theta_{k,h}$ positive bus voltage angle component across circuit k

$g_{n,h}$ generator output [MW]

$\mu_{i,h}$ non-served energy [MW]

Binary Variables

x_k investment variable

The most common implementation of the TNEP problem is a linearized formulation. The linearized formulation of the problem allows for the use of well characterized traditional optimization routines for linear and mixed integer linear programs (LPs and MILPs). In order to linearize the MINLP, Eq. 3 and Eq. 4 are simplified to Eq. 9 and Eq. 10 by assuming small angular differences in the network. When angular differences are small, the sine of an angle may be approximated the angle itself (in radians). The non-linearity left in Eq. 10 can be linearized for traditional optimization methods through the use of the disjunctive formulation such as in [8]. Other optimization methods, such as meta-heuristics, are blind to the non-linearity in Eq. 10 as they do not rely on linear relaxations of the full problem to establish bounds on the objective function for the algorithm to proceed.

$$\mathbf{Eq. 9} \quad f_{k,h} - 1/X_k (\theta_{i(k),h} - \theta_{j(k),h}) \cdot b = 0$$

$$\mathbf{Eq. 10} \quad f_{k,h} - 1/X_k (\theta_{i(k),h} - \theta_{j(k),h}) \cdot b \cdot x_k = 0$$

For clarity, the formulation presented thus far is the most basic expression of the problem. It does not include losses or the $n-1$ contingency considerations discussed above or any other reliability constraints. There are, however, linear representations of the

losses that may be integrated into MILP formulations of the optimal transmission expansion planning problem (see for example [1] [42] or [59]). Likewise, the $n-1$ contingency constraint has been included in transmission expansion planning models, but it is generally considered in an iterative fashion where the optimal transmission expansion planning model determines a test plan, $n-1$ reliability is examined, and new constraints are passed back to the optimal transmission expansion planning model (see for example [50] or [53]).

The most common optimal TNEP formulations are both deterministic and static. Deterministic models do not consider the many uncertainties that face power system planners, including uncertainty across fuel prices, policies affecting the power system and new generation locations. As deterministic models simplify the uncertainty facing planners, static models simplify the time horizons planners consider. Although new transmission lines have life spans of more than 40 years, static models consider only a single future year. This static-deterministic model formulation is the one presented above in Eq. 1 through Eq. 8.

Due to the computational complexity of the problem, the static-deterministic formulation is the most studied in the academic literature. Beginning with LL Garver's foundational planning paper in 1970, researchers have applied a variety of approaches to solve the problem, including linear optimization [26], dynamic programming [18], decomposition techniques [8], engineering heuristic guided searches [32][59] and meta-heuristics [54] [58][71]. The dominant approach for transmission expansion planning in the late 1990s to early 2000s used traditional optimization techniques for MILPs. Following the thrust on traditional optimization methods in the early 2000s has been an emphasis on meta-heuristic methods. A variety of meta-heuristic models ranging from genetic algorithms [58] to simulated annealing [54] have been applied to the transmission expansion planning, though no one meta-heuristic approach has been found to be most effective.

While the static-deterministic problem is the most studied formulation, the problem considered in this dissertation requires multi-stage stochastic modeling. The effects and

current literature of multi-stage transmission planning modeling are discussed in 1.2.1 and stochastic planning in 1.2.2. Finally, the combination of stochastic and multi-stage modeling are considered in 1.2.3.

1.2.1 Multi-Stage Planning

Transmission lines are capially intensive investments with economic lifespans of 40 or more years. Over the lifespan of these investments, the power system will continue to evolve. Existing generators will retire, and new generators will be added. Loads will increase and decrease with the economy and new development patterns. Fuel prices will fluctuate as will national policies and environmental regulations. All of these changes will affect the value of a transmission plan. A transmission expansion model, however, cannot capture all of these varying timescales. Instead, modelers simplify the number of time horizons to capture the most important details.

The most common simplification made in TNEP models is to consider only a single investment decision stage. With this simplification, modelers plan for a specific target year. Traditionally, plans have been constructed for mid-term horizons (5-10 years) or long-term horizons (15-20 years). The static case is the one most commonly studied in the literature (see above citations) and is also the most common in industry and policy studies. The most prominent renewable integration/transmission studies are examples of long-term static studies [17][42][43]. A less common approach is to consider a series of static time horizons. As shown in Figure 1-12, the static sequential approach plans for the first time horizon, carries over the new investments from the first stage as constraints to the second stage, plans for the second stage, etc. The WECC TEPPC planning approach using a 10 year planning process and integrating those lines as constraints into the 20 year horizon is an example of the static sequential approach [68].

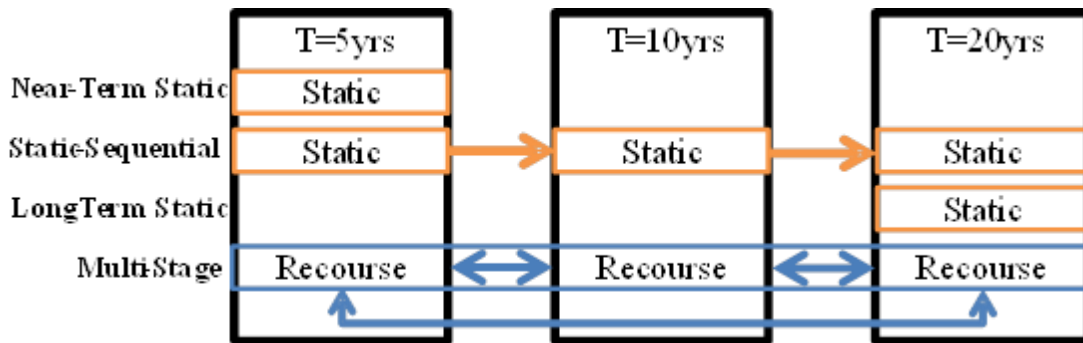


Figure 1-12 Comparison of Static, Static Sequential and Multi-Stage Modeling

The static horizon simplifications are problematic because either economies of scale are not captured in the modeling or pent up demand for new transmission capacity overstates the desirability of large lines. As an example of the economies of scale in the transmission system, on a thermal-capacity basis, a 345kV double circuit line (tower with two sets of conductors) costs \$1,333/MW-mile while a 765kV single circuit line (tower with one set of conductors) costs 70% less, \$413/MW-mile⁵. If a mid-term horizon is used, there is a limited time for demands to grow, and the need for larger lines is never recognized. On the other hand, with a long-term horizon and the assumption of no new transmission over 20-30 years, there is a pent-up demand for new transmission and large lines are almost exclusively selected.

This concept is illustrated in a simple three bus model in Figure 1-13. In this model, a single load exists at Bus A with generators at Buses B and C. There are also two investment options: a 750 MW line from Bus A to Bus B for an annualized cost of 5 million USD (MUSD) or a 1,500 MW line in the same corridor for 7.25 MUSD annually. In the traditional myopic formulation, only the load in year 10 would be considered. In this case, only 400 MW of transmission capacity is required to meet the load and the least cost investment is the smaller 750 MW line. If, however, a second time horizon, 25 years, is considered the demand has grown to 1,500 MW and now an additional 900 MW of capacity is required to meet the demand at Bus A. The lowest-cost option is now the

⁵ Assumes AEP Thermal Limits and Costs, [3]

more expensive but larger line. To include the effects of multi-stage modeling in optimal transmission expansion planning models, the objective function explicitly includes costs on multiple time horizons and inter-temporal investment constraints are added. An example of an objective function considering two time-horizons is given in Eq. 11. In Eq. 11, the costs are discounted by a factor, α_y , and summed across each annual time horizon, $y \dots Y$.

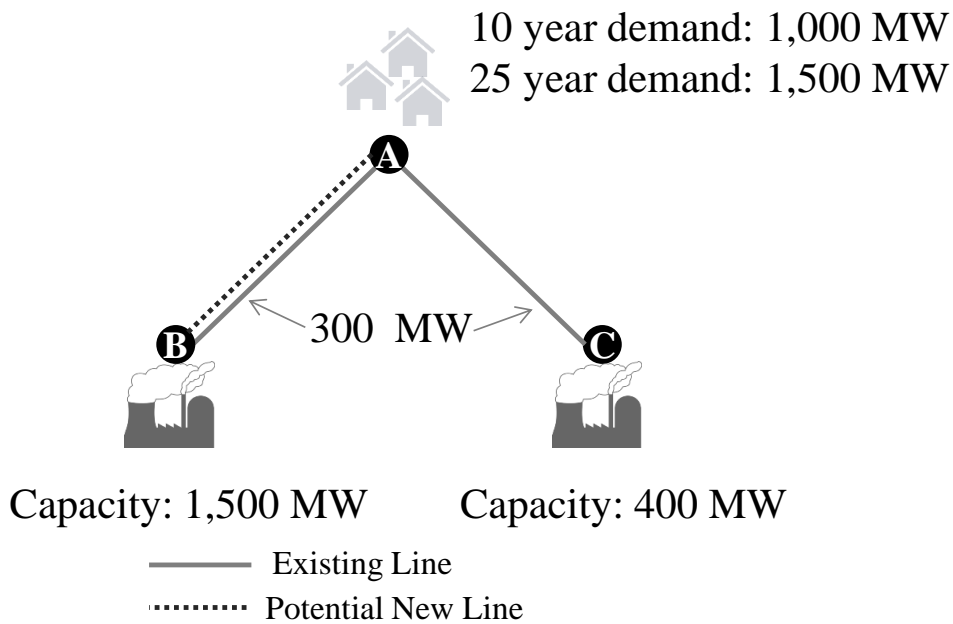


Figure 1-13 Test System for Time Horizon Simplification

Test system assumes cost and capacity characteristics approximating a 345kV double circuit line and 765kV single circuit line at 100 miles length.

$$\text{Eq. 11} \quad \text{Min} \sum_{y=1}^Y \left[ct \cdot x_{k,y} + \sum_{h=1}^H \left(\sum_{n=1}^N (cg \cdot g_{n,h,y}) + \sum_i^I (c\mu \cdot \mu_{i,h,y}) \right) \right] \cdot \alpha_y$$

These effects have also been explored in a case study on the Spanish transmission network. The case study considered a reduced-order system with 701 existing 400kV and 700kV lines. Five five-year periods were considered between 2015 and 2035. First, each

time period was planned statically (individually) with 279 candidate lines considering three different capacities assuming 1% load growth annually. As seen in Figure 1-14, there are five lines constructed in the 2015 plan that are no longer required in the 2035 plan. The planning exercise was then rerun using a more limited set of 16 potential transmission investments and two time horizons (2015 and 2035) to directly compare a multi-stage approach to the static-sequential results. As shown in Figure 1-16, the static and multi-period models produce different first stage decisions in 2015. The static approach selects seven first stage investments (four 400 MW lines, 0 750 MW lines and three 1500 MW lines) while the multi-stage model selects only five lines (one 400 MW line, one 750 MW line and three 1500 MW lines). This case study clearly illustrated that even in a small system, the static sequential approach does not mirror the results produced by the true multi-stage model.

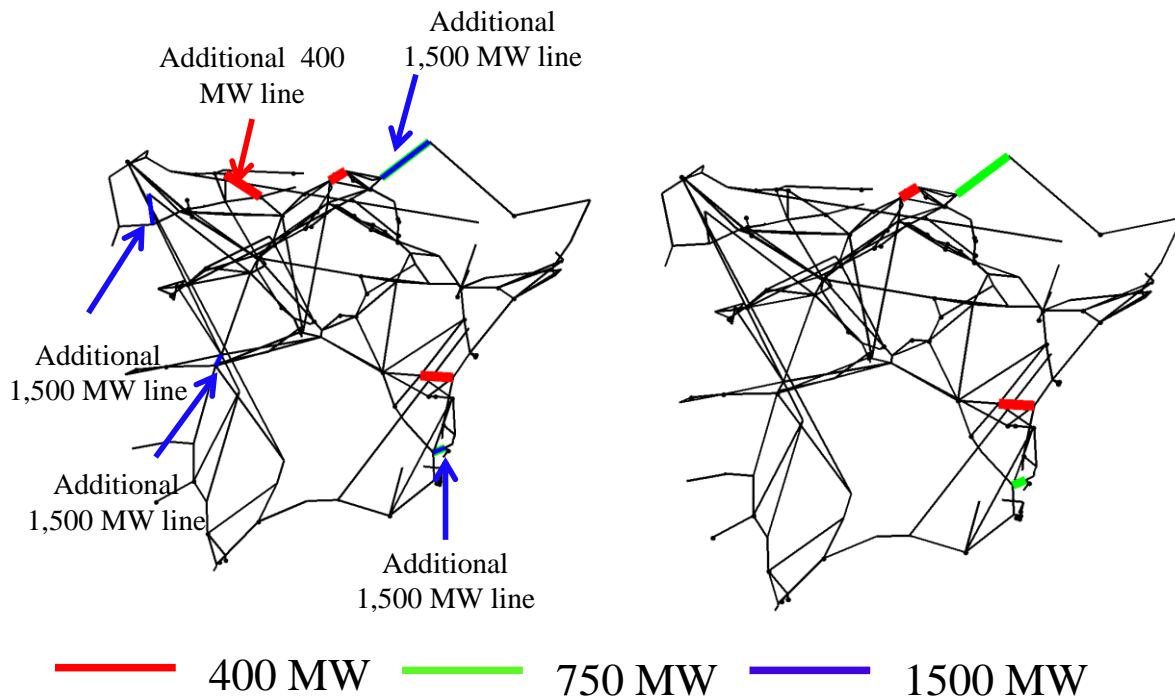


Figure 1-14 Spanish Case Study Static Results for 2015 (L) and 2035 (R)

Note that only lines which are not consistent across one or more time horizons are shown.

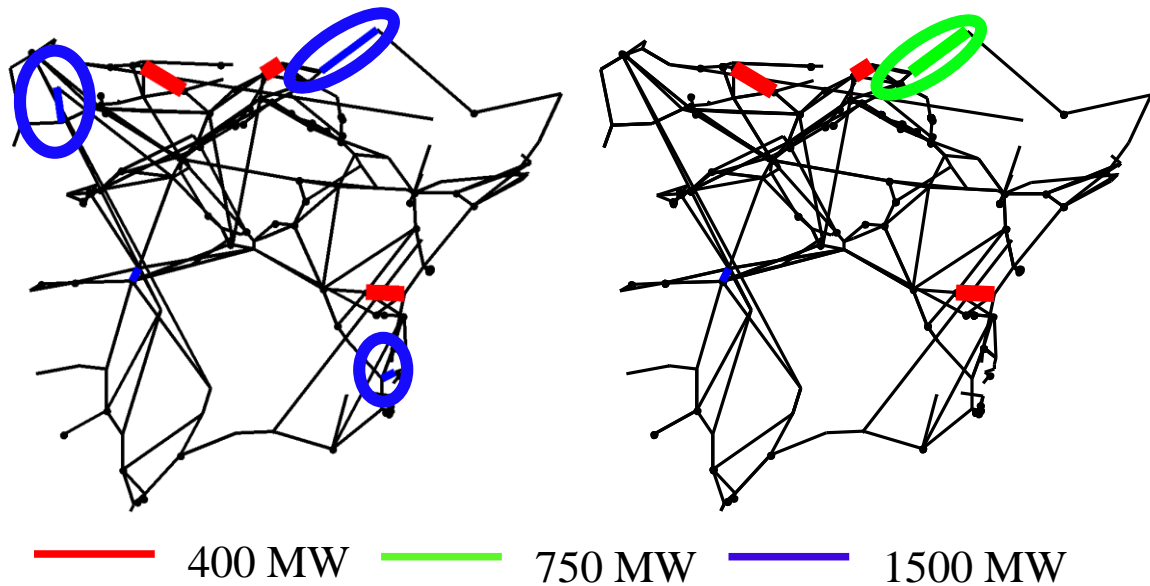


Figure 1-15 Spanish 2015 Case Study Dynamic Programming Results (L) and Static Planning (R)

This type of multi-stage modeling is relatively rare in the academic transmission expansion planning literature. Examples of the relevant literature are given [5], [9],[64], and in [35]. Common to these works is an emphasis on investment timing rather than capturing effects of economies of scale or rectifying issues with myopic planning. This may in part be due to the small test systems with limited investment options used in the academic studies. For example, [9] considers 23 investments, of which only 13 are transmission lines, across a system with 157 buses. By comparison, there are more than 2,000 buses in the electricity system in the western United States over 200kV. Each of these works is also focused on the demonstration of a new algorithm rather than an analysis of the results. The first three works demonstrate heuristic algorithms while [5] explores a branch-and-bound algorithm with a transportation-based transmission work (respects only Kirchoff's first law).

One reason that multi-stage models are not as popular in the literature is the computational complexity added by the intertemporal constraints. The addition of multiple time stages again increases the size of the optimization problem. The growth of the number of

possible plans that must be considered is demonstrated in a two-horizon expansion planning problem with two possible investments. As shown in Figure 1-16, if only the first stage problem is considered, there are four possible transmission expansion plans to evaluate. On the other hand, when two stages are considered, there are now nine possible two-stage transmission plans to consider.

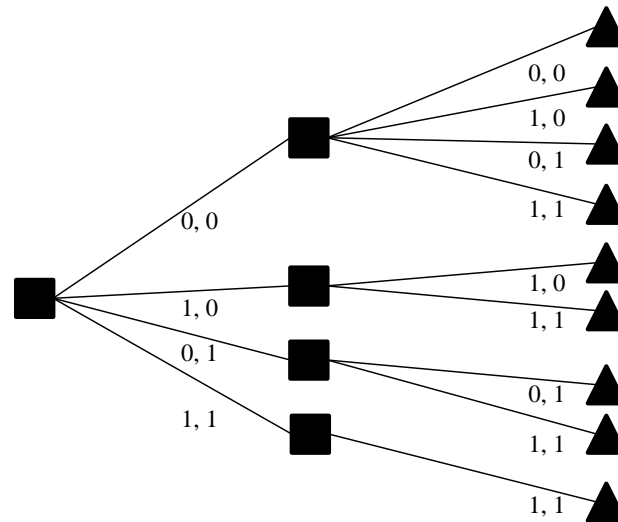


Figure 1-16 Expansion of the Decision Space Including Multiple Time Horizons

While true multi-stage modeling is not common to industry methods, heuristic methods have been used to span time horizons. For example, in the Southwest Power Pool's Integrated Transmission Planning, three time horizons are considered (near term, 10 years and 20 years) [61]. While the three timescales are never formally integrated, investments identified in longer terms studies are highlighted in the near term studies. Likewise, Red Electrica in Spain uses a long term planning models in to inform medium-term transmission planning [16].

1.2.2 Planning Under Uncertainty

Almost all industry and academic transmission expansion planning models consider specific scenarios with perfect knowledge. That is, the modeler assumes that all future demands, fuel prices, generator locations and reliability issues are known. In reality, of

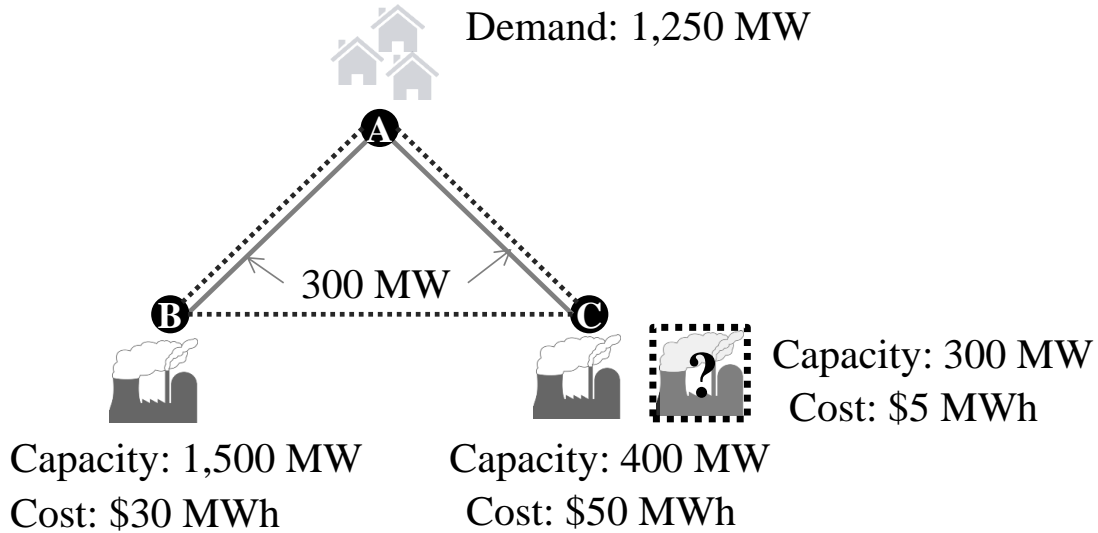
course, planners have very imperfect knowledge about the future. Natural disasters destroy infrastructure and cause major changes in fuel prices. Technological breakthroughs produce new generation types not even known to today's planners. In order to capture some of this uncertainty, planning models can try to produce transmission plans with low costs across a variety of different futures.

Including uncertainty in the objective function affects the operational costs. As shown in Eq. 12, the investment decisions made are not based on the individual scenario, s , but the operational costs all become scenario dependent. Each scenario is also weighted by its probability, p_s . There are both random and non-random sources of uncertainty to be considered in the transmission expansion planning problem. Random uncertainty, such as the outage of a component, the evolution of fuel prices, and demand growth can be modeled probabilistically based on future projections of historic data. Non-random uncertainties such as the location of new generation, regulatory changes, and the development of new technologies, however, are do not historic datasets on which to draw.

$$\text{Eq. 12} \quad \text{Min } ct \cdot x_k + \sum_{s=1}^S p_s \left[\sum_{h=1}^H \left(\sum_{n=1}^N (cg \cdot g_{n,h,s}) + \sum_i^I (c\mu \cdot \mu_{i,h,s}) \right) \right]$$

As an illustration of uncertainty's effects on transmission expansion in a small example, take the three bus example now shown in Figure 1-17. In this iteration of the model, a new low-cost power plant may be constructed at Bus C and three different investment options are presented by the planner. The total system costs assuming each transmission and generation scenario are given in Table 1-2. If the planner had perfect foresight and knew that the plant would not be built, the lowest cost investment option would be to build Lines A-B and B-C. Note that because the system is networked, the most intuitive solution, building Line A-B only, is not the lowest cost option in any scenario. Instead, building the additional line between Buses B and C allows the lower cost plant to meet the entire load at a lower cost by sending power across both the new lines and the existing line between Bus A and Bus C. On the other hand, if the planner had perfect foresight and knew the new generator would be built, the lowest cost option is to build only the additional capacity between Bus A and Bus C. If the planner assumes

the generation is built and only adds only Line A-B, but the generation developer pulls out, the system becomes very expensive, nearly a factor of five more, to operate.



	Cost (MUSD)	Capacity (MW)
Line A-B	5.0	750
Line A-C	5.0	750
Line B-C	3.0	500

Figure 1-17 Test System for Uncertainty Simplification

Table 1-2 Scenario Costs for Uncertainty Test System

System Costs	New Generator (MUSD)	No New Generator (MUSD)	Expected Value (MUSD)
Line A-B	268	369	318
Line A-C	104	500	302
Lines A-B, B-C	161	336	248

The most common place to find uncertainty considered in for transmission expansion planning is in the area of planning for reliability. As a field, probabilistic transmission expansion planning focuses on random uncertainties, specifically failure of components [14],[34]. With this approach, the uncertainty is treated as a constraint (eg the system must survive with at least a set probability) [14]. This framing with uncertainty as a

constraint is not appropriate considering uncertainties affecting operational costs. The reliability tests are binary, either the system fails or survives. With operational costs, there is not a failure metric but rather a distribution of operational costs. Plans can be differentiated by the expected value of these costs or other decision metrics of interest such as minimum regret.

Planning across uncertainties again increases the size of the planning problem. As a result, the approach used by most of the academic planning literature does not truly optimize across uncertainties. Instead, the general strategy is to generate a variety of plans by optimizing deterministic scenarios. The plan is then evaluated under different scenarios, and the plan with the best decision metric is selected. For example, Sozer generated scenarios using Monte Carlo sampling and then selected an optimal plan according a multiple comparisons to the base (MCB) metric [62]. Zhao also optimized individual scenarios to create transmission expansion plans and then tested the reliability of each transmission plan under the other scenarios [72]. Likewise, Bustamente-Cedeno and Arora used scenarios to generate potential transmission expansion plans [11]. The optimal transmission expansion plans was then selected from amongst the set using expected value and minimum regret decision metrics. These approaches reduce the computation burden of solving the stochastic problem, but likely miss the plan which performs best across uncertainty as demonstrated by [41].

Each hour of the year modeled in TNEP may also be framed as an uncertainty. For computational reasons, the full 8,760 hours each year are rarely modeled in an optimization problem. Instead, a selected number of hours (for example those representative of peak, intermediate and baseload demand) are modeled. The duration of each representative hour may be reconsidered as the probability for that load profile and the goal of the optimization is to produce the highest expected value. This is the framing used in [31] though not explicitly modeled in a stochastic framework.

1.2.3 Stochastic Multi-Stage Planning with Recourse

Combining multi-stage modeling and stochastic modeling explicitly models the full optimization problem of transmission expansion planning under uncertainty. The problem

now explores four major sets of tradeoffs summarized in Table 1-3. First, the optimization considers the balance of current and future costs. Second, the optimization considers the tradeoff between economies of scale and current costs. Larger lines capitalize on economies of scale, but they are also more expensive to build in the near term. By combining stochasticity and multi-stage modeling, the tradeoff between economies of scale and adaptability is also included. If large lines are constructed to capitalize on economies of scale but those interconnected areas experience less than predicted generation expansion, the investments may be underutilized. On the other hand, under-building and under-estimating generation may result in congestion, underutilization of the new generation and higher operational costs. Finally, the optimization respects the fundamental tradeoff between investment and operational costs.

Table 1-3 Summary of Tradeoffs Considered in Stochastic Multi-Stage Models

Tradeoffs		
Current Costs	vs	Future Costs
Current Costs	vs	Economies of Scale
Economies of Scale	vs	Adaptability
Investment Cost	vs	Operational Costs

By combining multi-stage and stochastic modeling, the structure of the objective function changes. An example of an objective function for a single investment with two decision-stages and two uncertainty-stages model is given in Eq. 13 and a graphic illustration of the problem in Figure 1-18. In Eq. 13, the cost to minimize is the cost of first stage transmission investments, the expected first-stage operational costs, OC_1 , and the discounted sum of second stage investment costs and second-stage expected operational costs, OC_2 . As shown in Figure 1-18, for a single investment the tree grows to 12 unique combinations of first stage investments, second stage investments, first stage uncertainty and second stage outcomes.

Eq. 13 $\text{Min } ct \cdot x_{k,1} + E[OC_1(x_1)] + \alpha_y(ct \cdot (x_2|x_1S_1) + E[OC_2(x_1, x_2, S_1)])$

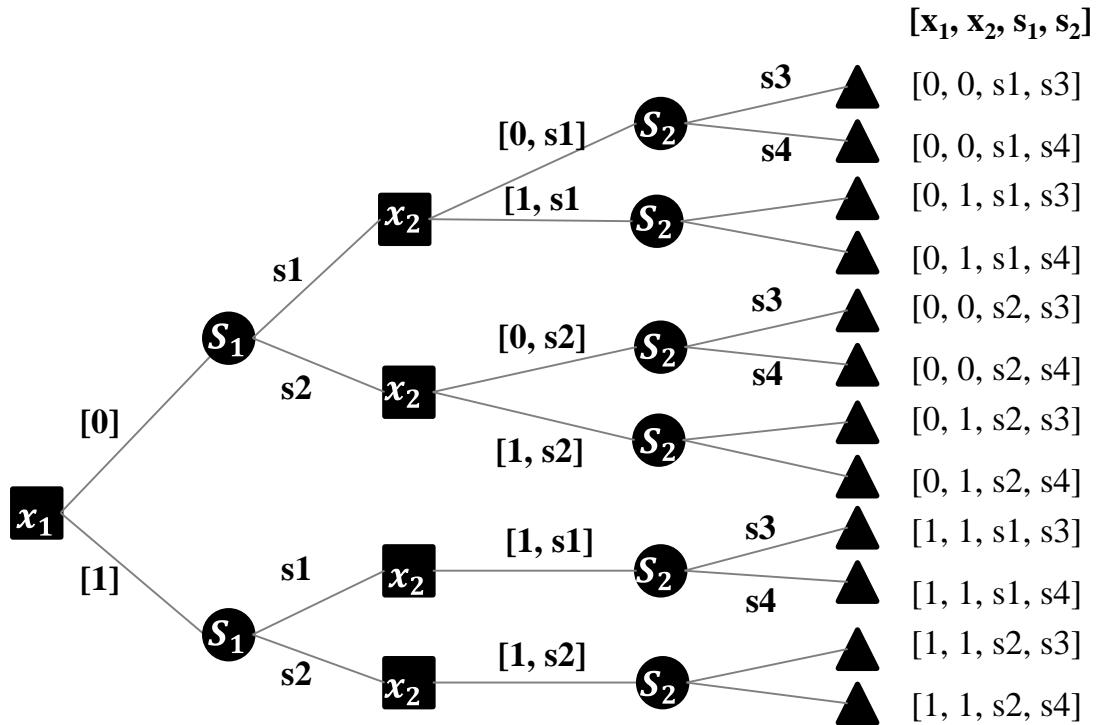


Figure 1-18 Single Investment, Two Decision-Stage, Two-Uncertainty Stage Problem Illustration

After each decision, x , the current system state is shown in brackets. NB, once the single line has been selected for investment, no further transmission investment options exist.

Important to note in Eq. 13 is that the second stage decision is contingent both upon the first stage decision and the outcome of the first stage uncertainty. This ability to reconsider investment decisions after the uncertainty has been revealed is referred to as recourse. Access to recourse is what differentiates this stochastic multi-stage modeling from the deterministic multi-stage modeling where a single trajectory for all time is selected. In this modeling context, the most important outcomes are the stage one decision investments. After these investments are made, the uncertainty is revealed, and the process is rerun given the new starting state of the world. Again, although only the

first stage decisions would be considered for investment, the additional decision stages provide foresight into possible future system developments and influence the first stage decisions.

While not explicitly discussed until this point, the location of new generation resources is not independent of new transmission investments. By design, new transmission lines increase access to areas with high electric prices, making areas serviced by the new transmission more desirable for generation expansion. This is especially true in the case of location-constrained generators which rely on transmission access to move their power from remote areas. As a result, the uncertainty shown in Figure 1-18 should be represented endogenously with a relationship between the transmission investments made and new generation investments as given in Eq. 14. In Eq. 14, the probability of a generation expansion scenario, p_s , is dependent on the first stage decision, x_1 . Likewise the probability of the second stage scenario is dependent on both the first and second stage investments as well as the uncertainty outcome from the first stage.

$$\begin{aligned}
 \text{Eq. 14} \quad & \text{Min } ct \cdot x_{k,1} + \sum_{s=1}^S [(p_s | x_{k,1}) \cdot OC(s | x_1)] + \\
 & + \alpha_y \left(ct \cdot (x_{k,2} | x_{k,1}, S_1) \right. \\
 & \left. + \sum_{s=1}^S [(p_s | x_{k,1}, x_{k,2}, S_1) \cdot OC(s | x_1, x_2, S_1)] \right)
 \end{aligned}$$

Exceptionally few academic works have explored multi-stage stochastic planning with recourse. On a line-by-line basis, stochastic multi-stage modeling has been explored using a real-options framing [10],[29]; however, when considering network planning, individual decisions (eg a single transmission line) cannot be evaluated individually due to network effects. For example, in Figure 1-19 the value of a line from Bus A to Bus B is contingent on the construction from a line from Bus B to Bus C. Munoz, Hobbs and Kasina have explored the approach using a Benders Decomposition approach for a two stage problem with three uncertainty scenarios [40],[41]. In their work, both transmission

and generation are co-planned in the first stage, and the uncertainty comes from full scenarios (demand growth, generation mix requirements, etc). When the scenario uncertainty is realized, the second stage offers access to recourse through both generation and transmission investments. In [40], the size of the problem is restricted by considering only three uncertainty scenarios and a single type of new investment in existing corridors.

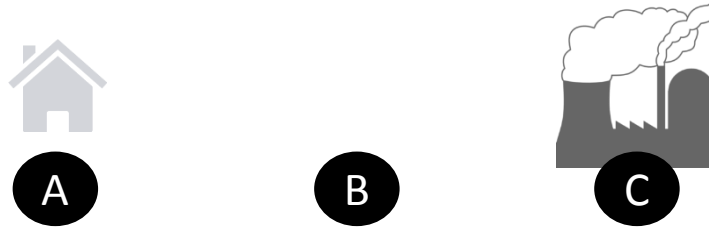


Figure 1-19 Network Effects Illustration

1.3 Dissertation Overview

A paradigm shift has taken place in transmission expansion planning. Rather than traditional questions of regional-system reinforcement, planners today must design network architectures for areas the size of continental Europe. The design of these new networks will determine the efficacy of expanding electricity markets and the ability to integrate high penetrations of renewable and other location-constrained generators. Planning these networks requires balancing the uncertainty in the evolution of the power system and the significant economies of scale inherent to transmission investments.

This new planning paradigm poses fundamental challenges to existing tools. These tools, designed for reinforcement problems, were developed to consider tens of investments and tens of nodes. The new paradigm, however, requires the analysis of thousands of investments and thousands of nodes. This change in dimensionality, before considering both the stochasticity and multi-stage aspects, overwhelms the capability of existing methods. The complexity of the problem, however, demands decision support tools to assist planners and policy makers.

Tackling the dimensionality of the wide-area multi-stage stochastic transmission expansion planning problem will require not a single model, but rather a comprehensive

approach. This work proposes one such suite of models. Chapter 2 focuses on directly reducing the dimensionality of the stochastic multi-stage transmission expansion problem through the development of a screening model. This work presents a method to reduce the number of investments considered in a wide-area transmission plan by greater than 90%. Chapters 3 and 4 propose novel heuristic approaches to solving the multi-stage stochastic problem. Chapter 3 proposes interpolation techniques based on image processing techniques and Chapter 4 embeds these interpolation techniques in an approximate dynamic programming framework. Finally, Chapter 5 presents broader conclusions and future work prompted by this research.

THIS PAGE INTENTIONALLY BLANK

2. St. Clair Screening Model: Dimensionality Reduction through Algorithmic Screening

The primary goal for transmission expansion planning is to decide which investments to make today. The investments considered for the mid-term planning horizon are new transmission lines, and each line is specified by its end nodes and rated voltage. When planning for regional power systems, local planners used expert judgment to select small numbers of potential investments. The number of possible investments in the interconnection and continental scale systems, however, is exponentially greater than the smaller regional systems. Due to this increased number of investments, uncertainty in future power system development and the necessity of multi-stage modeling, all lines may also not be considered in an optimization. For example, a system with 1,000 nodes has 500,000 unique connections between node pairs; the path between each node-pair or corridor, may house different types of investments. For example, a corridor between two nodes may contain a combination of 345kV, 500kV or 765kV rated transmission lines; if only a single circuit at each of the three voltage ratings is considered, a 1,000 bus system has more than 1.5 million potential investments to consider.

The number of transmission investments is a key driver for the complexity of the multi-stage stochastic transmission network expansion planning (MS-TNEP) problem. The number of investments exponentially (2^n) increases the computational size of the problem while the number of uncertainty scenarios increases the size of the problem polynomially. Thus reducing the number of transmission investments is an effective way to reduce the computational size of the problem. Despite the larger number of potential transmission expansion lines, few are selected for an individual plan; of the 1.5 million investments in the 1,000 node example, an expansion would contain only 10-100 lines or less than 0.01% of all possible lines. Many of the possible investments would never be constructed, for example a 1,000 mile 345kV transmission line would not be selected for both physical and economic reasons.

The most effective way to shrink the number of investments for planners to consider would be to solve the full stochastic multi-stage problem. This full problem, however, is

not computationally feasible and is simplified for the screening model. To identify likely transmission investment options and eliminate those which are not used under any scenario, this chapter proposes and demonstrates a novel screening model, the St. Clair Screening Model. An overview of the formulation and its implications of the St. Clair Screening Model are discussed in 2.1. Once formulated, the screening model is demonstrated on a test system described in 2.2. The results of the demonstration, including a 97% reduction in the number investments required for consideration, are given in 2.3, and the implications of these results and future work are given in 2.4.

2.1 St. Clair Screening Model Formulation

The goal of the St. Clair Screening Model is to reduce the total number of transmission investments for further consideration by solving a series of simplified problems. As shown in Figure 2-1 the structure of the model can be depicted in three main steps. First, transmission investments are characterized both physically and economically (discussed in 2.1.1.) in preparation for the optimizations run in the second step. Second, an optimization model identifies where new capacity should be added in the system. For a large system with significant uncertainty, the most computationally demanding aspects of the optimization problem are the integer investments and the stochastic modeling. In the St. Clair Screening Model, these two constraints on the problem are relaxed in order to create a model which solves in a reasonable amount of time. The method and implications of the simplifications are discussed in 2.1.2. Finally, in the third step, specific investments are identified to meet the capacity needs identified by the optimization model in step two.

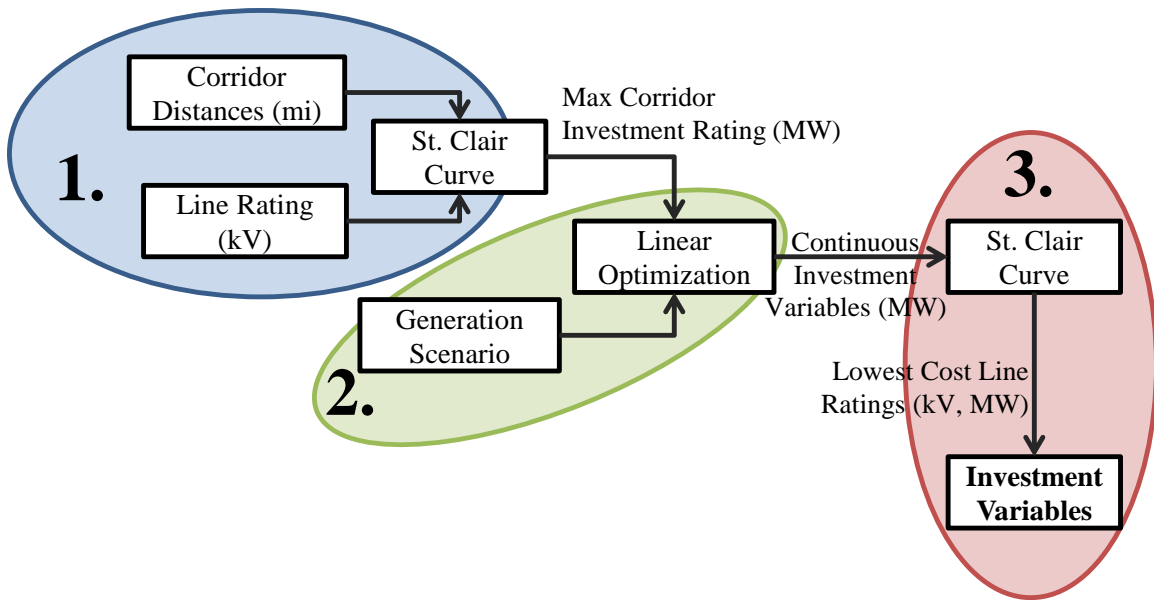


Figure 2-1 Schematic of St. Clair Screening Model

2.1.1 Characterization of Transmission Investments

Transmission lines must be characterized both physically and financially. In a smaller study, each transmission investment would be characterized uniquely; however, with thousands of investment options, this individual characterization would be overly time-consuming. Instead, the St. Clair Screening Model uses engineering heuristics to quickly characterize lines. Once the gross quantity of lines has been screened, individual financial and rating studies would be performed for future studies. These ratings would be updated and modified through the iterative processes with siting and routing discussed in Chapter 1.

Physically, a transmission line can be described by its maximum capacity to transmit power, measured in MW and its impedance, measured in Ohms. The impedance of a transmission line describes its opposition to power flow and is made up of the resistance and reactance. Resistance, typically much smaller than reactance in high voltage transmission lines, is inversely related to the line's conductivity and is a function both of the material properties of the conductors used and the length of the transmission line. The reactance of a transmission lines describes its opposition to alternating current and voltage, as in used in power transmission, and describes to the inductive and capaci-

tive properties of the transmission lines. The reactance of a transmission line is also related to both its material properties and length and can be calculated using literature values as shown in Eq. 15.

$$\mathbf{Eq. 15} \quad X_k = X(l) \cdot l_k$$

$X(l)$: Reactance per unit length

l_k : Length of line k

The capacity of a transmission line is determined by its operating voltage level, reactance, and the operational standards within the power system. One of the more common ratings on a transmission line is its thermal rating. The thermal rating describes the maximum capacity of a transmission line before it overheats. This rating reflects the active constraint for short lines (less than about 50 miles). For medium length (50-200 miles) and long lines (over 200 miles), power flow is constrained to lower quantities by voltage drop and then stability issues. These varying constraints are encapsulated in an engineering heuristic known as a St. Clair curve, first described by H.P. St. Clair in 1953 [63]. The St. Clair curve, shown in Figure 2-2, provides a relationship between the length of a transmission line and its carrying capacity normalized by Surge Impedance Loading (SIL). The SIL of a line is a function of the operating voltage level and the impedance of the transmission line. It represents the operation of the line when the line's capacitance and inductance are balanced, and the line neither supplies nor consumes reactive power.

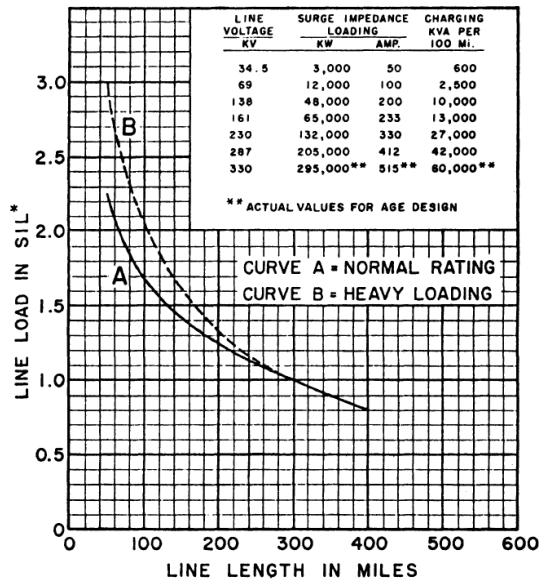


Figure 2-2 St. Clair Curve [63]

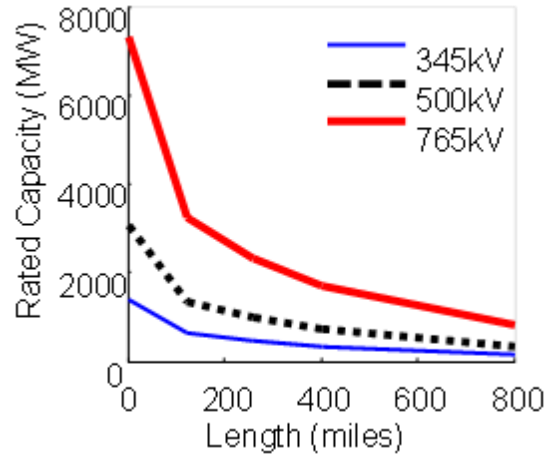


Figure 2-3 Piecewise St. Clair Curve Used
345kV double circuit, 500/765kV single circuit

Lines can be quickly rated for their maximum capacity using the St. Clair curves using literature values for surge impedance loading. The St. Clair curve can be operationalized as a piece-wise linear function, as shown in Figure 2-3; the SIL and thermal rating values used in Figure 2-3 are given in Table 2-1. For each potential investment, which is already specified at a specific voltage level, the length of the transmission line is calculated using the GPS coordinate of each node. The capacity is then read directly from the St. Clair curve. Again, the St. Clair curve provides a useful heuristic to relate the length of a transmission line to its rating; however, these ratings would be revised through additional study once a smaller set of lines is selected for analysis.

Table 2-1 Surge Impedance Loading and Thermal Ratings [3]

	345kV Double Circuit	500kV Single Circuit	765kV Single Circuit
Thermal Rating (MW)	1500	3000	8000
SIL Rating (MW)	400	880	2090

The cost of a new transmission line includes the land (right of way), transmission towers, conductors and labor. Like the reactance, the cost can be linearly scaled as a function of length as shown in Eq. 16. Typical line costs range from \$1.5-3.3 million per mile and vary by voltage rating, local geography and right-of-way costs. Characteristic costs are given in Table 2-2 for common line ratings in the United States. In Table 2-2, two different 345kV configurations are shown, single and double circuit. A single circuit configuration has one set of conductors on a single tower while a double circuit configuration has two sets of conductors on a single tower. While both single and double circuit lines have been constructed, the double circuit configuration is more prevalent and capitalizes on economies of scale.

$$\text{Eq. 16} \quad C_k = C(l, v) \cdot l_k$$

$C(l, v)$: Reactance per voltage rating and unit length

l_k : Length of line k

Table 2-2 Characteristic Transmission Line [3]

Line Type	Cost per mile (Million USD)
345kV single circuit	1.5
345kV double circuit	2
500 kV single circuit	2.9
765 kV single circuit	3.3

2.1.2 Linearization of the TNEP Problem

The first major simplification in the St. Clair Screening model is the relaxation of integer constraints for new investment in the optimization problem. This simplification exponentially decreases the solution time for the optimization model. Physically, this simplification means that rather than investing in a complete transmission line between two nodes, the model selects instead the amount of capacity between two nodes. This capacity can range from 0 MW up to the capacity of the full transmission line. For example, if a corridor was rated to mirror the capacity of a new 765kV transmission line

limited to by its thermal capacity, the model may select any capacity from 0 MW up to 8,000 MW.

This new formulation for the optimization model is given in Eq. 1 to Eq. 24. The relaxation of the investment variables is given in Eq. 4. By linearizing the transmission investment variables, the type of load flow in the model is also transformed from a DC load flow model to a hybrid load flow model. In the hybrid load flow model, flows on existing lines are governed by their reactance and the difference in voltage angles (Eq. 3), and flows on new lines are only constrained by the maximum capacity of the new investment (Eq. 4). This hybrid model has been used previously in transmission expansion planning [57].

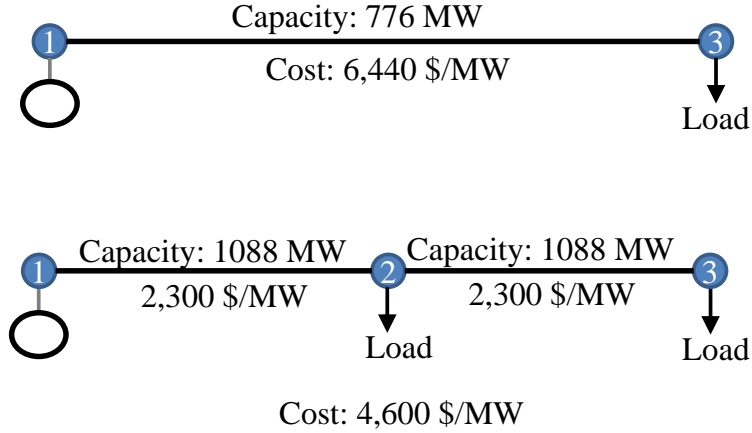
As a result of linearizing the investment variables, each type of line must be considered individually. With linearized transmission investment variables, the investment costs must also be linearized. Rather than the full investment cost for a new line, the transmission planning model considers costs on a cost per MW basis. An example of these linearized costs, which assume that each type of line is limited to its thermal capacity, is given in Table 2-3. On a \$/MW basis, the 765kV line is significantly less expensive than either the 345kV or 500kV transmission as demonstrated in Table 2-3. As a result, if all three line types were considered simultaneously, the linear optimization would always select capacity from the 765kV transmission investment pool first. In order to capture the cost characteristics as well as the capacity characteristics, the model must be run sequentially for each generation scenario, i.e. first with all 345kV capacity limits and costs, second with all 500kV capacity limits and costs and finally with all 765kV capacity limits and costs.

Table 2-3 Linearized Transmission Costs Assuming Thermal Limit Constraint (Non-Annualized Costs)

	345kV Double Circuit	500kV Single Circuit	765kV Single Circuit
Cost (\$/MW-mile)	\$1,333	\$967	\$413

The linearization of investment costs also introduces the potential for short-line bias in the model. The piece-wise linear St. Clair curve, shown in Figure 2-3, is an approximation of a negative exponential function with respect to length. While the capacity decreases dramatically up to 150 miles in length, the cost scales linearly. This mismatch introduces an unintentional short line bias. Take for example, the two test systems in Figure 2-4. In it a single 100 mile line has been characterized both economically and on a capacity basis for a 345kV double circuit investment. In the second system, a third node has been added to the system. Now the St. Clair procedure is used to rate two individual 50 mile segments instead of a single 100 mile transmission line. Each 50 mile segment costs \$2,300/MW (or \$4,600 for both segments) compared to the \$6,440/MW for the 100 mile line. This bias results from the fact that a 50 mile line has significantly higher capacity using the St. Clair methodology than a 100 mile line. Thus the same cost is divided by a higher capacity over a series of small lines. This bias is a fundamental issue when using the St. Clair curve. The St. Clair curve rates a specific segment of a transmission line rather than the whole path. The curve assumes that a short transmission line will be loaded on one end and discharged at the other. As a result, it received a higher rating than if it was explicitly modeled as part of a longer transmission path which would have a higher impedance and potentially voltage stability/angular stability issues.

The bias is mitigated to an extent by power electronics which effectively increase the capacity of short and medium lines in practice. The St. Clair approach assumes that each line is uncompensated, that is, there is no additional voltage or current support. Realistically, however, at each node there may be generation or load with the ability to control voltage and current. This support can effectively lower the reactance of a transmission line and raise its effective power transfer capacity. Because this effect is not captured in St. Clair curves, the St. Clair curve methodology may still underestimate the capacity of short to medium length lines.



Cost of investment normalized over 1,088 rather than 776 MW

Figure 2-4 Distortions in Per Unit Cost Due to St. Clair Assumptions and Linearization

Rather than solve the full stochastic problem, the St. Clair Screening Model solves a series of deterministic multi-stage models. The basic idea is to sample deterministic generation scenarios until no new corridors (path between two nodes) are selected for investment. By sampling many generation scenarios, the model captures the maximum number of transmission corridors possible while solving quickly. With this approach, however, the possibility that some ‘flexible’ corridors, those with mediocre value in a single scenario but high expected value across scenarios, will not be identified [40].

$$\text{Eq. 17} \quad \text{Min} \sum_{y=1}^Y \left[ct \cdot x_{k,y} + \sum_{h=1}^H \left(\sum_{n=1}^N (cg \cdot g_{n,h,y}) + \sum_i^I (c\mu \cdot \mu_{i,h,y}) \right) \right] \cdot \alpha_y$$

s.t.

$$\text{Eq. 18} \quad d_{i,y,h} = \mu_{i,y,h} + \sum_{k \in \Omega_i} f_{k,y,h} + \sum_{n \in \sigma_i} g_{n,y,h} \quad i=1, \dots, I$$

$$\text{Eq. 19} \quad f_{k,h,y} - 1/X_k (\theta_{i(k),y,h} - \theta_{j(k),y,h}) \cdot b = 0 \quad k \in K^0$$

Eq. 20	$x_{y,k} \leq f_{k,y,h} \leq x_{y,k}$	$k \in K^+$
Eq. 21	$d_{i,y,h} \geq \mu_{i,y,h}$	$i=1, \dots, I$
Eq. 22	$g_{n,y,h} \leq g_{n,y,h}^{max}$	$n=1, \dots, N$
Eq. 23	$x_{k,y} \leq f_k^{max}$	$k=1, \dots, K^0$
Eq. 24	$x_{k,2} \geq x_{k,1}$	$k=1, \dots, K^0$

Indices and Sets:

i	index of buses
k	index of circuits
i(k), j(k)	index of terminal buses of circuit k
h	index of load hours
n	index of generators
K^+	set of candidate circuits
K^0	set of existing circuits
Ω_i	set of all circuits connected to bus i
σ_i	set of all generators located at bus i
H	number of load hours
I	number of buses
M	number of candidate circuits
N	number of generators
Y	set of years considered

Parameters/Constants

cg	generator costs [\$/MW]
ct	annualized cost of candidate circuits [\$]
cμ	cost of non-served energy [\$/MW]
d	bus demands [MW]
f^{max}	circuit capacities [MW]

g^{\max}	generator capacities [MW]
X	circuit reactances [pu]
b	per unit base

Free Variables

θ	bus voltage angle
----------	-------------------

Positive Variables

f	circuit flow [MW]
θ	positive bus voltage angle component across circuit k
g	generator output [MW]
μ	non-served energy [MW]
x	investment variable

2.1.3 Transformation from Corridor Capacity to Investments

The goal of the St. Clair Screening Model is to identify a more limited set of investments for planners to consider. The relaxed linear model thus far, however, can only specify corridors which receive investment and the amount of investment they receive. In order to determine specific investments, these continuous capacities must be transformed back into lumpy integer investments. In the St. Clair Screening Model, this transformation is calculated via the St. Clair curves discussed earlier.

The transform from continuous investment variables to specific investments has two steps. First, for each corridor, the maximum capacity for each type of investment is calculated. Next, each continuous investment variable is compared to the capacity for each investment type, and the smallest and thus least cost investment is selected. For example, consider the hypothetical continuous investment variables for a 500 mile corridor shown in Table 2-4 and the capacities for the different investment types in Table 2-5. Cycling through the continuous investment quantities, the smallest investment to

provide the 250 MW is the 345kV line and the lowest cost investment to meet the capacity for the 2,000 and 2,200 MW capacity requirements is a 765kV line.

Table 2-4 Illustration of Linear Model Results

Corridor Capacities (MW)		
250	2,000	2,200

Table 2-5 Capacity for 500 Mile Lines by Investment Type

345kV	500kV	765kV
500 MW	1,100 MW	2,400 MW

2.1.4 St. Clair Screening Model Flow

The previous three subsections reviewed the assumptions and limitations of the various sub-modules in the St. Clair Screening Model. When combined, the flow of information and connections between modules is shown in Figure 2-5. The first step of the screening model is to rate all potential investments both financially and as discussed in 2.1.1. The model iteration shown in Figure 2-5 considers the three different line ratings characteristic to the United States, but other investment types may be considered. As discussed in 2.1.2, the linear optimization model must be run considering each investment type separately and the optimization is repeated for sampled deterministic generation expansion scenarios to identify an amount of new capacity (MW) in each corridor. The continuous investment variables from each linear optimization are then transformed to investment variables via a St. Clair curve as discussed in 2.1.3.

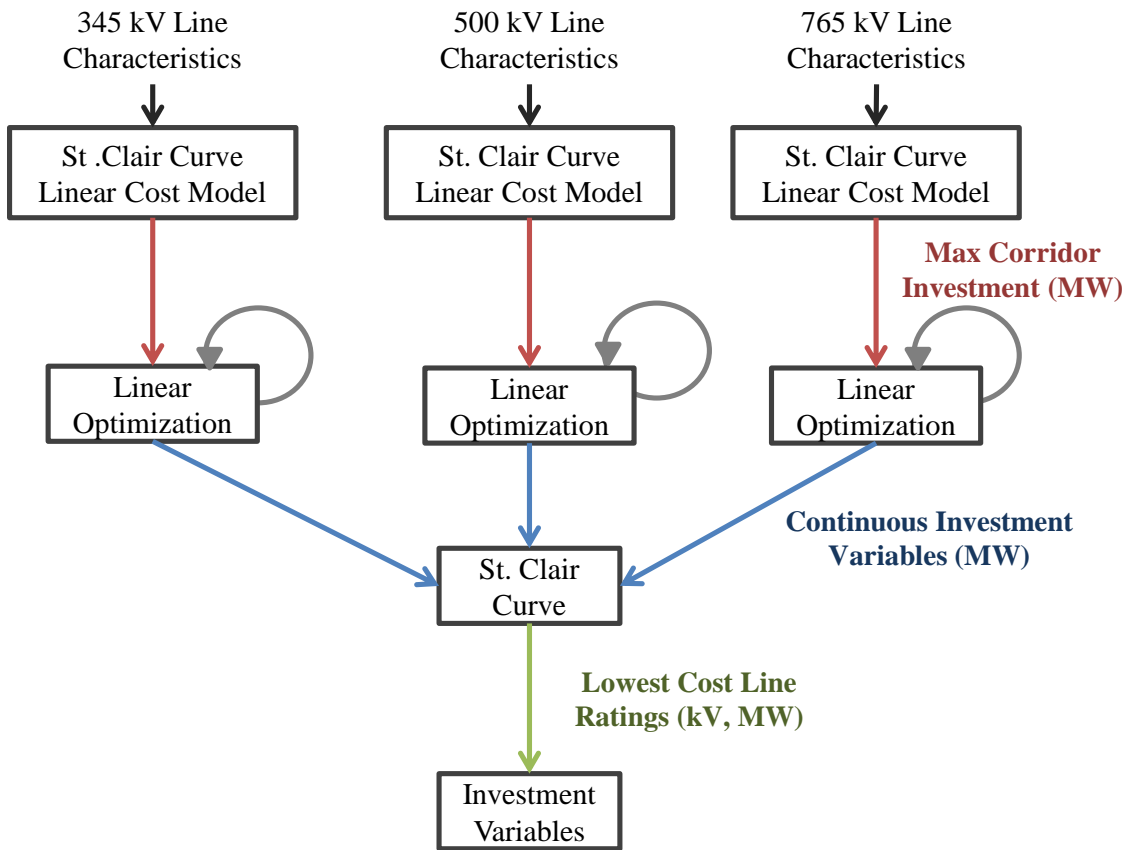


Figure 2-5 St. Clair Screening Model

2.2 Screening Model Demonstration Problem

The St. Clair Screening Model is demonstrated in the following sections on a reduced order model of the Western Electric Coordinating Council (WECC) footprint. WECC is one of the three electrically independent interconnections in the North America, shown in Figure 2-6. This model rather than existing IEEE test system models was used because the architecture of the wide-area extra high voltage network differs from distribution or small systems. Medium and low voltage distribution models, for example, are generally radial rather than meshed, and the extra high voltage distribution lines considered in the St. Clair Screening Model are too large to be deployed in a distribution setting.

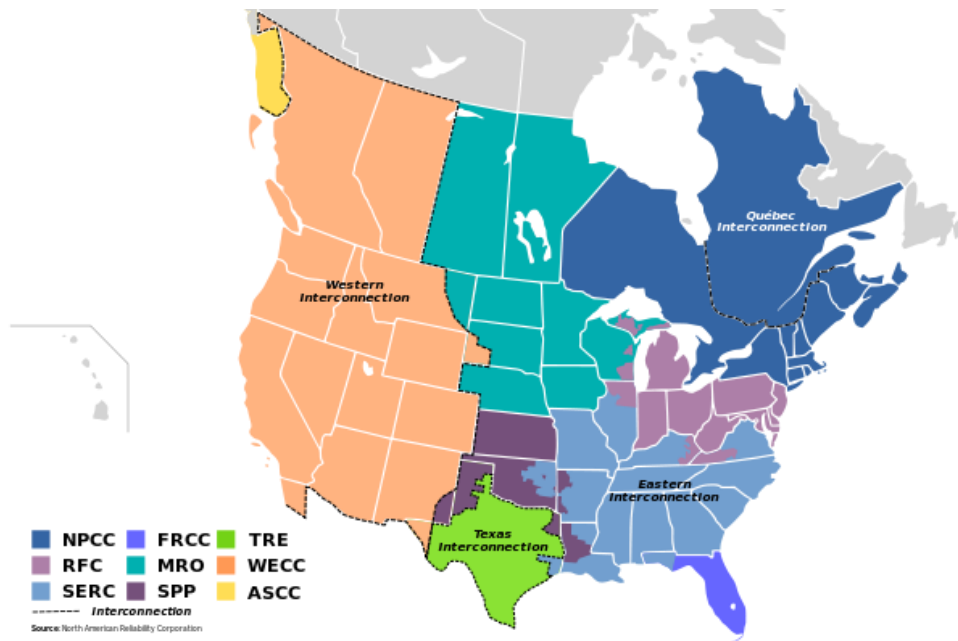


Figure 2-6 North American Interconnections

2.2.1 WECC System Model

The test system used in this model is a 240 bus reduced-order model of the WECC footprint, developed by the California Independent System Operator [49]. For this demonstration, the WECC model was further simplified by combining nodes located within five miles into a single geographically unique zone. This simplification was made to focus on major transmission investments rather than transformer capacity or small local reinforcements. The nodes within each geographically zone were connected with transportation lines, lines with directable flow in the hybrid optimal power flow model, with 20,000 MW capacity to avoid congestion. Collapsing the system in this way reduced the number of nodes from 240 to 113. It also reduced the number of existing lines by 10% from 329 to 296; these changes are summarized in Table 2-6. This base model was augmented with 53 unique Western Renewable Energy Zones (WREZs) developed by the Western Governors’ Association [69] for a total of 164 unique nodes. Each WREZ is a designated geographic area which contains a significant quantity of high quality renewa-

ble energy resources. These WREZs are the sole location for new renewable generation development in the model.

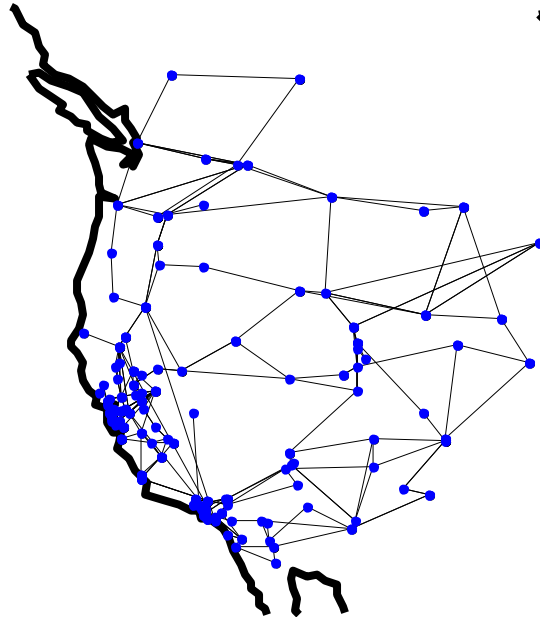


Figure 2-7 Existing Nodes and Susceptance Lines in the Modified WECC System

Table 2-6 Size Comparison of Original and Simplified WECC Models

	Nodes	Transmission Lines
Original Model	240	329
Simplified Model	113	296

The demonstration model considers 13,695 individual corridors. These corridors include all possible connections between existing WECC nodes, WECC nodes and WREZs, and between all WREZs. For each generation scenario, the relaxed TNEP problem was solved with three different corridor ratings: 345kV double circuit, 500kV single circuit and 765kV single circuit. The 345kV double circuit rather than 345kV single circuit was used to reflect industry preference for new investment at the 345kV rating. The SIL, thermal limits and costs for these different line types are summarized in

Table 2-7. When enumerated at all voltage levels, this leads to 41,085 unique investment variables (e.g. 3 possible investments for each of the 13,695 corridors) and $2^{41,085}$ possible unique plans if any combination of lines may be selected.

Investment in the transmission network is induced in the model through annual load growth. Load growth in the model is assumed to be 2% annually, and a stratified sample of 19 load hours is modeled in each time horizon. The load hours modeled were stratified seasonally, to reflect the changing load patterns across the wide geographic area as well as across load levels (peak, intermediate and baseload). The load hours were also stratified across hydro-production levels due to the high percentage of hydropower in the Pacific Northwest. To focus on transmission build-out prompted by uncertainty in renewable energy in this demonstration problem, existing thermal generation was also assumed to grow at 2% annually. While not a realistic assumption, the exogenously applied growth in thermal capacity allows for exploration of transmission investment based on changing location and quantities of new renewable generation without also running a generation expansion model to assure adequacy of generation supply. For problems focused on broader questions, this assumption may not be necessary and generation expansion may be handled differently.

Table 2-7 Summary of Transmission Line Properties

	345kV (double circuit)	500kV (single circuit)	765kV (single circuit)
SIL (MW)	400	880	2090
Thermal (MW)	1500	3000	8000
Cost (MUSD/mi)	2.0	2.9	3.3

2.2.2 Generation Sampling

The question driving this research is how to plan the transmission network in the face of uncertainty in the development of location-constrained generators. As a result, the uncertainty considered in the demonstration of the St. Clair Screening Modeling is the location and quantity of new wind and solar generation. The uncertain nature of future generation development was captured using 500 randomly sampled generation-expansion

scenarios. These scenarios are deterministic expansion paths characterized by the amount and location of new generation on 10 and 25 year time scales. While the only uncertainty in this model is the quantity and location of location-constrained generation, the framework of the St. Clair Screening model could also be used to examine other uncertainties.

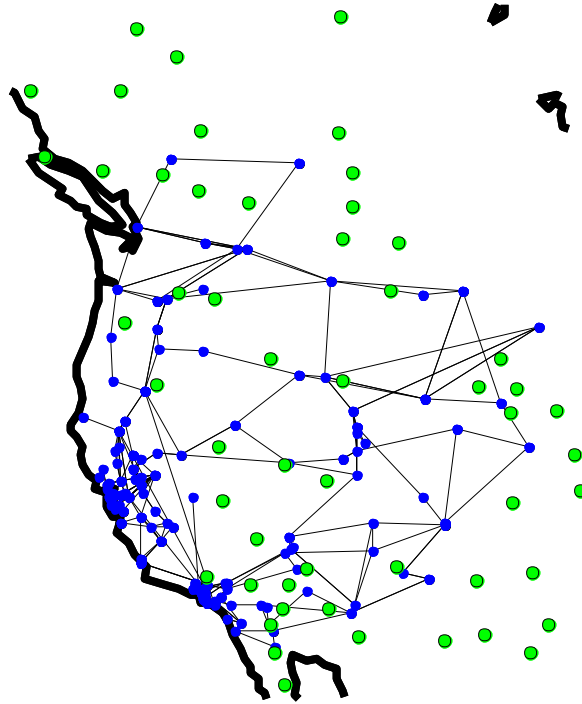


Figure 2-8 Existing Nodes (Blue), WREZs (Green) and lines in the Modified WECC System

The first 250 generation expansion scenarios used in the WECC demonstration problem are stratified by the number of WREZs developed. The location of these WREZs is shown with the location of existing nodes and transmission lines in Figure 2-8. Rather than directly sample a percentage of zones, each WREZ was assigned a probability of selection such that an expected number of WREZs were developed. For example, if there were 100 WREZs and the goal was to sample 10% of zones, each WREZ would be assigned a probability of selection of 0.1. With these probabilities assigned, a random number was generated for each WREZ and only those with a random number less than 0.1 would be selected for development. For the demonstration problem here, samples

were stratified to develop 20%, 40%, 60% and 80% of WREZs as shown in Table 2-8. A greater number of 20% WREZ samples were developed because a greater number of unique WREZ combinations are possible.

Table 2-8 Number of Generation Scenarios by Expected Percentage of WREZs

Expected Percentage of WREZs Developed	20%	40%	60%	80%
Equal Probability of Selection	100	50	50	50
Weighted Probability of Selection	100	50	50	50

The majority of WREZs contain modest quantities of potential wind and solar generation. Of the 53 WREZs considered, only 23% have potential for more than 1,000 MW new generation⁶. To ensure the inclusion of scenarios with higher penetrations of renewable energy, an additional 250 samples were included where the probability of a WREZ’s selection for development was directly proportional to its potential capacity of wind and solar generation development.

Once the WREZs were selected using the methods outlined above, the quantity of generation for each WREZ was sampled. For each selected WREZ, the quantity of new generation to be added, ranging from 0-100% of the total potential was randomly sampled. In the second stage, the potential generation for development in each WREZ was the total generation not developed in the first stage.

The set of randomly sampled generation scenarios were augmented with two additional scenarios. The first scenario assumed no new generation development in any WREZ, and the second scenario assumed complete development in each zone. All 502 samples were solved considering the full set of possible corridors in the WECC/WREZ model and three investment types for a total of 1,506 iterations.

2.3 Results

The primary goal and result of the St. Clair Screening Model is a reduced set of new transmission lines for consideration. This reduction, over 97% in the demonstration problem, is discussed in 2.3.1. The results provided, however, are much richer. An

⁶ 1,000 MW after capacity factor de-rate

iterative method, the screening model provides the frequency of each corridor's use. As discussed in 2.3.2, examining the frequency of investment can lead to further insights into potentially robust corridors and information to guide additional optimizations. The St. Clair Screening Model also provides insights into the use of linear relaxation, both the importance of considering multiple investment types on each corridor (2.3.3.1) and the validity of the linear relaxation for the screening model (2.3.3.2). Finally, as formulated, the model provides insight into high value renewable energy zones.

2.3.1 Reduction in Transmission Investments

Each iteration of the St. Clair Screening Model identifies the amount of new capacity in each corridor. These capacities, continuous variables, are transformed to unique investments specified by location (corridor) and rating (kV) as discussed in 2.1.3. In the demonstration problem, the 1,506 two-stage optimizations identified 491,173 non-zero investment variables. These continuous investment variables in turn identified 1,081 investments (unique by both location and rating). With 1,081 investments, only 0.2% of investment variables identify unique investments. This indicates, that across optimizations, the same investments are being identified repeatedly.

The reduced set of investments identified by the screening model represents a 97% reduction of investments to consider. Of the 1,081 remaining investments 629 (58%) are 345kV double circuit, 301 (28%) are 500kV single circuit and 141 (14%) are 765kV single circuit. It is not surprising that a greater number of 345kV double circuit (smallest investment option) lines are identified as 345kV lines may be identified in all runs. On the other hand, 500kV investments may only be identified in 2/3 of the runs and 765kV investments may only be identified in 1/3 of the runs. For example, in an optimization where all transmission corridors have been rated to a 500kV capacity, the optimization can select capacities up to the 500kV capacities. If, however, the selected capacity is less than (or equal to) the 345kV line capacity, it will be transformed to a 345kV investment. A 765kV investment, on the other hand, can never be identified in that same optimization because the maximum capacity allowed for selection is the 500kV capacity.

2.3.2 Frequency of Corridor Use

Examining the corridors where new investments take place can lead to further insight for planners. For example, the frequency of investment in each corridor can be used to guide follow-up optimization studies (i.e. assigning higher branching priorities to investments with higher frequencies) or identify corridors which are robust to uncertainty in generation expansion scenarios and should receive further routing and siting analysis. The first section here examines the frequency of development per corridor across time stages. In 2.3.1, the corridors with development in the first stage (mid-term planning horizon) are analyzed. These corridors are of greater interest because they would house the investments which must be decided upon today. Finally, in 2.3.2.2, corridors which may be considered robust for their frequency of development are discussed.

With each iteration of the screening model, the optimization selects corridors for development. In the WECC demonstration problem, the St. Clair Screening Model identifies 5% (629) of the possible corridors for investment. As shown in Figure 2-9, the number of unique corridors identified per iterations reduces with the number of iterations, and all unique corridors are identified within the first 1,175 iterations of the model. It should be noted that the samples in Figure 2-9 are not random but ordered as given in Table 2-9 with the no-generation and all-generation samples first.

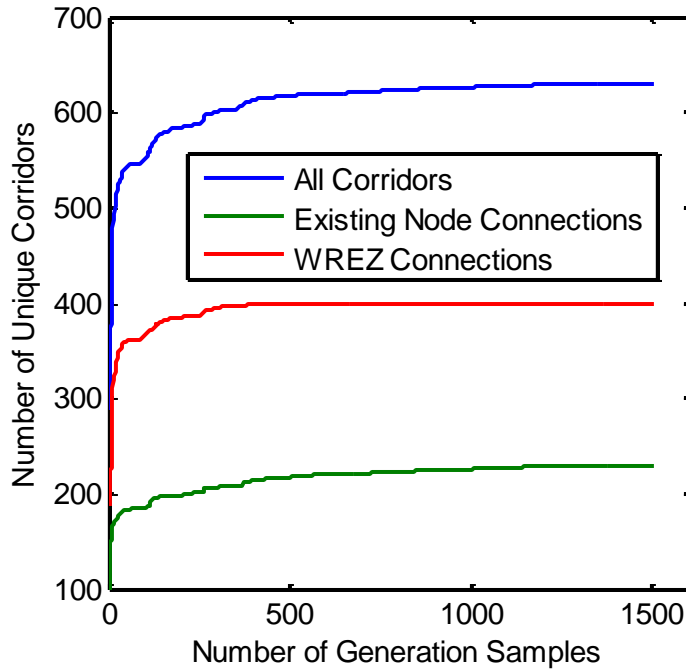


Figure 2-9 Number of New Unique Corridors with Additional Generation Samples

Table 2-9 X-Axis Ordering Given in Figure 2-9

x-axis number in Figure 2-9	Generation Samples	Transmission Rating
1-3	No Generation Developed	345kV, 500kV, 765kV
4-6	All Generation Developed	345kV, 500kV, 765kV
7-506	Randomly Sampled Generation	345kV
507-1,006	Randomly Sampled Generation	500kV
1007-1,506	Randomly Sampled Generation	765kV

The number of corridors, 629, is smaller than the number of possible investments, 1,081, because multiple types of investments may be made in each corridor. For example, one iteration of the screening model may identify a 500kV investment in a specific corridor while the next may identify a 765kV investment in the same corridor. Both iterations identify the same single corridor but produce two investment variables.

These corridors, paths between two nodes, are developed with frequencies in the demonstration problem range from 99.9% of generation scenarios to 0.033% of generation scenarios. As shown in Figure 2-10, of the corridors which are developed in some iteration, the majority are developed in less than 11% of all scenarios. Only 21% (132) of the corridors are developed in more than 50% of scenarios. These 132 corridors, less than 1% of the original 13,695 corridors, may be of further interest for routing studies.

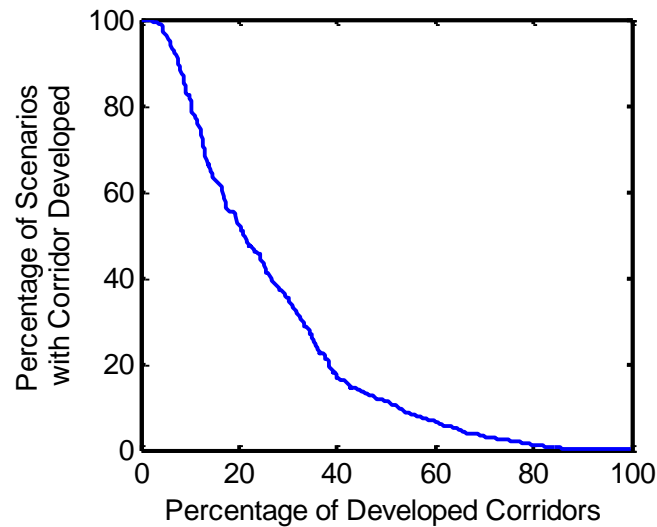


Figure 2-10 Frequency of Corridor Development by Percentage of Developed Corridors

2.3.2.1 First Stage Investments

Thus far, the frequencies of corridor investment have considered both the mid-term (first) and long-term (second) time horizons. These frequencies are useful for screening corridors for future planning studies, especially optimization planning studies. The information on decisions that planners must make in the near term, however, is found in the mid-term time horizon (10 years). In the demonstration problem, a smaller percentage of corridors, 3.4% (472 rather than 629) are selected for investment in the first stage. The distribution of frequencies for this smaller set of corridor is also steeper. As shown in Figure 2-11, fewer than 40% of corridors are developed in more than 10% of first stage scenarios. This continues to narrow the number of corridors that planners must character-

ize further. There are, for example, only 157 corridors developed between existing nodes, and only 40 of these are developed in at least 50% of the scenarios as also shown in Figure 2-11. This can continue to focus planners on corridors for further characterization.

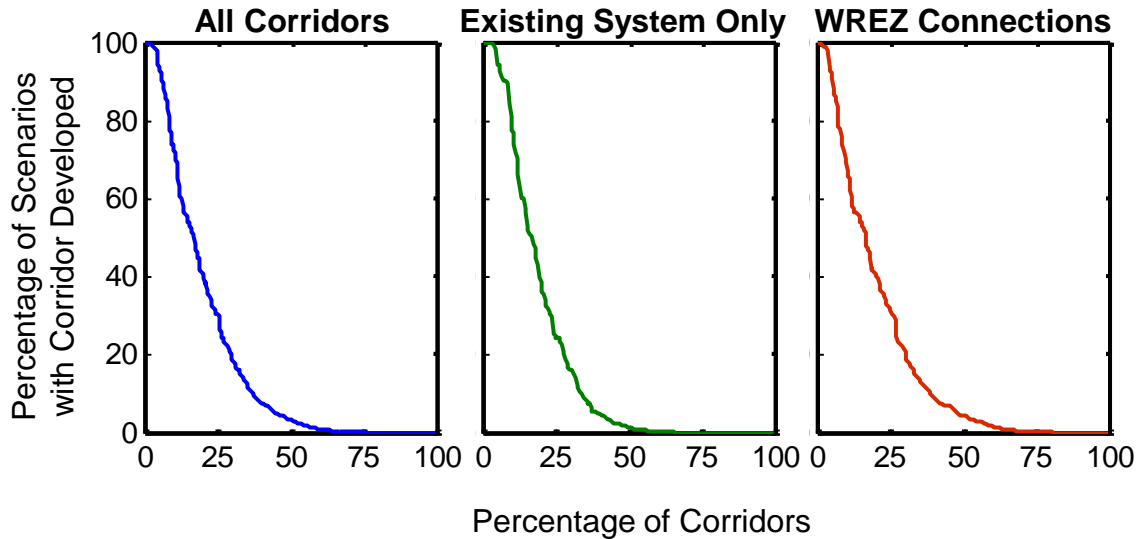


Figure 2-11 Frequency of Corridor Development by Percentage of Developed Corridors in First-Stage Scenarios

2.3.2.2 Robust Corridors

A small subset of corridors in the demonstration problem may be characterized as robust corridors. Robustness in this context means that the corridor receives investment in nearly every generation scenario; it is robust to future uncertainty. There are 41 such robust corridors in the demonstration problem. These corridors have investment in at least 90% of future scenarios and are shown in Figure 2-12. If the robustness criterion is increased to 95% of scenarios, 28 corridors may be considered robust, and if the robustness criterion is increased to 99% of scenarios, 19 corridors may be considered robust. These corridors can be singled out for further analysis or fixed in future planning studies.

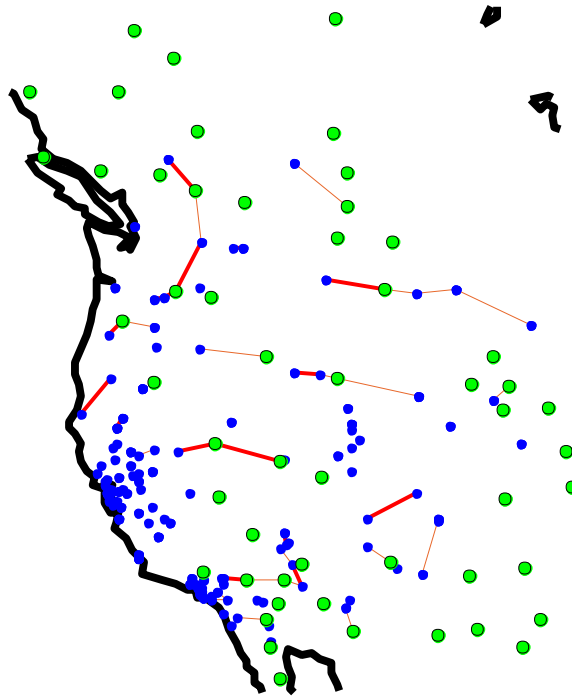


Figure 2-12 Map of Corridors Developed in At Least 90% of Scenarios

Thick red lines show corridors developed in 99% of scenarios or greater. Thinner orange lines show corridors that are developed in greater than 90% but less than 99% of scenarios.

Table 2-10 Percentage of Corridors Developed in At Least 90% of Scenarios by Type

	>90%	>95%	>99%
WECC Connections	19	11	10
WREZ Connections	22	17	9

Given the future uncertainty in generation development, it would be reasonable to expect that all robust corridors connect existing nodes. After all, with no certain new generation or demand at a node, there is no obvious reason to interconnect the node. As shown in Figure 2-12 and Table 2-10, however, more than 50% of the robust corridors (using the 90% criterion) connect either a WREZ to a WREZ or a WREZ to an existing node. These corridors are developed frequently for a combination of reasons. First, the short line bias discussed in 2.1.2 means that the model selects series of short segments over continuous long lines. An example of this short line bias in the Pacific Northwest is

highlighted in Figure 2-13. In Figure 2-13, the transmission segments A-C and C-B are constructed through the WREZ, marked C, rather than directly from the existing nodes, labeled A and B. Second, in this iteration of the screening model, all corridors were eligible for construction in each scenario. Thus, even if no new generation is constructed at the WREZ labeled C, corridors A-C and C-B may be developed. This secondary issue could be eliminated by allowing construction only between existing nodes and new nodes with generation development. In this case, corridors A-C and C-B would only be available for development if WREZ C was selected for generation expansion. This constraint was not made in the current iteration of the model, and the resulting information on potentially high value WREZs is discussed in 2.3.3.2.

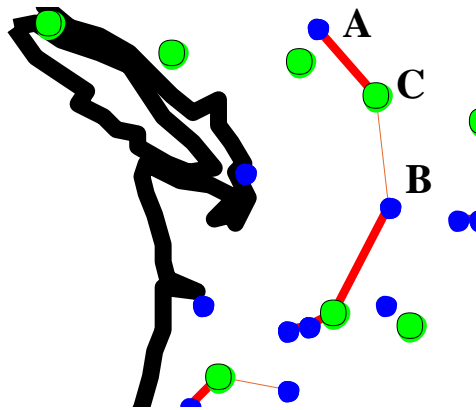


Figure 2-13 Illustration of Short-line Bias in Demonstration Case

In the demonstration case, there are several instances of robust corridors which form a pass-through. A pass-through is defined to be a sequence of corridors where a WREZ forms an intermediary point between two existing nodes, such as above in the A-C, C-B example. Logically, both the pass-through and the direct route (e.g. line A-B in the example above) should be included in the set of identified corridors. In the demonstration problem, there are eight such pass-throughs using the 90% of scenarios criteria for robustness. For each of these pass-throughs, the direct route was also identified as a corridor of interest by the screening model.

The WREZs used in the demonstration problem contain high quality of renewable energy potential and are mainly differentiated by the type and quantity of resource potential. Due to a combination of short-line bias and allowing investment to WREZs whether or not generation investment is made, several WREZs are interconnected in over 90% of scenarios. These paths through WREZs may signal to system planners that as power flow is desired along the path, these WREZs may be particularly advantageous to develop. Routing paths near these zones provide necessary reinforcement capacity for the system and economic benefit without WREZ development but also have the option to access these resources. As shown in Table 2-11, if all segments constructed greater than 99% of scenarios in first stage are developed, direct access to 10,000 GW of wind and solar power would be provided. If these segments are expanded to include those constructed at least 90% of the time, this would provide access to an additional 9,000 GW.

Table 2-11 WREZ Wind and Solar Resources (GW) Accessed by Lines in the First Stage

	Robustness Criterion	
	>99% Scenarios	>90% Scenarios
Wind (MW)	3,517	9,780
Solar (MW)	6,225	9,228
Total (MW)	9,742	19,009

2.3.3 Insights into Linear Relaxation

The linear relaxation of the transmission network expansion planning problem has been used both to warm-start mixed-integer optimizations and to screen investments [55],[56]. There are, however, limitations to the relaxation as discussed in 2.1.2. The output of the screening model demonstrates the importance of running the linear relaxation using multiple corridor ratings. In the following sections, the necessity of running the screening model optimizations using multiple investment characteristics and the ability of the linear relaxation to respect the lumpy nature of the problem are explored.

2.3.3.1 Importance of Multiple Voltage Characterizations

The goal of the St. Clair Screening Model is to identify the set of promising transmission investments for further analysis. Because multiple types of investments (e.g. 500kV and 765kV) cannot be directly compared in the linear investment model, each generation scenario in the St. Clair Screening Model is optimized assuming maximum investment capacity and cost characteristics for the three different investment types (345kV double circuit, 500kV single circuit, 765kV single circuit). As the optimizations using the 765kV corridor ratings are able to identify 765kV, 500kV, and 345kV investments, it may be non-obvious why the 345kV and 500kV iterations are necessary. After all, the 345kV corridor rating runs are only able to identify 345kV investments and the 500kV corridor rating iterations are only able to identify 345kV and 500kV investments.

The lower capacity iterations are necessary due to the relaxation of physical flow constraints in the linear optimization model. As discussed in 2.1.2, the linear relaxation allows for flows on new lines to be driven by economic rather than physical constraints (Eq. 4). This relaxation potentially allows for larger flow through new investments than would be allowed physically⁷. The constriction on new flows results from the differences in operating voltage level and physical properties of the lines. For example, consider the hypothetical example given in Figure 2-14. In the example, a 500kV line is considered to augment an existing 345kV line in corridor A-B. In the hybrid optimal power flow, the 500kV line adds 2,200 MW for a total of 3,200 MW capacity in corridor A-B. Physically, however, only an additional 1,420 MW are added. The effective capacity is smaller than expected due to restriction on angular differences. Without the new 500kV line the maximum angular difference (calculated using Eq. 25 and demonstrated in Eq. 26) is 0.5 radians or about 29 degrees. When the new 500kV line is added, the new maximum angular difference is constrained by the 500kV line and reduced to 0.11 or about 6.3 degrees (Eq. 27). With this new smaller maximum angular difference, the 500kV line may convey 2,200 MW but the flow for the 345kV line is reduced to 220 MW (Eq. 28).

⁷ The hybrid OPF also allows for counter-flows wherein the difference in angles would dictate flow in one direction, but the flow in the new lines runs opposite direction. The author is unaware of methods to mitigate this issue.

The net effect is to reduce the effective capacity to 2,420 MW or 76% of the expected 3,200 MW. In some cases, this constriction is sufficient to change the optimal investment.

Eq. 25
$$\Delta\theta_k = flow_k \cdot X_k$$

Eq. 26
$$\Delta\theta_{k-max} = 1000 \cdot 5 \cdot 10^{-4} = 0.5$$

Eq. 27
$$\Delta\theta_{k-max} = 22005 \cdot 10^{-5} = 0.11$$

Eq. 28
$$flow_{345kV} = 0.11 \cdot \frac{10^4}{5} = 220 \text{ MW}$$

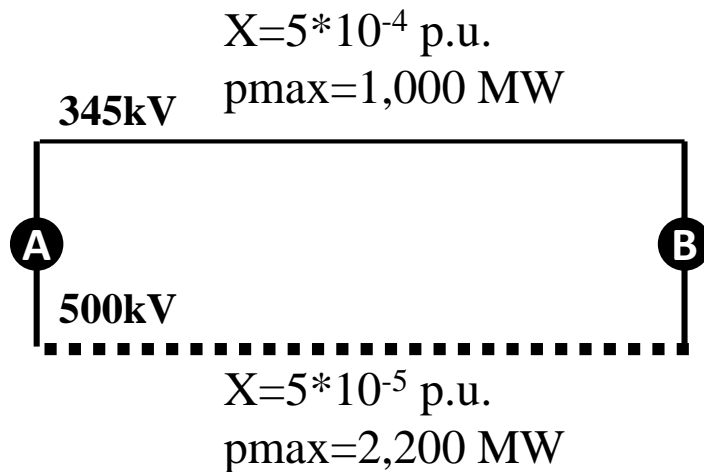


Figure 2-14 Artificial Network Constraint Example

The lower capacity optimizations would not be necessary if the 765kV iterations capture all of the investments generated by the 345kV and 500kV iterations. If this were true, the potential constrictions would not change the investment types. As shown in Table 2-12, however, the 765kV iterations capture only 57% of the investment variables generated. Specifically, it captures only 42% of the 345kV investment variables and 68% of the 500kV investment variables. This demonstrates the necessity of the lower capacity iterations.

Table 2-12 Number and Percentage of Lines By a Single Voltage Level St. Clair Screening Model

		Percentage of Total Lines Identified	Percentage of Lines Identified by Voltage Rating		
			345kV	500kV	765kV
Corridor Characterization	345kV	57	96	0	0
	500kV	62	62	97	0
	765kV	57	42	68	100

2.3.3.2 Validity of Linear Relaxation

One concern with the linear relaxation is that it will not accurately represent the lumpiness of the transmission expansion problem. To check whether or not the lumpiness was respected, capacity of all 491,173 non-zero continuous investment variables were compared to the smallest investment size considered in the problem (345kV double circuit). If most corridors are developed at a small percentage of the 345kV capacity, it would indicate that only a marginal investment was made in the corridor, that the full lumpy investment of a transmission line was not required and that the optimization is not accurately representing the TNEP problem. On the other hand, if the corridors are developed at a high percentage of their 345kV capacities, it would be expected that a full transmission line would be selected in the integer problem, and the lumpiness of the problem is still being respected. As shown in Figure 2-15, very few (6.3%) continuous investment variables are developed at less than 10% of the 345kV capacity, and a majority, 60% of investments, are developed at greater than 90% of the 345kV capacity. This indicates that in the demonstration problem, the linear relaxation is not perverting the inherent lumpiness of the problem. While this analysis confirms that the lumpiness of the problem is respected for the majority of investments, the other distortions of the linear relaxation (e.g. the hybrid optimal power flow) remain.

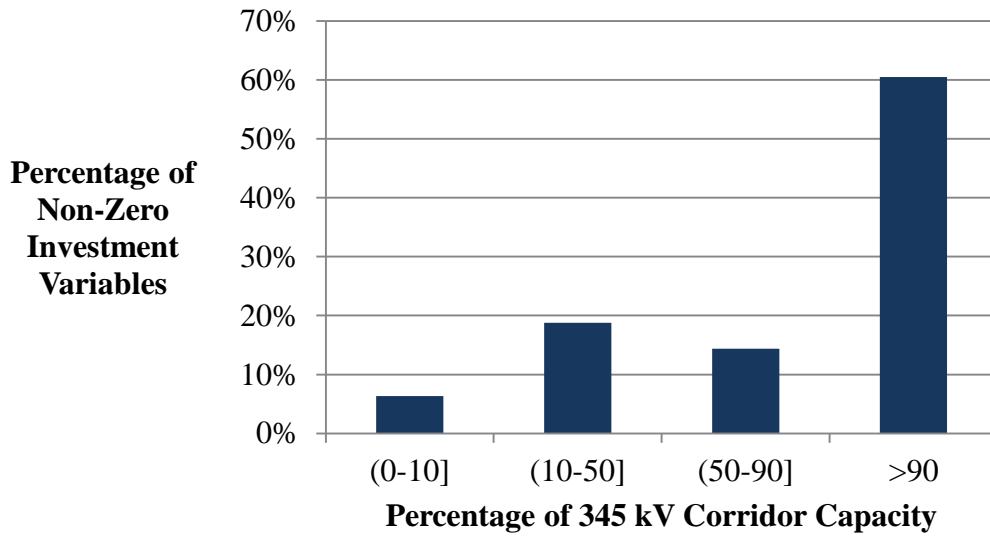


Figure 2-15 Histogram of Investment Variables by Percentage of 345kV Capacity

2.4 Conclusions and Future Work

The screening model presented here, the St. Clair Screening Model, is an effective method for reducing the corridors and investments required for consideration in stochastic wide-area transmission planning. The full problem is made tractable by transforming the multi-stage stochastic integer problem to a series of deterministic multi-stage scenarios. The integer problem is also relaxed to a linear problem with continuous investment variables translated into integer variables through the use of St. Clair curves. In a case study on the Western Electric Coordinating Council footprint, the St. Clair Screening Model was able to reduce the number of corridors, connections between specific node pairs, by 95%. The screening model also reduced then number of lines, connections between specific node pairs with a voltage rating, necessary for consideration by 97%. These dimensionality reductions are summarized in Table 2-13.

Table 2-13 Summary of Dimensionality Reduction Through St. Clair Filter

	Corridors	Lines
Percentage Selected	4.59% (629 of 13,695)	2.63% (1,081 of 41,085)
Reduction	95.31%	97.37%

In addition to reducing the number of corridors and lines for consideration in an optimization or other technical study, the St. Clair Screening Model provides information about the frequency of investment in individual corridors. This information could be embedded in a tree search, such as a branch and bound optimization. It can also be used to identify corridors developed in all or nearly all scenarios which may be robust to uncertainty. In the WECC case study 41 corridors were identified as developed in at least 90% of first-stage scenarios. These 41 corridors connected both existing system nodes as well as WREZs and as a result may indicate WREZs which are advantageous to develop, as they lie along economically advantageous pathways.

There are several future refinements possible for the St. Clair Screening Model. As discussed, the St. Clair curves are inherently biased toward shorter connected transmission segments than single long lines. This bias may be exacerbated by using line only costs rather than line costs as well as additional substation costs for both line compensation and transformers. Further refinement of the model may explore the effect of including these costs on the corridors selected. Additionally, the optimal power flow used in this work is lossless; however, investment in new transmission is prompted by changes in both magnitude and location of generation, load and losses in the system. This work focused on the geographic location and quantity of generation; however, the work should be extended to examine the impact of losses on the corridors selected by the filter.

THIS PAGE INTENTIONALLY BLANK

3. Dimensionality Reduction and Interpolation for Multi-stage Stochastic Transmission Expansion Planning Algorithms

The multi-stage stochastic transmission expansion planning problem is computationally intractable using traditional optimization methods. Even when the full set of investments has been reduced from many thousand potential investments to approximately 1,000 potential investments, such as was shown in Chapter 2, the problem remains beyond the scope of traditional methods in a reasonable time frame. When traditional methods, such as Branch and Bound or Bender's Decomposition, cannot be used, heuristic algorithms are used instead. Meta-heuristics such as Simulated Annealing or Tabu Search have been used for smaller transmission expansion planning problems and are blind to the computational problems associated with traditional optimization methods [36],[58]. These methods, however, are also blind to problem specific characteristics which may make the problem easier to solve and provide further insight into the structure of the problem.

Rather than meta-heuristics, a third approach is to use heuristic algorithms guided by problem specific knowledge. For example, CHOPIN is a heuristic transmission expansion planning model built around a search tree. This tree search is guided by information derived from optimal power flows [31]. More broadly, classes of algorithms such as approximate dynamic programming are flexibly structured to integrate problem specific knowledge. These models all rely on interpolation, estimating the value of new transmission plans using information from previously characterized transmission plans. Both the dimensionality and the networked nature of the wide-area MS-TNEP problem effectively preclude the use of existing interpolation methods. Instead, a method to predict these costs must reduce the dimensionality of the MS-TNEP problem and then predict a value in the reduced dimensional space. Unlike the dimensionality reduction in Chapter 2 which was designed to reduce the number of independent variables (investments), the dimensionality reduction here is a transformation from a search space based on the investments to another space.

This work proposes new approaches to approximate the MS-TNEP problem combining dimensionality reduction methods from the image processing literature with

traditional interpolation techniques. This chapter proceeds in six major sections. Section 3.1 describes the challenges of developing a reduced order model using existing methods. Next, Section 3.2 introduces two dimensionality reduction methods from the image recognition literature and 3.3 discusses two interpolation techniques. Finally, 3.4 evaluates combinations of the dimensionality reduction and interpolation techniques as reduced order models for transmission expansion planning. Conclusions are presented in 3.5 and future work in 3.6.

3.1 Approximating the Multi-Stage Stochastic Transmission Expansion Network Planning

The existing transmission planning literature has focused on reinforcement planning, adding a small number of lines to existing well developed systems. In this context, new transmission investments are unlikely to change the major flow patterns in the network. As a result, methods which consider transmission lines on a one-by-one basis may be considered. These methods implicitly consider the investments to be independent, that is, the value of adding transmission line A and transmission line B is approximately the same as adding both at the same time. The wide-area problem, however, is fundamentally different. In this problem, a great number of alternatives are considered, with the value of a new investment dependent on many other possible investments. Unlike the reinforcement problem, there are also a majority of plans which produced high levels of non-served energy.

The MS-TNEP problem presents fundamental problems for constructing and exploiting a reduced order model. The shape of a response surface for the transmission expansion planning problem can be thought as a landscape of steep canyons as in Figure 3-1. Of the trillions of possible transmission expansion plans, most have high costs because demand is not met and non-served energy is very expensive as illustrated in Figure 3-2. A very few plans, those in the canyons, have all demand met. The goal of an optimization algorithm is to identify those areas of the surface where all demand is met and then find the optimal solution in this smaller area.



Figure 3-1 Aerial of the Grand Canyon [36]

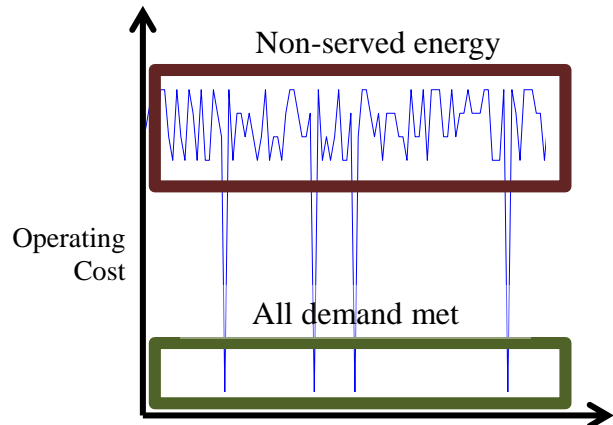


Figure 3-2 Illustration of Costs

A reduced order model of the transmission expansion planning model for the wide-area should be designed to capture different patterns of investment. These patterns could be specific combinations of lines or more general observations of north-south and east-west investment. Because the networked problem produces non-intuitive investment combinations, defining what will characterize the structure of these patterns a priori is not possible. The difficulty of a priori assigning the patterns or neighborhoods of solutions was noted early in transmission planning research by Duschonnet in 1972 [18]; however, no generalized solution was offered by Duschonnet and research in the transmission planning field moved away from dynamic programming methods and toward linear programming and mixed-integer methods.

One of the key difficulties of defining investment patterns as combinations of lines is the sensitivity of operational costs to incremental investments. For example, a very advantageous investment pattern in the test problem shown in Figure 3-3 could be a build-out of a key east-west 500kV transmission line. This line allows demand to be met at each of the three demand nodes with minimum cost. The value of plans containing this line, however, is dependent upon the construction of two of the three smaller connector lines (Line 1, Line 2 and Line 3 in Figure 3-3). With at least two lines constructed, all three load centers have access to the generation. With only one of these three lines, one

load center remains unconnected to the generator and has unmet demand. Indicative costs demonstrating these cases are shown in Figure 3-4. As shown in Figure 3-4, plans with two or three connector lines have very low operational costs. Plans with only zero or one lines have dramatically higher operational costs due to the costs of non-served energy. It is important to note that in the transmission planning context, high non served energy is realistically unacceptable and that operational costs are measured in the billions of dollars. It is tempting to assume that these smaller connections can easily be identified, but as demonstrated in Figure 3-3, there many possible advantageous and disadvantageous combinations even in a very small system.

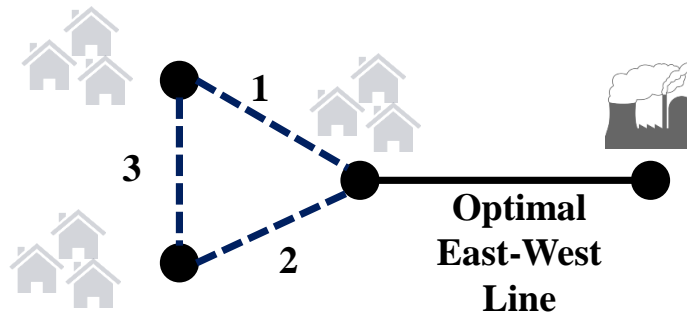


Figure 3-3 Feature Identification Test System

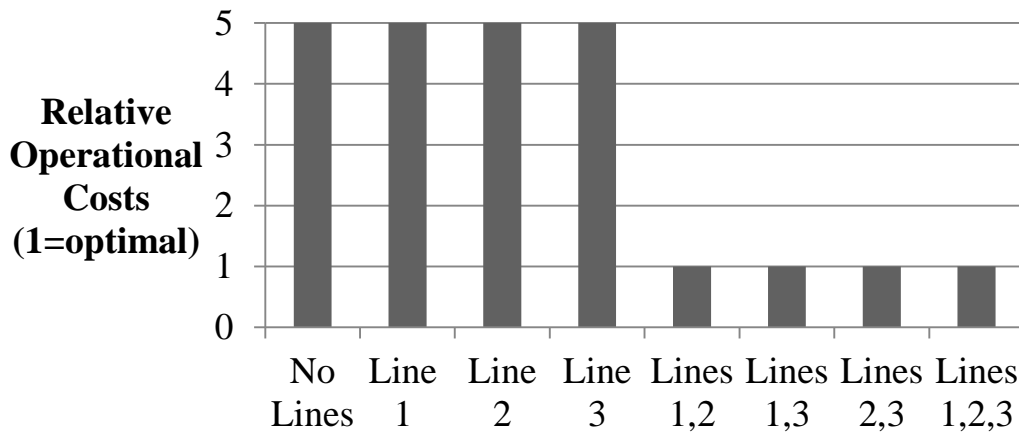


Figure 3-4 Relative Costs for Feature Identification in This Test System

This sensitivity is a complication for many standard methods. For example, one common way to characterize the difference between two binary strings is the Hamming Distance. The Hamming Distance quantifies the number of symbols which are different in a string. In transmission expansion planning, a binary string would describe which investments were made as in Figure 3-5 (1 = investment, 0 = no investment); however, as also shown in Figure 3-5 plans with equal Hamming Distance can have wildly different values. In Figure 3-5, both plans two and three have a Hamming Distance of two. Plan Two has a different combination of small lines than Plan One, but all demand is met and has a low cost. Plan Three, however, has no small connecting lines and demand is not met. Despite the difference in operational costs, the plans are indistinct using the Hamming Distance.

	East-West Line	Line 1	Line 2	Line 3	Hamming Distance from Plan One
Plan One	1	1	1	0	0
Plan Two	1	1	0	1	2
Plan Three	1	0	0	0	2

Figure 3-5 Hamming Distance Example

Regression methods likewise fail due to the networked nature of the problem. If, for example, the east-west line in Figure 3-3 was treated as an independent variable with the cost as a dependent variable, the regression would fail to recognize its value. Half of the plans (four of eight) with the line have very high costs. Capturing all of the dependencies necessary to accurately capture the value of the line quickly devolves into full enumeration of all line combinations, the dimensionality of the problem is not reduced and all states must be explored.

3.2 Dimensionality Reduction

The first challenge in building an approximation for the MS-TNEP problem is transforming the high dimensional problem to a lower dimensional space. This transformation is required because the dimensionality of the MS-TNEP problem prevents interpolation

based purely on the combinations of investments themselves. One of the most natural ways to capture the information in a transmission expansion plan is graphically, through a transmission map. Unlike mathematical network representations, such as node incidence matrices, transmission maps quickly and intuitively display patterns of investment. This work explores two novel dimensionality reduction techniques from the image recognition literature inspired by this insight.

Image recognition algorithms have been explicitly developed to transform data rich images into smaller sets of comparable summary data. These approaches allow for dimensionality reduction while still maintaining much of the information discussed above. Here, two image recognition approaches are explored, the Method of Moments and Principal Component Analysis. These two approaches are described in 3.2.1 and 3.2.2.

3.2.1 Method of Moments

The Method of Moments (MOM) is an image recognition technique with roots in the Optical Character Recognition literature. The method works by calculating the statistical moments of an image. These moments are then compared to a database of existing moment profiles to identify the letter or number. The statistical moments treat an image as a distribution and describe the image's axial symmetry, skew, etc. As a dimensionality reduction technique, the MOM satisfies the requirements presented in 3.1. MOM identifies patterns of investment based on a plan's image rather than on individual investments, and these patterns emerge from analyzing many plans rather than being identified a priori.

Much research on MOM has been on identifying invariant moments. These measures, such as Hu's moment invariants [30], allow images to be recognized regardless of rotation or scale. For the transmission planning problem, however, both rotation have meaning (i.e. an east-west line rotated 90 degrees is a north-south line and the two should not be equated). As a result, the simplest of moment formulations may be used. As shown in Eq. 29, each moment is characterized by its x degree (p), y degree (q) and total degree ($p+q$). While an arbitrarily high number of moments may be calculated, moments of

degree three are ordinarily sufficient to characterize an image. The weighting function inside the integral, $f(x,y)$, shown in Eq. 29, indicates where new transmission lines have been constructed.

$$\text{Eq. 29} \quad m_{p,q} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q f(x,y) dx dy$$

3.2.1.1 Interpretation of Moment Values

Each moment calculated is itself a scalar quantity but encodes physical information about the system. The moments of projection ($m_{p,0}$ and $m_{0,q}$) have intuitive meaning for the transmission expansion problem. These moments of projection up to order two are discussed briefly below both with their classical meaning in image recognition and the translation of that meaning to describing a transmission expansion plan. Moments with interaction terms ($p>0$ and $q>0$) and greater than order three do not have the same intuitive meanings appealing in the lower order moments, and while only five moments are discussed below, all ten moments up to order three ($p+q\leq 3$) were considered for dimensionality reduction.

The zeroth moment ($m_{0,0}$) describes the shaded area of a black and white image. For the transmission expansion planning problem, the weighting function, $f(x,y)$, can be scaled such that the moment of degree zero indicates the total MW-miles of new transmission miles. For this weighting function, the weighting for each transmission line is the capacity divided by length. The weighting function can also be scaled to reflect the different line ratings (e.g. 765kV weighting >500kV weighting >345kV weighting); however, alternate weighting functions will not have the same physical meaning. The technical procedure used to transform an expansion plan into an image is given in Appendix A.

The first moments ($m_{1,0}, m_{0,1}$) describe the center of mass of an image or alternatively the axial symmetry of the image. For transmission planning, a positive first moment with respect to the x-axis ($m_{0,1}$) indicates that the center of mass of the new transmission is in the northern quadrants of the map. More simplistically, it indicates that more new trans-

mission is in the north than the south. Likewise, a positive first moment with respect to the y axis ($m_{1,0}$) indicates more investment in the east than west (and vice versa for a negative first moment). A transmission expansion plan with a first moment ($m_{0,1}$) equal to zero would have equal megawatts of capacity added both in the north and south

The second order projection moments ($m_{0,2}$, and $m_{2,0}$) indicate the moments of inertia of the image. Thinking about this applied to transmission expansion planning, it indicates the centrality of investment to the axes. A low second order moment indicates investment is near the axis considered (low moment of inertia about the axis). A high second order moment indicates investment is far from the axis (high moment of inertia about the axis).

3.2.2 Principal Component Analysis

Developed by Karl Pearson in 1901, Principal Component Analysis (PCA) is a tool to identify patterns in high dimensional data and then store those patterns in a lower dimensional form. The method identifies patterns through analysis of the problem's covariance matrix. The patterns are summarized by the principal components or eigenvectors of the covariance matrix. Each principal component is orthogonal and captures a descending quantity of variance in the data; the dimensionality of the data can then be reduced by tracking a reduced number of principal components.

PCA is commonly used as an image recognition tool and to compress images described by thousands of pixels into lower dimensional storage. Applied to the transmission expansion planning problem, PCA identifies patterns of investment, and the ultimate goal is to correlate these patterns of investment with expected costs. Rather than literally transform the transmission map into an image, investment matrices are analyzed.

The first step of applying PCA to the MS-TNEP problem is to construct the matrix for analysis. For the MS-TNEP problem, the matrix to be analyzed is the investment matrix. Each row of the investment matrix represents a specific expansion plan and each column represents a specific investment. As shown in Table 3-1 and Table 3-2, the investment matrix can be represented as a binary matrix or a capacity based matrix. The binary representation in Table 3-1 treats each line equivalently, regardless of size. The

capacity representation in Table 3-2, on the other hand, emphasizes patterns in higher capacity lines.

Table 3-1 Binary Matrix Representation

	Line 1	Line 2	Line 3	Line 4
Plan 1	1	0	1	0
Plan 2	0	1	0	1
Plan 3	1	1	0	0
Plan 4	1	0	1	1
Plan 5	0	0	0	0

Table 3-2 Capacity Matrix Representation

	Line 1	Line 2	Line 3	Line 4
Plan 1	900	0	100	0
Plan 2	0	100	0	100
Plan 3	900	100	0	0
Plan 4	900	0	100	100
Plan 5	0	0	0	0

The covariance matrix is a compact manner to organize and analyze relationships between various data sets. The covariance matrix for Table 3-1 is shown in Table 3-3 and covariance matrix is symmetric about the diagonal, and for clarity only the upper triangle is shown. The diagonal entries of Table 3-3 show the statistical variance for a specific line and the off-diagonal entries reflect the statistical covariance between each of the lines. For example, the construction of Line 1 and Line 3 in Table 3-1 are related; both are constructed in Plan 1 and Plan 4. In Table 3-3, this relationship is shown as a positive covariance of 0.2. On the other hand, in the plans here the incidence of Line 3 and Line 3 are anti-correlated; Line 2 is only constructed in Plan 1 and Plan 4 while Line 3 is only constructed in Plan 2 and Plan 3. This relationship is shown in the covariance matrix with a negative score of -0.2. Rather than looking for patterns across what would be an exceptionally large covariance matrix, the data is summarized through the eigenvectors of the covariance matrix also known as principal components.

Table 3-3 Covariance Matrix for Table 3-1*Data Centered Before Calculation*

	Line 1	Line 2	Line 3	Line 4
Line 1	0.3	-0.05	0.2	-0.05
Line 2		0.3	-0.2	0.05
Line 3			0.3	0.05
Line 4				0.3

The principal components for the full problem can be used to reconstruct the initial data set. Each principal component, shown in Table 3-4, captures a certain quantity of the variation, and the quantity of this variation captured is reflected in the eigenvalues, which are calculated alongside the eigenvectors. For the test problem here, the individual principal components shown in Table 3-5, individually capture at most 51% of the variation in the data and as little as 1%.

The quality of the reconstructed data improves with the number of principal components used for the reconstruction. As shown in Table 3-6, reconstructing the data in Table 3-1 using only the first principal component captures two of the major patterns. In the plans which exhibit the patterns with high covariance factors (Plan 1, Plan 2, Plan 4) have Lines 1, 2, and 3 well represented. Plans 3 and 5 as well as Line 4 throughout do not contain these patterns and are poorly represented. For example, in the reconstructed data, Line 4 has values which range from 0.3 to 0.6 rather than near an integer value. Adding the second principal component, shown in Table 3-7, however, accounts for the variation in Line 4 and provides a much better representation. If all principal components are used, the investment matrix may be reconstructed perfectly. Thus, a trade-off must be made between the number of dimensions used and the accuracy of the approximation.

Table 3-4 Principal Components for PCA Test System

PC 1	PC 2	PC 3	PC 4
0.53	-0.10	0.71	-0.46
-0.53	0.10	0.71	0.46
0.67	0.25	0.00	0.70
-0.062	0.96	0.00	-0.29

Table 3-5 Eigenvalues for PCA Test System

	Eigenvalue	Percentage Variation Explained
PC 1	0.61	51%
PC 2	0.32	27%
PC 3	0.25	21%
PC 4	0.02	1%

Table 3-6 Reconstructed Test PCA Data Using One Principal Component

	Line 1	Line 2	Line 3	Line 4
Plan 1	1.0	0.0	1.0	0.3
Plan 2	0.1	0.9	-0.2	0.5
Plan 3	0.4	0.6	0.2	0.4
Plan 4	1.0	0.0	0.9	0.4
Plan 5	0.4	0.6	0.2	0.4

Table 3-7 Reconstructed Test PCA Data Using Two Principal Components

	Line 1	Line 2	Line 3	Line 4
Plan 1	1.1	-0.1	0.9	0.0
Plan 2	0.0	1.0	-0.1	1.0
Plan 3	0.5	0.5	0.1	0.0
Plan 4	0.9	0.1	1.1	1.0
Plan 5	0.5	0.5	0.1	0.0

3.3 Interpolation

The two dimensionality reduction techniques overviewed above reduce the number of predictive variables to describe the relationship between a transmission plan and its cost. By themselves, however, these methods do not provide a relationship between the predictive variables and expected cost of a transmission expansion plan. This relationship is given by an interpolation method. An interpolation method takes the set of variables provided by the dimensionality reduction techniques and creates a reduced order model to predict the expected cost of new transmission plans. Two interpolation methods are proposed and explored in this work: nearest neighbor and linear regression. The relationships in these models are not pre-specified by the modeler but instead arise from training the interpolation method on the cost of known transmission plans and updating the models with more information as it is generated.

The first method, nearest neighbors, is the most natural for object recognition techniques. Historically, object recognition algorithms have been used to classify an unknown object into categories of known objects [13]. In the character recognition algorithms, for example, an unknown character is compared to each of known letters of the alphabet. The

unknown character is then matched to the closest known letter. The determination of which characterized plan is the nearest neighbor of a proposed plan is made by computing the Euclidian distance. The Euclidian distance between two plans is the ordinary distance between the estimators of each plan. Any number of nearest neighbors can be used to estimate the cost of a new plan. If a single nearest neighbor is used, the new plan is estimated to have the same cost as that nearest neighbor. If more than one nearest neighbor is used, the cost can be estimated using an unweighted mean cost of the neighbors or by weighting the neighbors inversely proportional to distance (e.g. close neighbors have higher weights and further neighbors have lower weights).

The second approach uses traditional linear regression. With this approach, the variables produced by the object recognition techniques are the explanatory variables and the expected costs are the dependent variable. The estimated cost of a new plan is then a function of the linear regression coefficients determined using the characterized plans.

3.4 Evaluation of Dimensionality Reduction and Interpolation Techniques

The combined dimensionality reduction and interpolation methods can be used to guide a search in multiple ways. The predicted costs could be used globally to select the overall optimal plan. The predicted costs could also be used in a more limited manner, for example to decide between two branches in a decision tree. Tests for both of these possible uses are included in 3.4.1. and 3.4.2. All methods are evaluated using a set of 1,000 transmission plans were created using the algorithm outlined in Appendix A. Of these 1,000 plans, 750 were used to train the dimensionality reduction and interpolation techniques and the remaining 250 were used to test each method.

The Western Electric Coordinating Council (WECC) 240 bus model discussed in Chapter 2 is also used as the test system here. For initial exploration of the dimensionality reduction techniques, the test system has been simplified to a single peak load hour and a single time horizon of 10 years out. The set of 220 reinforcement corridors, corridors which already connect existing nodes, are considered in the model and each corridor is rated physically and financially at a 500kV single circuit capacity using the St. Clair

curve described in Chapter 2 and a cost of \$2.9 million/mile. Uncertainty in generation development was again introduced through probabilistic generation developed at specific Western Renewable Energy Zones (WREZs). In this test system, generation development was only considered at the 10 largest WREZs and each WREZ was assumed to connect, without congestion, into its nearest neighbor WECC node. These WREZs are shown in green in, and the corresponding WECC nodes are shown in red. Two pairs of WREZs share a nearest neighbor, leaving eight unique WECC nodes with potential WREZ power injections.

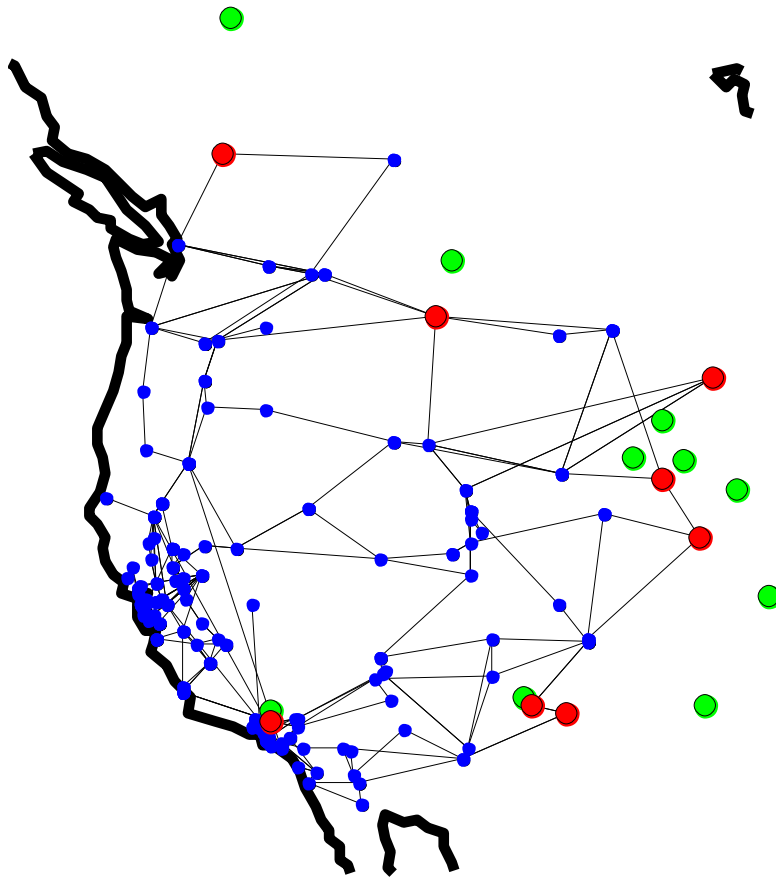


Figure 3-6 Test System for Value Function Approximation Testing

3.4.1 Cost Prediction

The first set of trials test the ability of each approximation method to predict the total expected cost of a transmission expansion plan. These costs include both the operational and investment costs. Rather than compare the predicted cost to the true cost directly, a normalized mean (mean of costs minus the minimum expected total cost) was used to evaluate predictions. The normalized mean was used because there is a high minimum operational cost required to meet demand. The variation in plans compared relative to this minimum cost is small (1,189 million USD compared to 14,434 million USD); however, the variation itself remains large in absolute terms, over one billion USD.

3.4.1.1 Moment Results

Six different prediction techniques using the method of moments were explored to predict the cost of each plan. These permutations, shown in Table 3-8, vary both by the number of moments and the number of neighbors used. For the trials using 10 moments, all moments up to those of degree three are used. For comparison, the trials using six moment used only moments up to order two (shown in Table 3-9). All nearest neighbor trials treated neighbors equally and were not weighted by distance. As shown in Table 3-8, the average errors range from 252 million USD (21.1%) to 409 million USD (34.4%). Generally, the methods using 10 moments rather than 6 performed better; however, this improvement is marginal. These errors are likely sufficiently large not to allow differentiation globally between plans; 78 of all 250 samples (31%) lie within 21% of the normalized mean value.

Table 3-8 Cost Prediction: Moment Results

Nearest Neighbor	Number of Moments	Number of Neighbors	Absolute Average Error (M-USD)	Percentage Average Error of Normalized Mean
	<i>10</i>	<i>5</i>	<i>328</i>	<i>28%</i>
	<i>10</i>	<i>1</i>	<i>399</i>	<i>34%</i>
	<i>6</i>	<i>5</i>	<i>353</i>	<i>30%</i>
	<i>6</i>	<i>1</i>	<i>409</i>	<i>34%</i>
Regression	Number of Moments		Absolute Average Error (M-USD)	Percentage Average Error of Normalized Mean
	<i>10</i>		<i>252</i>	<i>21%</i>
	<i>6</i>		<i>282</i>	<i>24%</i>

Table 3-9 Moments Used in Six Moment Set

Total Degree	0	1	1	2	2	2
p	0	0	1	1	0	2
q	0	1	0	1	2	0

3.4.1.2 PCA Results

As with the method of moments, multiple combinations of PCA outputs and interpolation techniques were tested. As shown in Table 11, PCA analysis was performed on both the capacity and binary representations of the investment matrices. Each PCA analysis produces 220 eigenvectors (equal to the number of investment variables); however, for both the binary and capacity representations, only the top ten eigenvectors were used. Thus the dimensionality for all PCA techniques was reduced from 220 to 10.

For both the nearest neighbor and regression interpolation techniques, the distances and independent variables are the product of the eigenvector and the plan. This produces a scalar quantity. In regression and the unweighted nearest neighbor interpolation techniques, these variables are used directly as the independent variables. Unlike the moments, where there is no intuitive reason to weight one more than the other, each eigenvector resulting from the PCA analysis also produces an eigenvalue weighting factor. To explore whether or not these weighting factors improve the interpolation, nearest neighbors were explored using these weights as shown in Eq. 30.

$$\text{Eq. 30} \quad \text{coeff}_i = \text{eig_val}_i \cdot \text{eig_vector}_i \cdot \text{plan}$$

The errors for the PCA trials are generally lower than those for the MOM trials. The mean errors for the predicted costs range from a minimum of 13.5% to a maximum of 27.0% with three combinations achieving 14% average error. Unlike MOM, this error may be sufficiently small to allow differentiation amongst plans globally.

Table 3-10 Cost Prediction: PCA Results

Nearest Neighbor	Eigenvalues	Number of Neighbors	Absolute Average Error (M-USD)	Percentage Average Error of Normalized Mean
	<i>Binary, W</i>	<i>5</i>	<i>164</i>	<i>14%</i>
	<i>Binary, UW</i>	<i>5</i>	<i>163</i>	<i>14%</i>
	<i>Binary, W</i>	<i>1</i>	<i>213</i>	<i>18%</i>
	<i>Binary, UW</i>	<i>1</i>	<i>223</i>	<i>18%</i>
	<i>Capacity, W</i>	<i>5</i>	<i>232</i>	<i>20%</i>
	<i>Capacity, UW</i>	<i>5</i>	<i>250</i>	<i>21%</i>
	<i>Capacity, W</i>	<i>1</i>	<i>282</i>	<i>24%</i>
	<i>Capacity, UW</i>	<i>1</i>	<i>322</i>	<i>27%</i>
Regression	Eigenvalues		Absolute Average Error (M-USD)	Percentage Average Error of Normalized Mean
	<i>Binary</i>		<i>161</i>	<i>14%</i>
	<i>Capacity</i>		<i>266</i>	<i>22%</i>

3.4.2 Ranking Prediction

The ranking prediction trials tested the approximation’s ability to correctly predict the lower cost of two plans. For these trials, the actual cost predicted is not as important as long as the relative ranking of the plans is correct. To test each approximation’s ability to correctly predict lower cost plans, 500 random pairs of the 250 test plans were sampled. The true lower cost plan was identified and each technique was evaluated on its ability to select these lower cost plans.

3.4.2.1 Moment Results

The results for the moment trials, given in Table 3-11, indicate that all combinations select the lower cost plan in at least 74% and up to 81% of trials. These techniques all perform significantly better than random and reduce the dimensionality of the problem from 2^{220} to at most 10 moment values. As with the cost prediction trials, the ranking trials using ten moments generally performed better than the trials using six moments, however, the differences were again marginal.

Table 3-11 Moment Results

Nearest Neighbor	Number of Moments	Number of Neighbors	Number Correctly Predicted	Percentage Correctly Predicted
	10	5	403	81%
	10	1	372	74%
	6	5	382	76%
	6	1	379	76%
Regression	Number of Moments		Number Correctly Predicted	Percentage Correctly Predicted
	10		385	77%
	6		377	75%

3.4.2.2 PCA Results

The PCA results for ranking mirror the PCA results for predicting values. Overall, the PCA trials predicted between 71% and 87% of the pairs correctly. While the worst performing PCA method was outperformed by MOM, the top three performing methods, all based binary investment matrices, each outperformed the best MOM combination. As with the prior results, predictions improved with increasing numbers of neighbors. The only case where this trend did not hold was in the weighted capacity based investment matrices, where both methods predicted 78% of pairs correctly.

Table 3-12 PCA Results Trials

UW: unweighted by eigenvalues, W: weighted by eigenvalues

Nearest Neighbor	Eigenvalues	Number of Neighbors	Number Correctly Predicted	Percentage Correctly Predicted
	<i>Binary, W</i>	5	415	83%
	<i>Binary, UW</i>	5	421	84%
	<i>Binary, W</i>	1	396	79%
	<i>Binary, UW</i>	1	377	75%
	<i>Capacity, W</i>	5	388	78%
	<i>Capacity, UW</i>	5	387	77%
	<i>Capacity, W</i>	1	390	78%
	<i>Capacity, UW</i>	1	355	71%
Regression	Eigenvalues		Number Correctly Predicted	Percentage Correctly Predicted
	<i>Binary</i>		436	87%
	<i>Capacity</i>		365	73%

3.5 Conclusions

Heuristic methods must be investigated for the multi-stage stochastic transmission expansion planning problem as the problem is computationally intractable using traditional optimization methods. Rather than focus on meta-heuristics which are blind to problem specific characteristics, this work explores creating a reduced order model or approximation of the MS-TNEP problem. Developing an approximation for the transmission expansion planning problem is specifically challenging due to a large number of categorical and integer variables (one for each type transmission line investment between two buses).

This chapter proposed combining dimensionality reduction techniques from the image processing literature and two interpolation techniques to form a reduced order model. Two methods from the image processing literature, Principal Component Analysis (PCA) and Method of Moments (MOM), were introduced to reduce the dimensionality from the number of network configurations (2^n where n is the number of investment options) to 5 to 10 dimensions. These reduced dimensional spaces were then combined with a nearest neighbors or a linear regression approach to interpolate between the system-costs (transmission cost and generation costs) of known transmission expansion plans to new transmission expansion plans.

In order to test the dimensionality reduction and interpolation techniques, a new method was developed to produce a sample set of transmission expansion plans. The method combines depth first search for non-served energy, generation sampling, and traditional optimization methods to ensure both diversity of transmission expansion plans and low expected total costs. The method was tested on a 240 bus model of WECC to successfully produce 1,000 unique transmission expansion plans with no non-served energy. The different combinations of dimensionality reduction techniques were tested by training on the first 750 plans and then predicted costs of the remaining 250 plans and ranking between 500 pairs of the remaining 250 plans.

Both the PCA and MOM approaches have strong potential to be used in guiding searches. The best combinations of interpolation and dimensionality techniques were able

to identify greater than 80% of rankings correctly. The highest performing combination, PCA and regression, identified 87% of rankings correctly. Generally, the PCA methods performed better than the MOM methods; however, the differences were not significant. In the cost prediction trials, however, PCA methods out performed MOM. The errors for MOM ranged from a minimum of 21% to a maximum of 34%, likely too high to accurately discriminate between plans. Errors for the PCA methods ranged from 14% to 27% which may be sufficiently small to discriminate between plans. The true test, however, for each method would be to embed it in a heuristic algorithm.

3.6 Future Work

The methods presented here were tested on a system expanded using a single voltage level and only allowed a single investment in each corridor. Further testing should be done on systems which allow for multiple voltages of transmission line and multiple investments per corridor. Additionally, further technical details should be integrated into the testing problem, for example losses and an assessment of n-1 reliability. As above, the true test for these approximation methods will be to embed them in heuristic optimization methods.

There are many possible refinements to the work presented here. For example, in the PCA trials, 10 principal components were used. Greater differentiation may be gained using more principal components. Likewise, greater than 10 moments may be explored for MOM. There are also additional trade-offs to explore, such as the need to recalculate principal components against the general lower efficacy of MOM, which has static interpolation coefficients.

The work presented in this chapter is an initial exploration into problem specific approximate value functions for the transmission expansion planning problem. Two image processing techniques were explored; however, there are many other problem dimensionality reduction techniques which may be useful in this context. For example, the image processing techniques are based on geographic distances, however, exploring electrical distances may perform better in the transmission problem [15]. Additional interpolation

techniques, such as Moving Least Squares, may also be able to better capture the non-linear characteristics of the problem (see for example [67]).

THIS PAGE INTENTIONALLY BLANK

4. STEP Model: An Approximate Dynamic Programming Approach for Multi-Stage Stochastic Transmission Network Expansion Planning

Thus far this work has developed a screening model to reduce the dimensionality of the multi-stage stochastic transmission network expansion planning problem (MS-TNEP) in Chapter 2 and explored interpolation techniques to guide a heuristic search in Chapter 3. This chapter builds on the insights and methods of these previous chapters to develop a heuristic search algorithm to solve the MS-TNEP problem, titled the STEP (Stochastic Transmission Expansion Planning) model. The search algorithm proposed here uses the framework of Approximate Dynamic Programming (ADP). ADP avoids the curses of dimensionality that plague other solution methods through Monte Carlo methods and value function approximation.

This chapter proceeds in three major sections. First, 4.1 describes the Approximate Dynamic Programming family of methods. Second, 4.2 works through the development of the STEP model. Finally, the 4.3 explores the STEP model on a small demonstration system. Conclusions and future research are given in 4.4.

4.1 Approximate Dynamic Programming

Approximate Dynamic Programming (ADP) is an extension of Dynamic Programming (DP) methods. The general approach of DP and ADP is to break a complex problem into many less-complicated subproblems. Each subproblem describes a discrete state of the system (e.g. for transmission planning a state would include the set of constructed transmission lines, set of generators, fuel prices, and demand levels). These states are connected as Markov chains, with each state containing all information required to make future decisions. Because of this general approach, DP is very flexible and naturally lends itself to both integer and non-linear problems.

In the DP framework for decision making under uncertainty each sub-problems represents the cost or value contribution for each discrete decision-stage. The decision-stages (or time steps) are then connected via the Bellman equation and solved using backward recursion. The Bellman equation, shown in Eq. 31, describes the value of being in a

specific state as the sum of the current value of the state and the expected future value. That is, the value of a decision today depends both the cost incurred today and the costs incurred tomorrow. This is especially important in the transmission expansion planning context because the decisions are irreversible. A transmission line constructed today will not be taken out of service in ten years. As a result, constructing a new line today limits the future options to only plans which include that new line.

$$\mathbf{Eq. 31} \quad V_t(S_t) = \min_{x_t \in X_t} C_t(S_t) + \alpha \cdot (E[V_t(S_{t+1}|S_{t+1})])$$

S_t	state variable
C_t	current stage costs
x_t	actions
α	discount factor

With a Markov state-based solution system, DP can also include the endogenous relationship between generation and transmission. For example, if a large line is constructed to a new generation area, the probability of generation development matching the capacity of that new line can be increased. If on the other hand no line is constructed to that new generation area, the probability of new generation can be set to zero. These are relationships that are natural to model in the DP framework but cannot be modeled in other frameworks such as stochastic programming.

The major impediment to using the DP approach directly is the thoroughly characterized curse of dimensionality [47][51]. The curse of dimensionality can be characterized in the state, action and uncertainty states. In the MS-TNEP problem, the size of the state space for each stage is 2^n with n again describing the number of investments. The size of this state space for each decision stage then increases linearly with the number of uncertainty scenarios. Likewise, the action space, or set of all potential decisions for a given state, will also have a size of 2^n . These effects multiply and make the problem computationally intractable. For a single stage problem considering 1,000 investments, there are

$2^{1000} \times 2^{1000}$ state action pairs to be characterized. Again, this problem is realistically infeasible to solve due to computational time and memory constraints.

A relatively new family of methods, Approximate Dynamic Programming (ADP), overcome these dimensionality problems through a combination of Monte Carlo methods and value function approximation [7],[47]. While the traditional DP solution methods require recursively solving every possible state, action and uncertainty combination, ADP methods iteratively sample state, action and uncertainty combinations to build a reduced order model of the value function. The ADP model then optimizes this reduced order model, more commonly called the value function approximation. Both the strength and weakness of ADP approach is its flexibility. Rather than a set algorithm which can be treated as a black box solver, ADP is a flexible approach which must be adapted to each individual context. Using an ADP approach, previously intractable may be solved; however, a new problem-specific algorithm will need to be developed. As with any heuristic method, the solution found by an ADP algorithm will not be provably optimal; however, the problems are intractable using methods with provably optimal solutions.

4.1.1 Approximate Dynamic Programming Framework

While there is great variety in ADP models, a common framework is a double-pass algorithm. A double-pass algorithm consists of a forward pass, sampling forward in time through the decision tree and a backward pass, updating Bellman values using the information gained through the forward pass. This generic double-pass algorithm is described in Table 4-1.

During the forward pass of the algorithm, demonstrated in Figure 4-1, a path is sampled through the decision tree. In each stage, a decision is sampled and then an uncertainty realization is sampled. If an endogenous relationship between the generation and transmission is being modeled, the uncertainty realization (location and quantity of new generation) sampled will be affected by the transmission investment decisions made.

In the backward pass, the decisions sampled states from the forward pass are evaluated. Starting with the last time stage, the costs incurred in each time stage are evaluated. Once the costs incurred in some stage t are evaluated, the current estimate of Bellman

cost for stage $t-1$ can also be updated. For the MS-TNEP problem, the costs incurred in each decision stage are both the operational costs (generation costs plus non-served energy costs) and the investment costs for all transmission lines constructed up to that point.

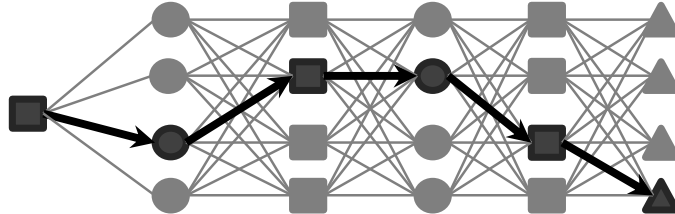


Figure 4-1 Illustration of Forward Pass in ADP Algorithm

Table 4-1 Generic Description of ADP Double-Pass Algorithm

<i>Step Description</i>	<i>Variable/Equation</i>	<i>Step in MS-TNEP Problem</i>
Forward Pass		
<i>for</i> $t = 1, 2, \dots T$;		
1. Sample Decision	x_t^i	<i>Sample new transmission lines</i>
2. Sample Uncertainty	ω_t^i	<i>Sample new generation</i>
3. Update State Variable	S_t^i	<i>State: All transmission lines (including new) and all generation (including new)</i>
Backward Pass		
<i>for</i> $t = T, T - 1, \dots 1$;		
4. Evaluate Current Stage Costs	$C_t^i(S_t^i)$	<i>Evaluate generation and investment costs</i>
5. Update Estimated Bellman Cost for the Current State	$\hat{v}_t^i(S_t^i) = C_t^i + \alpha \cdot \hat{v}_{t+1}^i(S_{t+1}^i)$	<i>Update Bellman cost for specific plan and generation scenario</i>
6. Update Approximation of Value Function	\bar{C}_t^i	

Indices

- i iteration
- t stage (time stage)

Variables

S_t^i	state variable
C_t^i	current stage costs
\bar{C}_t^i	approximate value function
\hat{v}_t^i	estimate of Bellman cost

Constants

α	discount factor
----------	-----------------

4.1.2 Exploration and Exploitation Phases of ADP

While an approximate value function, or reduced order model, is refined over each iteration, there are often two distinct phases of an ADP model, explore and exploit. During the exploration phase of the model, the goal is to gather information to build the value function approximation. To characterize the value function approximation, a wide variety of decisions and uncertainties are explored. The second phase, exploitation, uses the model developed during the exploration phase to find the actions which result in the highest Bellman values.

Consider the simple one-stage stochastic illustrative problem given in Figure 4-2. In this simple problem, the goal is to minimize the expected value of the first stage decision. There are only three possible actions and two equally weighted probability scenarios for a total of six states. The problem structure and outcomes are shown in Figure 4-2.

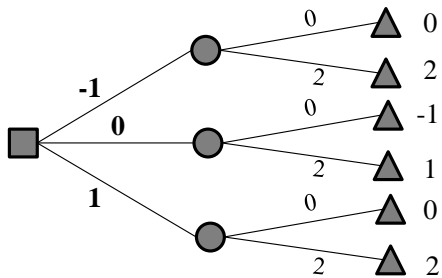


Figure 4-2 Decision Tree for Sample Problem

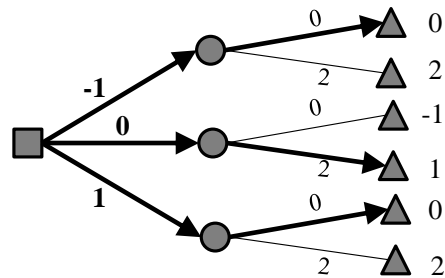


Figure 4-3 Exploration Phase Sample Paths Through the Tree

Six iterations of the algorithm and the resulting value estimates are shown Figure 4-3. The first three iterations of the algorithm are the exploration phase; each of the actions is sampled and an estimate of its value is based on the sampled uncertainty state. The second three iterations of the algorithm are the exploit phase. During these iterations, the optimal decision is selected as shown in Figure 4-5. With each iteration, the estimates of each action's value changes (Figure 4-4) and as a result the optimal action also changes (Figure 4-5). In the illustrative example, all states are explored; however, in an ADP algorithm, interpolation is used to estimate value of states not yet visited.

x	$\hat{v}_i(x)$					
	i=1	i=2	i=3	i=4	i=5	i=6
-1	0	0	0	1	1	1
0		1	1	1	1	0
1			0	0	1	1

Figure 4-4 Estimated Expected Values for ADP Example Problem

Values during updated during each iteration are bolded

	i=4	i=5	i=6
x^*	-1	1	0
ω_i	2	-1	2

Figure 4-5 Optimal Actions and Uncertainty Outcomes for the ADP Example Problem

4.2 ADP for the MS-TNEP Problem

The STEP model, developed in the following sections, is based on the Approximate Dynamic Programming framework and includes one of the approximation techniques discussed in Chapter 3, Method of Moments. As a result of its structure, developing an ADP model using Method of Moments requires adaptation of the double-pass algorithm outlined in Table 4-1. This section is laid out in four sections. The first section, 4.2.1, describes the specific problem the STEP model is structured around. The second section, 4.2.2, describes how the interpolation techniques explored in Chapter 3 are integrated into the ADP framework. The actual algorithm is outlined in two parts. The exploration phase of the algorithm is described in 4.2.3 and the exploitation phase is described in 4.2.4.

4.2.1 Problem Structure

The STEP model in the following section is explicitly formulated for a two-stage stochastic transmission expansion planning problem. The problem, diagrammed in Figure 4-6, has two decision stages and a single uncertainty stage. The first-stage decision, the new transmission investments made today, is made in the face of uncertain future generation development. The second-stage investment decision is made once the generation uncertainty is realized. This second stage decision is the recourse in the problem. For both development work and ease of explanation this algorithm is structured for two stages, but it could be expanded to an arbitrary number of decision stages.

As shown in Figure 4-6, the optimal solution structure for the problem posed has a single first-stage plan, p_1^* , and multiple optimal second-stage plans dependent on the realization of generation uncertainty, $p_{2|\omega_n}^*$. The first-stage plans represent the best decisions made in the face of uncertainty. In the second-stage, additional lines are selected to adapt the network to the realized generation uncertainty. As also shown in Figure 4-6, the state variable that the algorithm will be optimizing is the post-decision state variable. For transmission expansion planning problems, the post-decision state variable reflects the build-out of the transmission system after new lines have been constructed but before the uncertainty is realized. In the two-stage problem, the first-stage state-variable, S_1 , contains only the new transmission build-out. In the second stage, the state-variable, S_2 , includes both the complete transmission build-out and the generation uncertainty realization as the optimal second stage plan will depend on where generation is constructed.

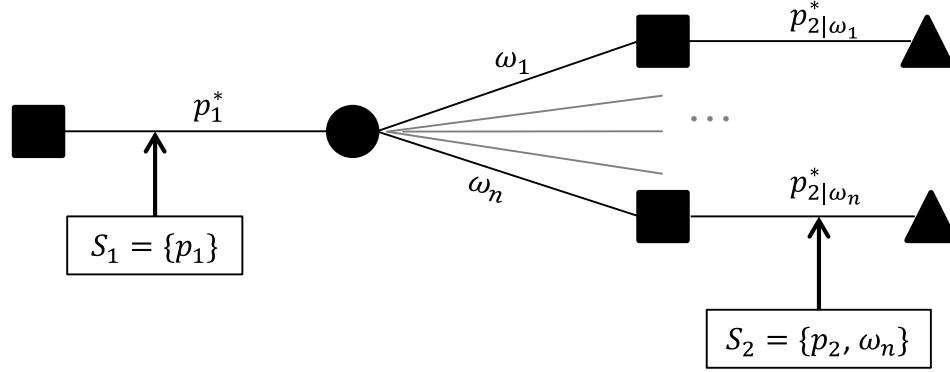


Figure 4-6 Optimal Solution Structure and State Variables

The algorithm presented in Table 4-1 minimizes the Bellman costs for first-stage decisions. The classical form of the Bellman equation, balancing current and expected future costs, is repeated in Eq. 32. For the MS-TNEP problem, the Bellman costs represent minimizing the sum of today's investment costs, expected operational costs and future investment and operational costs as shown in Eq. 33. For the two stage problem presented in Figure 4-6, the second stage costs are deterministic rather than stochastic, as shown in Eq. 35. As a result, Eq. 33 can also be simplified in the first-stage to remove the expectation as shown in Eq. 34.

Eq. 32 $V_t(S_t) = \min_{x_t \in X_t} C_t(S_t) + \alpha \cdot (E[V_t(S_{t+1}|S_{t+1})])$

Eq. 33 $V_t(S_t) = IC(S_t) + E[OC(S_t)] + \alpha \cdot \min E[(IC(S_{t+1}|S_t) + E[OC(S_{t+1}|S_t)])]$

Eq. 34 $V_1(S_1) = IC(S_1) + E[OC(S_1)] + \alpha \cdot \min E[IC(S_2|S_1) + OC(S_2|S_1)]$

Eq. 35 $V_2(S_2) = IC(S_2) + OC(S_2)$

V_t	Bellman value
S_t	State variable
α	Discount factor
$IC(S_t)$	Investment costs

4.2.2 Approximate Value Function

Most approximation methods operate in the same space as the decision variables. For example, a regression is performed on the decision variables as the independent variables or the hamming distance is calculated using strings of decision variables. With these methods, the dimensionality issues addressed may be caused by the range of values each decision variable can take. In the transmission expansion planning, this problem is reversed. Each investment variable can take on only one of two values, one or zero; however, there may be thousands of variables. As discussed in Chapter 3, these variables are not separately-additive and cannot be considered independently. Due to these issues, the value function approximation using the methods in Chapter 3 is performed in a reduced dimensionality space.

The transformation from expansion plans into the lower dimensional states are straightforward calculates. There are not, however, direct transformations from the low dimensionality spaces back to the high dimensionality spaces. For example, if the value function approximation uses Method of Moments and regression, the resulting regression equation can be optimized, but there is no way to translate directly from the moments the optimization identifies back into a transmission expansion plan. Instead, the approach taken in the STEP algorithm is to use interpolation to select the best candidate plan among a small set. This small set is generated using MILP as will be discussed in 4.2.4.

4.2.3 Exploration Phase

The first part of the STEP model is designed to identify areas of the solution space where demand is met under generation scenarios. In the explore-exploit terminology, this is the exploration phase of the algorithm. No attempt is made to find the globally optimal solution; instead, individual plans are identified in each stage which can be built upon during the exploit phase of the algorithm. As discussed in Chapter 3, the vast majority of randomly selected expansion plans result in high non-served energy costs. To avoid this issue, the exploration phase of the ADP algorithm proposed here uses MILP techniques

to identify transmission plans with low to zero expected non-served energy costs. Using the canyon analogy from Chapter 3, this phase identifies points on the canyon floor by optimizing specific scenarios. It then explores nearby areas of the canyon floor to develop a map of the low cost spaces.

The exploration phase of the model proceeds in four steps, outlined in Figure 4-7. The first step is the identification of optimal two-stage (or more generally n-stage) transmission expansion plans for scenarios. The second step is the identification of additional plans with low to zero expected energy non-served (EENS) constructed with limited sets of lines. For clarity, these plans with low to zero non-served energy will be referred to as valid plans. This terminology is used to indicate that these plans may be considered by the system planner; it does not indicate that plans outside this set are infeasible from an optimization perspective. With valid plans identified, the third step is to calculate estimates of expected operational costs for each state. Finally, with current-stage contribution cost estimates for each state, Bellman values are calculated.

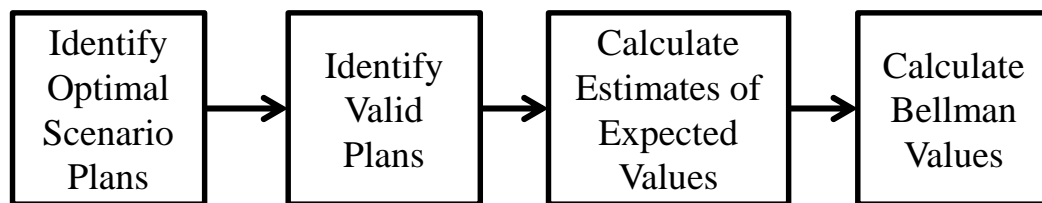


Figure 4-7 Four Phases of the Explorative Phase

The first step of the explore algorithm is to identify optimal plans for generation expansion scenarios. This is a straight-forward procedure where a generation expansion scenario is randomly sampled and a two-stage MILP optimization is solved for the deterministic scenario all potential transmission investments. Because the goal in this stage of the algorithm is to explore the solution space rather than find the optimal plan, the optimality criteria (e.g. MIP gap) can be very loose, for example 10%. The plans identified during the first step of the explore-phase are relatively computationally expensive. These plans are computationally expensive to identify because the full set of transmission investments are considered.

In the second step of the exploration phase, the set of candidate transmission lines is reduced to lessen the computational burden of exploring the solution space and also to force diversity into the resulting plans. During this step, shown graphically in Figure 4-8, valid plans for scenarios are identified as in step one. In step two, however, the full set of transmission lines is not considered for investment. Instead, a subset of candidate transmission lines is developed based on an existing plan. This is analogous to starting from a known position on the canyon floor and exploring in a random direction. First, a random plan from step one is selected and the investments in this plan become the set of candidate transmission investments. Second, a subset of candidate investments are randomly removed and replaced with investments from the full set of plans. This replacement guarantees that the same transmission plan will not be repeated in both step one and step two. Finally, these transmission investments are optimized for a random generation expansion scenario.

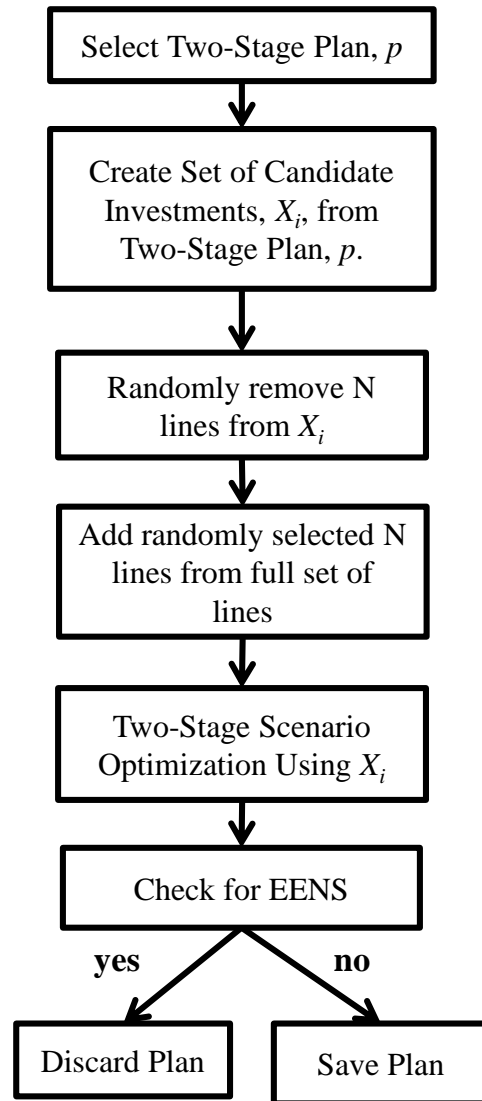


Figure 4-8 One Iteration of the Second Step of the Exploration Phase

It is important to note here that the candidate transmission lines must not all be selected for investment in the step two plans. Instead, the candidate lines are the set from which the two-stage optimization may select. In essence, the optimization is used to identify a reasonable plan from a subset of transmission investments. Rather than randomly selecting lines, however, the candidate set is based on a known valid plan.

The valuation of the plans in steps one and two were based on a two-stage scenario. The problem at hand, however, is stochastic. In order to value the plans identified in step two, the operational costs are evaluated across randomly sampled generation scenarios. These evaluations are completed using a DC optimal power flow which is computationally inexpensive. For stage one, several generation scenarios are necessary to develop an expected operational cost. In the second stage, however, each generation scenario and plan combination reflects a new state.

Finally, in the fourth step of the exploration algorithm, the Bellman values are calculated for each state. As discussed in 4.2.1, the Bellman values for the final time stage are the investment and deterministic generation costs for each plan and scenario combination. The Bellman values for the first-stage include the expected current and future stage contributions and are calculated recursively.

4.2.4 Exploitation Phase

The second half of the algorithm exploits the knowledge gained during the exploration phase to optimize decisions. The exploration phase generally follows the double-pass algorithm described in Table 4-1. The forward pass of the algorithm samples a path through the tree, and the backward pass evaluates those states sampled in the forward pass and updates the problem information. Like the exploration phase of the algorithm, the exploitation phase also uses MILP techniques to identify valid solutions. In the exploration phase, however, the MILP techniques are combined with interpolation to select plans with high expected values.

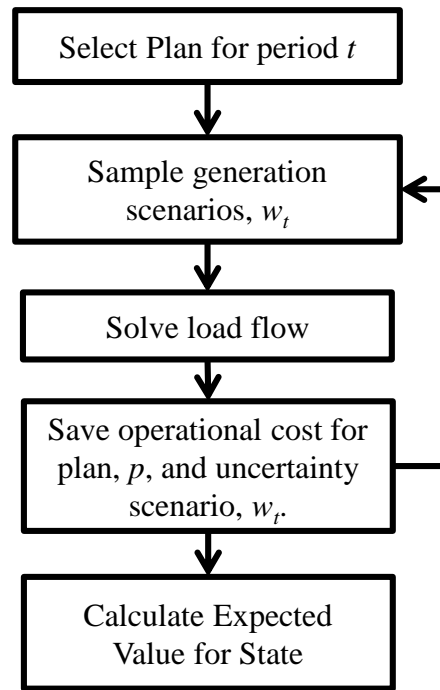


Figure 4-9 Process for the Generation of Additional Valid Plans

The first step of the forward pass, shown in Figure 4-10, creates a set of new transmission investment candidates for consideration. First, the lowest-cost (Top N, where N is a parameter set for individual problems) plans are tested for feasibility given the current state. During the first planning stage, all future transmission plans are possible because no new transmission has been built. During the second planning stage, however, some future plans will not contain all of the lines from the first-stage plan. Because transmission lines are not removed from service, these plans are not feasible for the second stage decision. For example if a 500kV line appears in the stage one plan, all feasible second stage plans must contain that same 500kV line. As shown in Figure 4-11, the algorithm here identifies two random feasible plans included in the Top N plans. If there are fewer than two feasible plans, random plans are sampled from the Top N plans. The new investments from the two sampled plans are then combined to create a candidate investment set.

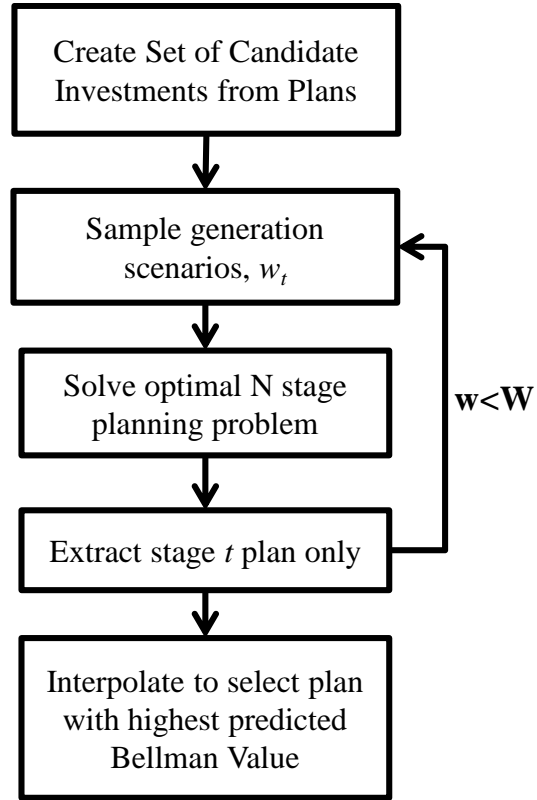


Figure 4-10 Overview of forward pass

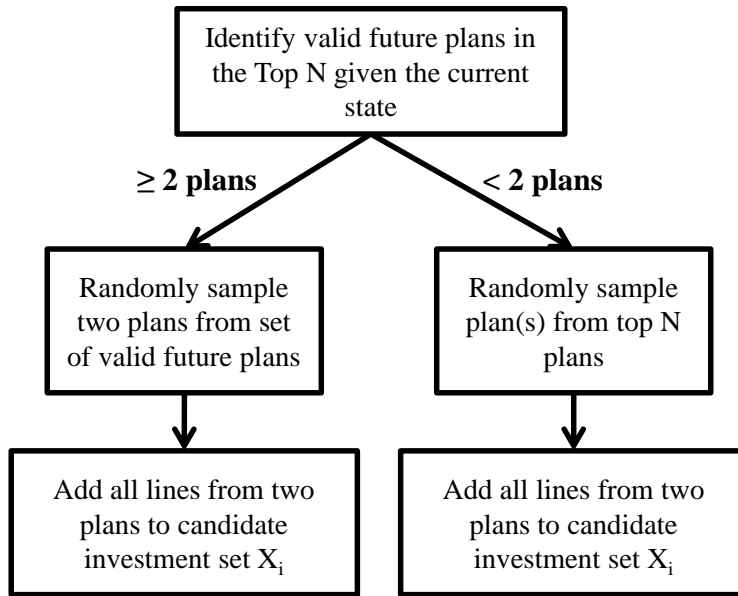


Figure 4-11 Selecting a Set of Candidate Lines

With the set of candidate investments selected, a series of candidate transmission plans are created. These plans are created by iteratively optimizing the same set of candidate investments for different deterministic generation scenarios. In the second-stage, only a single plan needs to be created as the plan is for a deterministic state. If the problem is sufficiently large such that the two-stage optimization requires an unreasonable time, simplifications such as using a transportation power flow model in the second stage, could be explored. These simplifications are allowable because only the selections from the first-stage are retained from the optimization, and the second-stage is included in the optimization to prevent the algorithm from greedily choosing lines for the current stage and ignoring future investment and operational costs.

Up to this point, the ADP algorithm proposed has not followed the structure proposed of the standard double-pass algorithm. In the standard approach, the exploit-phase of the algorithm would simply optimize the approximate value function to identify a single new sample rather than identify several candidates for exploration as done here. The reason for this is straightforward. As discussed in, Chapter 3 identifying the small subset of valid plans is a non-trivial problem, even when starting from an existing valid

plan. Thus, both the exploratory and early steps of the exploit algorithm are intended to identify the valid regions of the search space, a step which ordinarily does not need to be undertaken. The candidate plans identified in this section are not guaranteed to be valid; however, because the set of lines are derived from two valid plans, the probability of non-served energy is low.

From this point in the algorithm, the more standard double-pass algorithm is followed. With several candidate plans identified, interpolation is used to select the plan with the highest expected value. At this stage, there is a trade-off between the number of candidate plans identified during each iteration and run-time. With more candidate plans, interpolation can be more effectively used, requiring fewer total states to be explored. The identification of plans, however, requires solving a MILP which is computationally expensive compared to running the optimal power flows used to assess plans once selected. The balance will need to be tuned for individual problems. Again, in the final deterministic stage, only a single optimization is completed as there is no uncertainty and interpolation is unnecessary.

The backward pass evaluates the plans selected in the forward pass and updates the expected stage cost contributions and Bellman values. As shown in Figure 4-12, the backward pass has five steps. If the state is new, e.g. the state has not yet been selected for evaluation, the interpolation database is updated to include the new state. Next, an estimate of operational costs is calculated. This estimate includes running new optimal power flows and updating the previous estimate of operational costs. While there are several ways to update the estimate, the values here are averaged as shown in Eq. 36. With updated operational cost estimates, the Bellman values can be updated as shown in Eq. 37. The second stage costs are only updated if the estimated cost for the second-stage is lower than the current best. Finally, with new Bellman values, the list of Top N plans can be updated.

$$\text{Eq. 36} \quad \widehat{oc}^i(S_t^i) = \widehat{oc}^{i-1}(S_t^i) \frac{n-1}{n} + OC^i(S_t^i) \frac{1}{n}$$

Eq. 37 $\hat{v}^i(S_1^i) = \widehat{oc}^i(S_1^i) + IC(S_1^i) + \alpha \cdot \min E[IC(S_2|S_1^i) + \widehat{oc}^i(S_2|S_1^i)]$

i	iteration
n	number of times state (S_t^i) has been selected
OC^i	operational costs found during iteration i
$\hat{v}^i(S_t^i)$	Estimate of the Bellman cost for the stage one stage selected in iteration given iteration i
$\widehat{oc}^i(S_t^i)$	Estimate of the operational costs for the stage one stage selected in iteration given iteration i

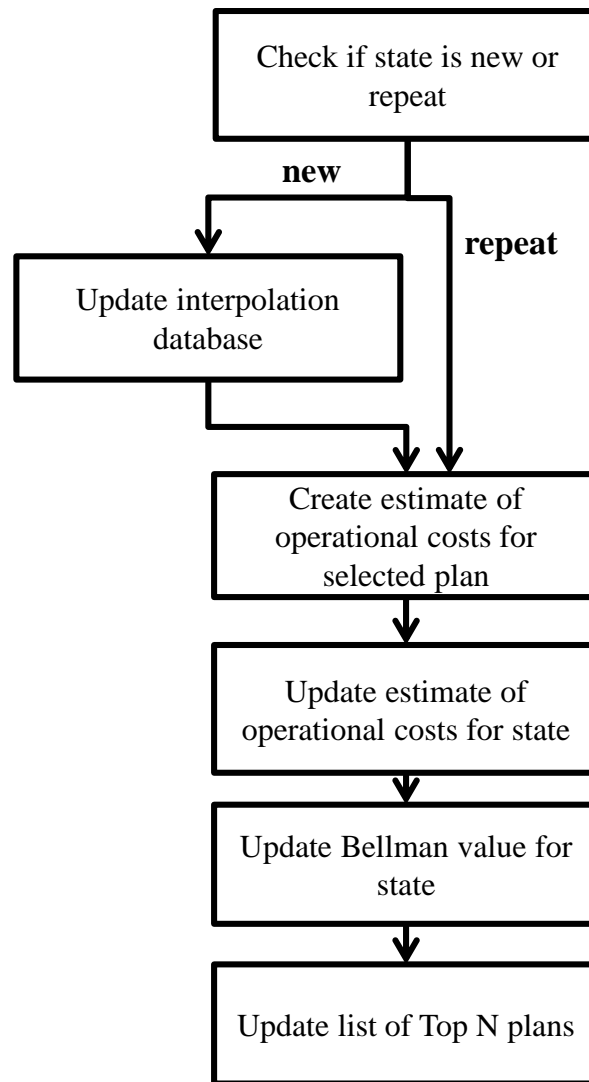


Figure 4-12 Backward Pass

At first blush, it may seem odd that a new estimate of operational costs is calculated during the backward pass. After all, the optimization used in the forward pass to identify candidate plans is optimizing the sum of operational and investment costs. Using these costs, however, would never provide a good estimate for the operational costs and would result in a systematic bias toward plans optimized for high renewable energy. The creation of this bias is demonstrated in Figure 4-13. Assume that for some problem, there are only two generation scenarios, one with high renewable energy penetrations and one with

low renewable energy penetrations. The optimization in the forward pass will produce two plans, one roughly optimized for the high scenario and one for the low scenario. Assuming both plan are valid, the interpolation should select *Plan A* as it will have a lower total cost. These lower costs results from the fact that operational costs dominate investment costs and the new renewable energy generators have zero operational cost. While *Plan B* may have a better expected value, it is never evaluated for the high penetration scenario and *Plan A* is never evaluated for the low penetration scenario. By sampling generation scenarios to update the operational costs, the plans are explicitly evaluated under different generation scenarios and the bias in operational costs is mitigated.

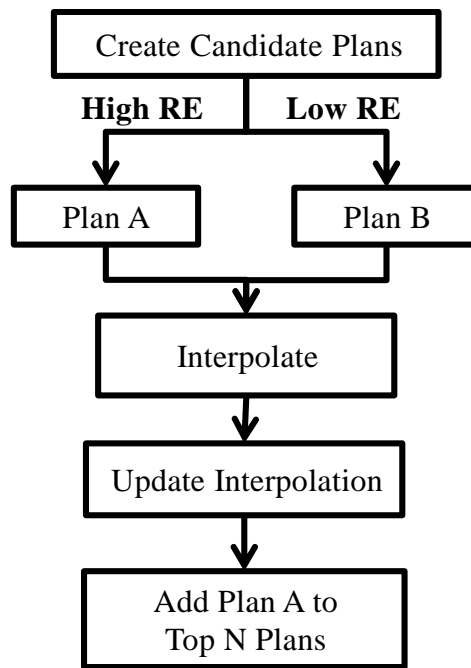


Figure 4-13 Process for Creation of Bias in ADP Algorithm

4.3 Evaluation of the STEP Model

The STEP model’s performance was assessed on a demonstration-scale problem. This problem, described in 4.3.1, was also solved using a Branch and Bound algorithm to benchmark the STEPP model. These results are presented in sections 4.3.2 through 4.3.4.

Across sections, the first-stage plans are used for cost comparison and analysis. These first-stage plans are used because they are the actionable decisions produced by the model. After stage-one decisions are made, the state is reassessed and the model is rerun to identify new stage-one decisions. Section 4.3.2 compares the cost obtained by the branch and bound solution to those obtained by the STEP model. Section 4.3.3 examines the composition of the lowest cost plans and convergence of the STEP model is discussed in 4.3.4. For the test problem described in 4.3.1, the STEP model was executed according to the parameters in Table 4-2.

Table 4-2 STEP Model Parameters

Exploratory Phase Replacement Coefficient	20
Number of Top Plans Tracked	30
Exploration Iterations	206 (time limited to 1,000 seconds)
Exploitation Iterations	5,000
Candidate Plans for Interpolation	2
Optimality Gap for MILP optimization	10%

4.3.1 Test System

The STEP model was evaluated on a small-scale demonstration problem. This test problem follows the structure outlined in Figure 4-14 with two decision stages and one uncertainty stage. The base test system is the same 240 bus WECC model used in both Chapter 2 and Chapter 3. As in Chapter 3, only reinforcement investments in existing corridors were considered. For this test system, however, both 345kV and 500kV investments were included for a total of 440 potential investments. The two uncertainty scenarios modeled reflect 25% and 50% investment in top 10 WREZs, and a single load hour is modeled for each decision stage. Again, this test system is intentionally small such that the STEP results can be benchmarked using existing methods.

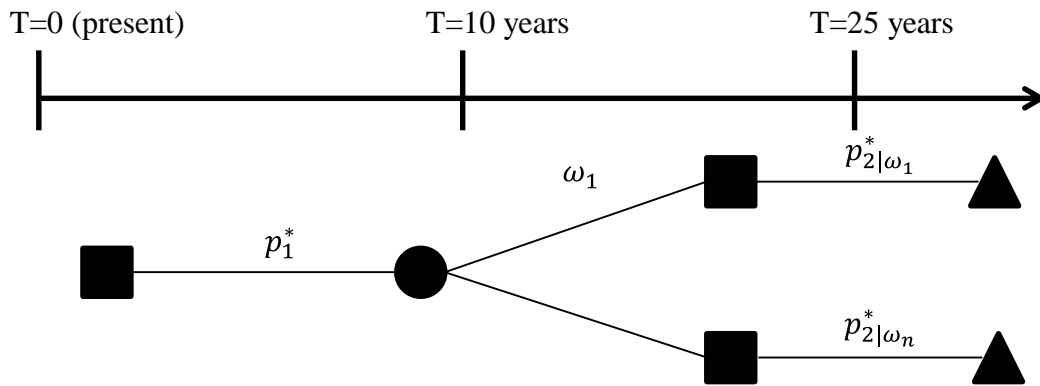


Figure 4-14 Optimal Solution Structure for STEP Test Problem

To provide the benchmark solution, the test problem was formulated as a deterministic equivalent and solved using the CPLEX Branch and Bound algorithm. The solution identified obtained using Branch and Bound algorithm had an optimality gap of 1.05%. In addition to solving the stochastic problem, the two-stage scenarios were also solved. For the test problem, the stochastic solution differs from the two scenario results. Each of the plans include a common set of 16 transmission lines; however, as shown in Table 4-3, both scenario plans both include lines not in the stochastic solution and excluded lines present in the stochastic solution.

Table 4-3 Comparison of Stochastic and Scenario Stage-One Results

	Scenario 1 (50% generation)	Scenario 2 (25% generation)	Stochastic
Number of Lines	31	24	25
Lines in Scenario but not Stochastic	11	6	
Lines in Stochastic but not Scenario	5	7	

4.3.2 Cost of Plans

The STEP model identifies a first-stage solution with a total expected cost 0.03% higher than the optima branch and bound solution. In the MS-TNEP problem, this difference is negligible due to the simplifications made in the model. As shown in Figure 4-15, the cost of the best plan in the STP model improves from 1.7% of the theoretical lower bound cost provided by the branch and bound optimization to 1.1%. Likewise, the mean cost of the Top 30 plans reduces from 2.3% to 1.2%.

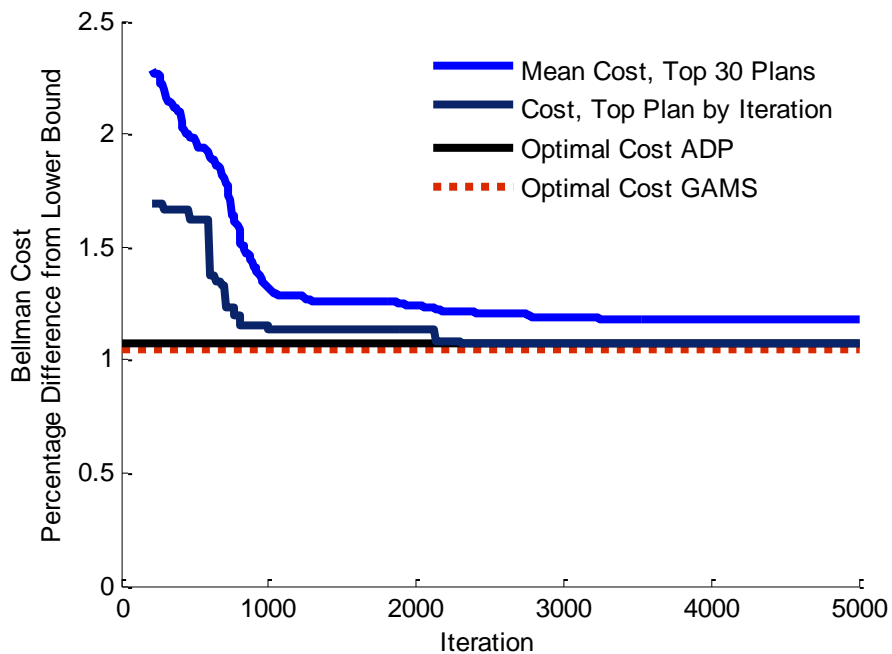


Figure 4-15 Cost Comparison of STEM and Branch and Bound Solution

While identifying the optimal plan for a stage-one expansion plan is a useful benchmark, there are many plans with similar costs. As shown in Figure 4-16, 19% (215) of the 1,136 plans are within 2% of the theoretic lower bound. Realistically, the mid-term and long-term planning models do not have sufficient operational details to differentiate between these plans. As a result, it is more useful to examine the key features, common lines or structures of the best performing plans than consider only the optimal solution.

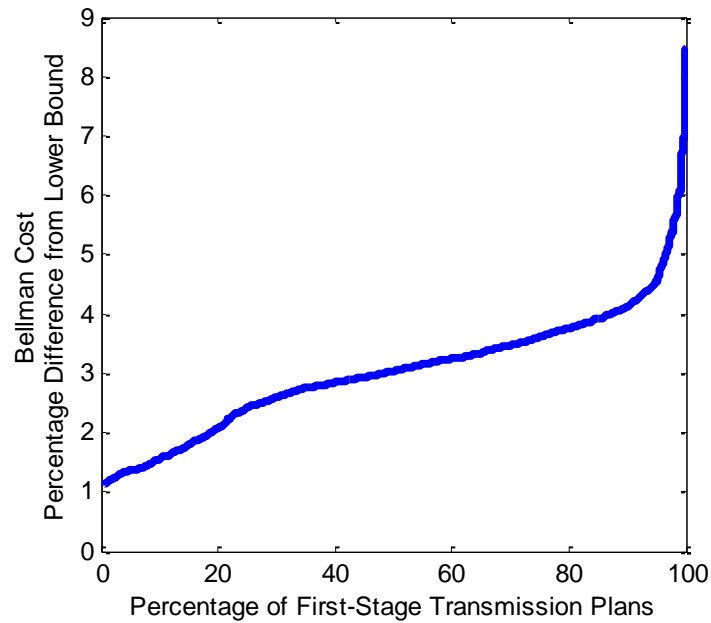


Figure 4-16 Distribution of Expected First-Stage Costs

4.3.3 Composition of First-Stage Plans

As discussed both in Chapter 3 and above in section 4.3.2, transmission expansion plans developed by models such as STEP will require further analysis due to their limited operational detail and siting concerns. Because operational-reliability constraints will likely require modifications to the plan produced by such models, the patterns of investment across the lowest-cost plans are more important than the specific expansion plan.

Of the possible 440 investments, only 8% (36) are used in the first-stage Top 30 plans. The reduction in the number of transmission lines occurs both in the exploratory and exploitative phases of the STEP model. During the exploratory phase of the model, 206 two-stage plans were identified containing 197 unique investments. Because these 197 investments are the only investments available for future first-stage transmission investments, the exploratory phase decreases the number of potential first-stage investments by 55%. More realistically, the investments in the Top 30 lines after the exploratory phase will be those selected for further analysis. This reduces the number of

possible investments 81% from 440 to 84. The exploitative phase of the algorithm then reduces the number of transmission investments again from 84 to 36. This process is shown iteratively in Figure 4-17.

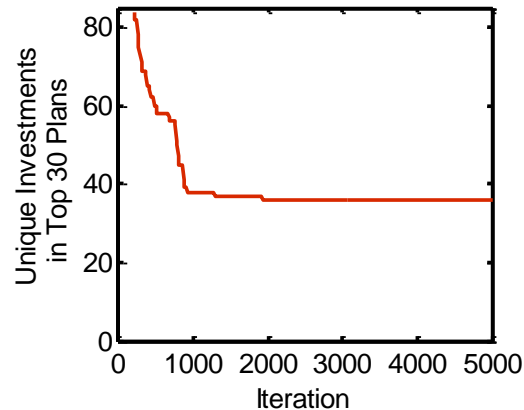


Figure 4-17 Reduction in the Number of Potential Transmission Investments

With 36 possible investments, there are still 2^{36} possible transmission expansion plans. As shown in Figure 4-18 and Figure 4-19, the 36 transmission investments are not used in equal frequency across the plans. In this smaller set of potential investments, 44% (16) are used across all Top 30 first-stage plans. These investments, shown in Figure 4-23, are regional reinforcements rather than an east-west or north-south connection across large geographic areas. These east-west corridors, however, emerge when the lines appearing in at least 50% or 75% are added to the transmission maps as shown in Figure 4-21 and Figure 4-22. Only 4 of the 36 investments appear in fewer than 50% of the Top 30 Plans.

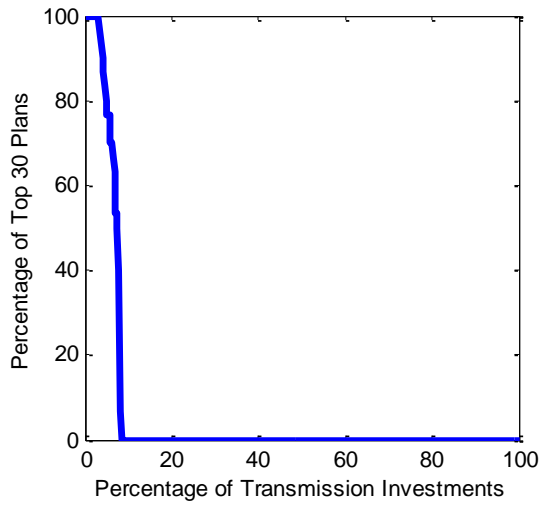


Figure 4-18 Distribution of All Transmission Investments

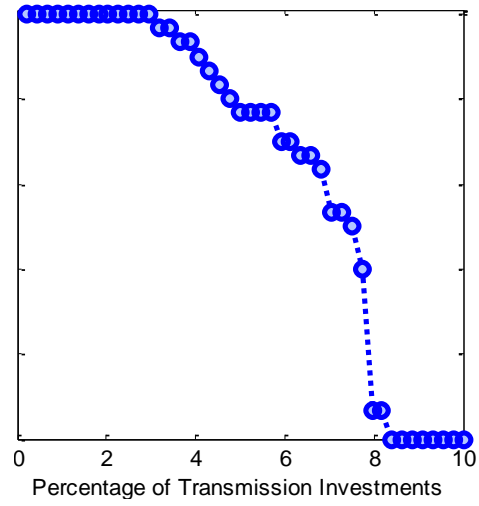


Figure 4-19 Distribution of the Most Frequent Transmission Investments

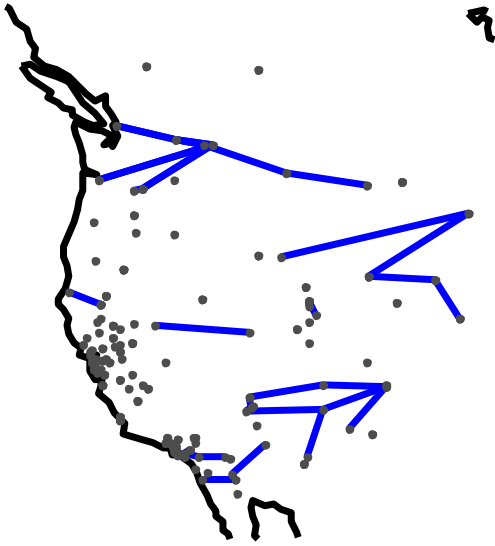


Figure 4-20 All Lines in the First-Stage Top 30 Plans

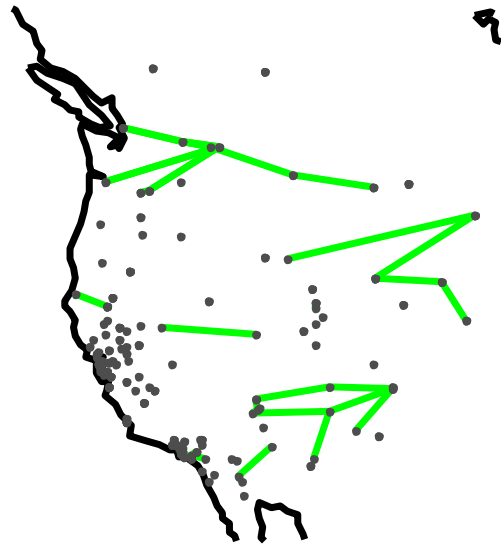


Figure 4-21 Lines in at least 50% of Top 30 First-Stage Plans

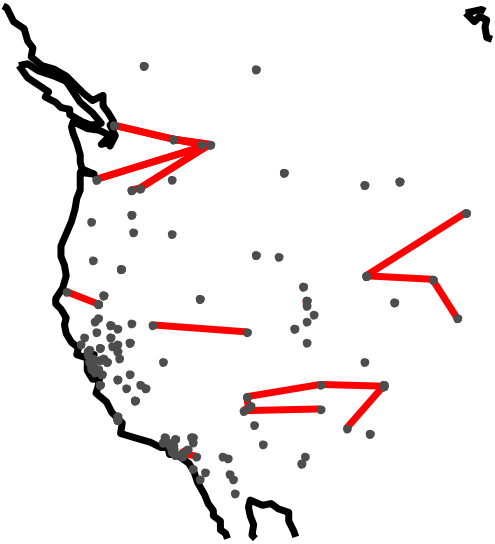


Figure 4-22 Lines in at least 75% of Top 30 First-Stage Plans



Figure 4-23 Lines in All Top 30 First-Stage Plans

4.3.4 Convergence

Convergence in the STEP model can be examined both in the first-stage decisions and the cost of these first-stage decisions. As shown in Figure 4-15 in 4.3.2, the mean cost for the Top 30 Plans converged after 3,420 iterations. The first 1,500 samples accounted for most of the change in the cost, approximately 1% drop, while the remaining 3,500 samples accounted for only a 0.1% additional drop in cost. In the action space, the set of Top 30 plans for the first-stage does not conclusively converge. The last change in the set occurs after 4,460 of 5,000 iterations as shown in Figure 4-24. While the set of Top 30 plans continue to change, these changes have minimal effect on the costs, and as discussed in 4.3.3, there are only 36 lines to select between for the final 3,000 iterations. As a result, the structure of the new plans will not change significantly.

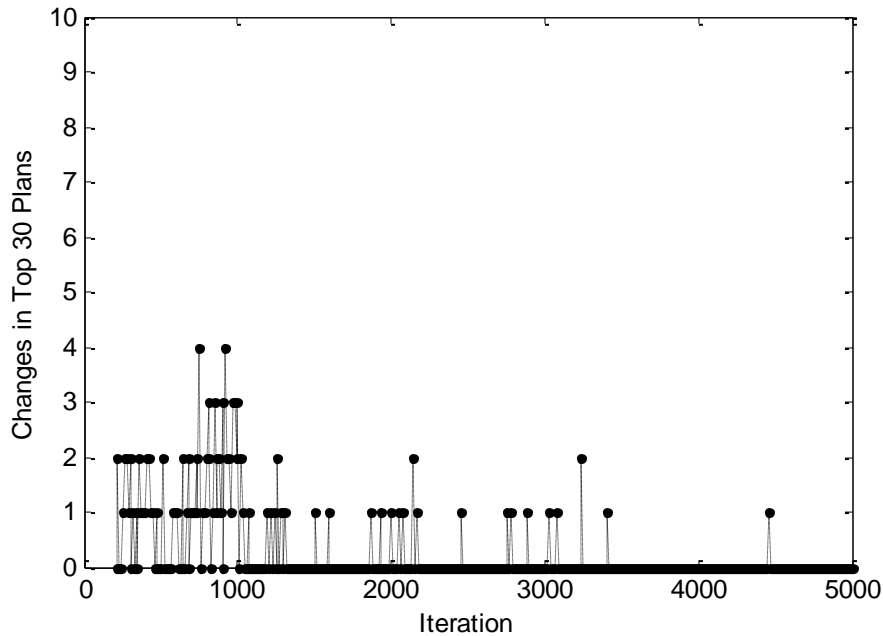


Figure 4-24 Action Convergence
Changes Tracked Across Groups of 10 Iterations

4.4 Conclusions and Future Work

This chapter develops the STEP (Stochastic Transmission Expansion Planning) model. The multi-stage stochastic transmission expansion planning problem naturally fits into

the Markov state-based Approximate Dynamic Programming framework. The STEP model adapts a double-pass ADP framework to include the interpolation methods developed in Chapter 3. In the model, mixed integer linear programming is used to identify new plans, and the interpolation techniques from Chapter 3 to choose amongst these potential plans.

The STEP model was tested on a demonstration-scale problem. The demonstration problem included 440 potential transmission investments, two decision stages and two uncertainty scenarios. In this problem, the STEP model identified a first-stage transmission expansion plan with a cost 0.03% higher than the optimal plan identified using a branch and bound solver. The STEP model also identified 215 first-stage expansion plans within 2% of the theoretic lower bound. These low-cost plans all include a set of 16 transmission lines and the 30 lowest-cost plans use a total of 36 lines. These results mirror the results from the St. Clair Screening Model in Chapter 2 where a small number of lines are used across of scenarios and most potential investments are not used.

There are many future research directions for the STEP model. The major research thrust should be directed on increasing the size of the test system. The model here was tested on a demonstration-scale problem with a single load hour and two uncertainty scenarios. To understand the model's behavior, experiments should be conducted with increasing numbers of both load hours and uncertainty scenarios. Furthermore, a single approximation technique was tested; however, Chapter 3 includes several additional options for exploration. Finally, the effects of the model parameters should be explored. These trials include exploring the trade-offs of tracking fewer or greater plans in the set of top solutions and varying the lengths of the exploration and exploitation phases.

THIS PAGE INTENTIONALLY BLANK

5. Conclusions

The development of location-constrained generation resources, integration of electricity markets and increasing connectivity between regional electric systems are forcing transmission planners to consider larger areas than ever before. Rather than areas the size of European countries or states in the United States, transmission planners are planning for areas the size of continental Europe or west of the Rocky Mountains in the United States. Complicating the planning is the uncertainty in the development of new location-constrained generation, especially wind and solar generation. Transmission lines are large capital investments, characterized by their economies of scale, which are made in context of an ever evolving electric system. Planning these lines requires trading off these economies of scale and the uncertainties in the ever evolving power system. The tools to support this multi-stage stochastic wide-area transmission expansion planning, however, do not yet exist.

This dissertation presents the framing and building blocks to develop these tools by blending both heuristic and mathematical programming approaches. The wide-area multi-stage stochastic transmission network expansion planning problem (MS-TNEP) is challenging from an algorithmic perspective due to lumpy investments, physically driven network flows and dimensionality. For n transmission investments, there are 2^n unique transmission expansion plans possible. Reducing this dimensionality is key to developing tools for realistically sized MS-TNEP problems. To manage the dimensionality of the problem, this work develops a screening model, interpolation methods and an approximate dynamic programming model.

5.1 Dissertation Summary

The first building block is the St. Clair Screening Model, a tool to reduce the number of transmission investments for consideration. In small systems, expert judgment can be used to select a small number of investments for consideration. In wide-area planning, however, there are too many possible network configurations and investments to manually screen the investments. The St. Clair Screening Model, developed in Chapter 2,

combines linear optimization and generation sampling to identify corridors, paths between two nodes, and lines, specific investment types in each corridor, for future investment. In a test system, the St. Clair Screening Model was able to reduce the number of corridors of interest by 95%, from 13,695 to 629. The model also reduced the number of investments by 97%, from 41,085 to 1,081.

In addition to reducing the number of corridors and lines for consideration, the St. Clair Screening Model provides information about the frequency of investment in individual corridors. In the test study, 41 corridors were identified with development in at least 90% of scenarios. These 41 corridors connected both existing system nodes as well as nodes with new location-constrained generation potential. The corridors which interconnect nodes with new generation potential, representative of geographic regions with wind and solar resources, are advantageous to develop as they lie along economically advantageous pathways. These corridors may be considered robust and emphasized in future planning studies.

The second building block is interpolation methods to guide heuristic searches. Even with the reduction of investments made through the St. Clair Screening Model, the MS-TNEP problem is too large to solve via traditional methods. Instead, Chapter 3 combines dimensionality reduction and interpolation techniques to predict the cost and ranking of new transmission expansion plans. The dimensionality reduction techniques, Method of Moments and Principal Component Analysis, are drawn from the image recognition algorithm. These choices were inspired by the ease of pattern recognition in transmission maps and their ability to reduce the dimensionality of the problem from a function of the number of lines, 2^n , to 10. To interpolate within the reduced dimensionality spaces, the dimensionality reduction techniques were paired with both regression and nearest neighbor techniques. When combined, these methods were able to select the lower cost transmission plans in over 80% of trials and at best, predict costs within 14%.

Finally, the third building block is a heuristic algorithm to embed the methods developed in Chapter 4. The framework developed here is based on Approximate Dynamic Programming (ADP). ADP manages dimensionality by combining Monte Carlo methods

with value function approximation. The STEP model developed in Chapter 4 uses mixed-integer linear programming techniques to identify transmission plans with low non-served energy and then selects high value plans through the interpolation techniques from Chapter 3. In a limited test problem, this model was able to identify a transmission expansion plan with costs within 0.15% of the optimal plan identified using a deterministic equivalent mixed-integer linear program.

5.2 Policy Implications

This dissertation demonstrates the feasibility of algorithmic wide-area transmission planning. Although this particular piece of work focuses on developing tools for planning, further works can and will add the necessary operational detail to produce transmission expansion plans for full power systems. The modeling discussed in this work focuses on reframing the transmission expansion planning question from one of local reinforcement to one of broad patterns of investment. Both the reframing of wide-area planning and the demonstration of tools for wide-area planning carry implications for policymakers focused both on regional and wide-area network design.

The first major policy implication is that transmission expansion planning on a realistic scale is within reach. Until this point, decision makers have relied on plans put together using expert judgment from stakeholders; however, this work demonstrates that neutral decision support models can be developed to guide the policymaking. Unlike existing models which are designed for generation expansion and have very coarse detail of the grid, the St. Clair and STEP models are specifically designed to capture the transmission system.

This work also demonstrates the feasibility of looking beyond scenario analysis to planning a network robust to future generation development. Most planning authorities produce a variety of scenario-specific transmission plans; these scenarios often result in very different network topologies which are narrowly tailored for the scenario. Instead, this work demonstrates that networks can be planned to provide value across future

generation scenarios and that oversight authorities should request the development and use of models which move beyond single scenarios.

Finally, for both regional and wide-area policymakers, the results from the St. Clair filter provide a set of corridors which are consistently developed across scenarios. These robust corridors provide a means of proactive planning for decision makers and can be made a regional or national priority. For each corridor identified, environmental, litigious and siting decisions can be identified and resolved early.

5.3 Future Work

In addition to the detailed future research recommendations given in each chapter, there are three broad areas of future work for all tools developed in this dissertation: more realistic systems, exploring a greater number of uncertainties, and finally increasing the speed of each algorithm's executions.

The building blocks developed in this work have been created to solve real-world problems for real-world systems. As first-of-a-kind tools, they were designed and tested on demonstration scale problems with limited operational details. Future research should focus on testing and refining these tools on larger systems. These test systems should include greater technical detail, for example losses, and also a greater number of load hours.

The research in this dissertation was also specifically motivated by the growth of location-constrained generation and its impact on the transmission network. There are, however, a great number of uncertainties which affect the value of transmission investments. These range from the retirement of current generators and changes in fuel prices to changing load patterns. The effects of these other sources of uncertainty should be included in research going forward.

Finally, the computation time for the St. Clair Screening Model and STEP model could be reduced. The prototype codes developed for the research here were not well automated and their efficiency could be dramatically improved upon. The speed could be

increased both through parallelization and also through pre-computation of constants, such as the moment values for each transmission investment.

THIS PAGE INTENTIONALLY BLANK

6. Works Cited

- [1] Alguacil, N. and A.J. Conejo. *Multi-period optimal power flow using Benders decomposition*. IEEE Transactions on Power Systems, Vol.15, No.1, pp.196,201, Feb 2000.
- [2] American Electric Power. *Encroachments on Transmission Rights of Way*. No Date. Available: <https://www.aepohio.com/global/utilities/lib/docs/info/facts/AEPOhio-Encroachment-onTransROW.pdf>
- [3] American Electric Power. *Transmission Facts*. No Date. Available: <http://www.aep.com/about/transmission/docs/transmission-facts.pdf>
- [4] American Electric Power. *Interstate Transmission Vision for Wind Integration*. Electricity Today. September 2007.
- [5] Asada, E.N., Carreno, E., Romero, R., and A.V. Garcia. *A Branch-and-Bound Algorithm for the Multi-Stage Transmission Expansion Planning*.
- [6] Bahiense, L., Oliveira, G.C., Pereira, M., and S. Granville. *A Mixed Integer Disjunctive Model for Transmission Network Expansion*. IEEE Transactions on Power Systems, Vol. 16, No.3. August 2001.
- [7] Bertsekas, D.P and J. Tsitsiklis. *Neuro-Dynamic Programming (Optimization and Neural Computation Series, 3)*. Athena Scientific. 1996. ISBN: 978-1886529106
- [8] Binato, S., Pereira, M.V.F., and S. Granville. *A New Benders Decomposition Approach to Solve Power Transmission Network Design Problems*. IEEE Transactions on Power Systems. Vol. 15, No.2, May 2001.
- [9] Braga, A., Silvestre D. and J.T. Saraiva. *A Multiyear Dynamic Approach for Transmission Expansion Planning and Long-Term Marginal Cost Computation*. IEEE Transactions on Power Systems. Vol.20, No.3, pp.1631-1639. 2005.
- [10] Boyle, G., Guthrie, G. and R. Meade. *Real Options and Transmission Investment: the New Zealand Grid Investment Test*. New Zealand Institute for the Study of Competition and Regulation (ISCR). 2006. <http://www.iscr.org.nz/n172.html>

- [11] Bustamante-Cedeno, E. and S. Arora. *Stochastic and Minimum Regret Formulations for Transmission Network Expansion Planning Under Uncertainties*. Journal of the Operational Research Society. Vol. 59. 2008.
- [12] Buygi, M.O., Balzer, G., Shanechi, H.M., and M. Shahidehpour. *Market-Based Transmission Expansion Planning*. IEEE Transactions on Power Systems. Vol. 19, No. 4. November 2004.
- [13] Cash, G.L., and M. Hatmain. *Optical Character Recognition by the Method of Moments*. Computer Vision, Graphics and Image Processing. Vol. 39, pp. 291-310. 1987.
- [14] Choi, J., Tran, T, El-Keib, A.A., Thomas, R., Oh, HS, and R. Bilinton. *A Method for Transmission System Expansion Planning Considering Probabilistic Reliability Criteria*. IEEE Transactions on Power Systems. Vol., 20, No.3. August 2005.
- [15] Cotilla-Sanchez, E., Hines, P.D.H., Barrows, C. and S. Blumsack, (2012). *Comparing the Topological and Electrical Structure of the North American Electric Power Infrastructure*. IEEE Systems Journal. Vol. 6, No.4, pp. 616-626. December 2012.
- [16] de Dios, R., Saiz, A., Melsion, J.L and A. Bassey. *Strategic Transmission Network Planning*. Report. Red Electrica de Espana.
- [17] dena. dena *Grid Study II: Integration of Renewable Energy Sources in the German Power Supply System from 2015-2020 with an Outlook to 2025. Summary of the main results by the project steering group*. Available: http://www.uwig.org/summary_dena_grid_study_ii.pdf
- [18] Dusonchet, Y.P. and A. El-Abiad. *Transmission Planning Using Discrete Dynamic Optimizing*. IEEE Transactions on Power Apparatus and Systems. Vol. PAS-92, No.4, pp. 1358-1371. July 1973.
- [19] Eastern Interconnection Planning Collaborative. Last accessed: November 27, 2013. <http://www.eipconline.com/>
- [20] Electric Reliability Council of Texas. Last accessed: November 27, 2013. <http://www.ercot.com/>

- [21] Energy Information Administration. *Frequently Asked Questions: How Much Electricity is Lost in Transmission and Distribution in the United States?*
<http://www.eia.gov/tools/faqs/faq.cfm?id=105&t=3>
- [22] Escobar, A.H., Romero, R.A., and R.A. Gallego. *Transmission Network Expansion Planning Considering Multiple Generation Scenarios*. IEEE. 2008.
- [23] European Environment Agency. *Europe's Onshore and Offshore Wind Energy Potential: An assessment of environmental and economic constraints*. ISSN: 1725-2237. June 2009.
- [24] European Union. *Photovoltaic Solar Electricity Potential in European Countries*. 2012. Available: <http://re.jrc.ec.europa.eu/pvgis>
- [25] Federal Electric Regulatory Commission. Order No. 1,000: Transmission Planning and Cost Allocation. October, 2012.
<http://www.ferc.gov/industries/electric/indus-act/trans-plan.asp>
- [26] Galbraith, K. *Lack of Transmission Lines is Restricting Wind Power*. New York Times. January 20, 2011.
- [27] Garver, L.L., *Transmission Network Estimation Using Linear Programming*. IEEE Transactions on Power Apparatus and Systems. Vol.PAS-89, No.7, pp.1688-1697, Sept. 1970
- [28] Haffner, S., Monticelli, A., Garcia, A. and R. Romero. *Specialised Branch-and-Bound Algorithm for Transmission Network Expansion Planning*. IEE Proceedings – Generation, Transmission and Distribution. Vol. 148, No.5, pp482-488. September 2001.
- [29] Hedman, K.W., Gao, F. and G. B. Sheble. *Overview of Transmission Expansion Planning Using Real Options Analysis*. Proceedings of the 37th Annual North American Power Symposium, 2005.
- [30] Hu, M-K., *Visual Pattern Recognition by Moment Invariants*, IRE Transactions on Information Theory. Vol.8, Issue 2, pp. 179-182., Feb. 1962.

- [31] International Energy Agency, Task 25. *Design and Operation of Power Systems with Large Amounts of Wind Power: Final Report, IEA WIND Task 25 Phase 1: 2006-2008*. ISBN 978-951-38-7309-7
<http://www.ieawind.org/AnnexXXV/PDF/Final%20Report%20Task%2025%202008/T2493.pdf>
- [32] Latorre-Bayona, G.; Perez-Arriaga, I.J., *CHOPIN, a heuristic model for long term transmission expansion planning*, IEEE Transactions on Power Systems. Vol.9, No.4, pp.1886,1894, Nov 1994
- [33] Leite de Silva, A.M., Rezende, L.S., Honorio, L.M., and L.A.F Manso. *Performance Comparison of Metaheuristic to Solve the Multi-Stage Transmission Expansion Planning Problem*. IET Generation, Transmission and Distribution. Vol. 5, Issue 3. 2011.
- [34] Li, W. and P. Choudhury. *Probabilistic Transmission Planning*. IEEE Power and Energy Magazine. September/October 2007. DOI: 10.1109/MPE.2007.9004765
- [35] Maghouli, P., Hosseini, S.H., Buygi, M.O. and M. Shahidehpour. *A Scenario-Based Multi-Objective Model for Multi-Stage Transmission Expansion Planning*. IEEE Transactions on Power Systems. Vol. 26, No.1. February 2011.
- [36] Mariner Consulting Services. *Why an Energy Imbalance Market will make the Western Interconnection more reliable*. July 28, 2011.
- [37] Masters, J. *Untitled*. Accessed November 27, 2013.
http://www.wunderground.com/wximage/viewsingleimage.html?mode=singleimage&orig_handle=JeffMasters&orig_number=88&handle=JeffMasters&number=87&album_id=61
- [38] Mori, H. and H. Kakuta. *A New Meta-heuristic Method for Probabilistic Transmission Network Expansion Planning*. IEEE Transmission and Distribution Conference and Exposition. New Orleans, Louisiana. April 19-22, 2010.
- [39] Moulin, L.S., Poss, M. and C. Sagastizábal. *Transmission Expansion Planning with Re-Design*. Energy Systems. Vol. 1 No. 2 pp. 113-139. 2010.

- [40] Munoz, F.D., Hobbs, B.F. and S. Kasina. *Efficient Proactive Transmission Planning to Accommodate Renewables*. Proceedings of the Power and Energy Society General Meeting, pp.1-7. July 2012.
- [41] Munoz, F.D. Hobbs, B.F., Ho, J.L. and S. Kasina. *An Engineering-Economic Approach to Transmission Planning Under Market and Regulatory Uncertainties: WECC Case Study*. IEEE Transactions on Power Systems. Issue: 199. September 2013.
- [42] National Renewable Energy Laboratory. *Eastern Wind Integration and Transmission Study*. February 2011. NREL/SR-5500-47078
- [43] National Renewable Energy Laboratory. *Western Wind and Solar Integration Study*. May 2010. NREL/SR-AAM-8-77557-01
- [44] National Renewable Energy Laboratory. U.S. 80m Wind Resource Maps. Accessed November 27, 2013. Available: <http://www.nrel.gov/gis/wind.html>
- [45] National Renewable Energy Laboratory. Solar Maps. Accessed December 18, 2013. Available: <http://www.nrel.gov/gis/solar.html>
- [46] Norbiato dos Santos, T., and A.L. Diniz. *A Dynamic Piecewise Linear Model for DC Transmission Losses in Optimal Scheduling Problems*. IEEE Transactions on Power Systems. Vol.26, No.2, pp.508,519, May 2011
- [47] Northeast Utilities. *The Middletown-Norwalk Transmission Project*. Accessed November 27, 2013. <http://www.transmission-nu.com/residential/projects/middletown/default.asp>
- [48] Powell, W.B. *Approximate Dynamic Programming: Solving the Curses of Dimensionality*. Wiley-Interscience. 2007. ISBN: 978-0470171554
- [49] Price, J.E. and J. Goodin, *Reduced Network Modeling of WECC as a Market Design Prototype*. Proceedings of the Power and Energy Society General Meeting, San Diego, CA, 2011
- [50] PSR. *OPNET-Network Transmission Expansion Planning and Analysis Methodology Manual*. Version 2.7. September, 2012.

- [51] Puterman, M.L. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. Wiley-Interscience. 2005. ISBN: 978-0471727828
- [52] Rious, Vincent, Jean-Michel Glachant, and Philippe Dessante. "Transmission network investment as an anticipation problem." European University Institute Working Paper. 2010. Available:
http://cadmus.eui.eu/bitstream/handle/1814/13081/RSCAS_2010_04.pdf
- [53] Roh, J.H., Shahidehpour, M, and L. Wu. *Market-Based Generation and Transmission Planning with Uncertainties*. IEEE Transactions on Power Systems. Vol.24, No.3. August 2009.
- [54] Romero, R., Gallego, R.A., and A. Monticelli. *Transmission System Expansion Planning by Simulated Annealing*. IEEE Transactions on Power Systems. Vol. 11, No. 1. February 1996.
- [55] Romero, R. and A. Monticelli. *A Hierarchical Decomposition Approach for Transmission Network Expansion Planning*. IEEE Transactions on Power Systems. Vol. 9, No.3, pp-373.380. February 1994.
- [56] Romero, R. and A. Monticelli. *A Zero-One Implicit Enumeration Method for Optimizing Investments in Transmission Expansion Planning*. IEEE Transactions on Power Systems. Vol. 9, No. 3, pp-1381-1391. August 1994.
- [57] Romero, R., Monticelli, A., Garcia, A. and S. Haffner. *Test Systems and Mathematical Models for Transmission Network Expansion Planning*. IEE Proceedings – Generation, Transmission and Distribution. Vol. 149, No. 1. January 2002.
- [58] Romero, R., Rider, M.J, and I. de J. Silva. *A Metaheuristic to Solve the Transmission Expansion Planning*. IEEE Transactions on Power Systems, Vol. 22, No.3. November 2007.
- [59] Sánchez-Martín, P., Ramos, A., and J.F. Alonso. *Probabilistic Midterm Transmission Planning in a Liberalized Market*. IEEE Transactions on Power Systems, Vol. 20. No 4. November 2005.

- [60] Short, W., Sullivan, P., Mai, T., Mowers, M., Uriarte, C., and A. Martinez. Regional Energy Deployment System (ReEDS). 2011. NREL/TP-6A20-4653. Golden, Colorado.
- [61] Southwest Power Pool. *ITP20: 2010 Integrated Transmission Plan 20-Year Investment Assessment*. 2010. Available:
http://www.spp.org/publications/ITP20_Report_01-26-11.pdf
- [62] Sozer, S., Park, C.S. and J. Valenzuela. *Considering Uncertainty in Transmission Expansion Planning for Restructured Markets*. Available online:
<http://www.eng.auburn.edu/~jvalenz/NSFLMP/Sozer-Park-Valenzuela.pdf>
- [63] St. Clair, H.P. *Practical Concepts in Capability and Performance of Transmission Lines*. Transactions of the American Institute of Electrical Engineers. 1953
- [64] Sum-Im, T., Taylor, G.A., Irving, M.R., and Y.H. Song. *Differential Evolution Algorithm for Multistage Transmission Expansion Planning*. IET Generation, Transmission and Distribution. Vol. 3, Issue 4. 2009.
- [65] United States Department of Energy. *National Transmission Grid Study*. May 2002. Available: <http://certs.lbl.gov/ntgs/main-screen.pdf>
- [66] Wald, M.L. *Ideas to Bolster Power Grid Run Up Against the System's Many Owners*. New York Times. July 12, 2013.
- [67] Webster, M., Santen, N., and P. Parpas. *An Approximate Dynamic Programming Framework for Modeling Global Climate Policy under Decision-Dependent Uncertainty*. Computational Management Science. Vol.9, pp.339-362. 2012.
- [68] Western Electric Coordinating Council. *2013 Plan Summary*. September 19, 2013. Available:
http://www.wecc.biz/committees/BOD/TEPPC/External/2013Plan_PlanSummary.pdf
- [69] Western Governors' Association. *Western Renewable Energy Zones – Phase 1: Mapping concentrated high quality resources to meet demand in the Western Interconnection's distant markets*. 2009. Available:
<http://www.westgov.org/component/content/article/102-initiatives/219-wrez>
- [70] Western Grid Group. Sub-regional planning map. Accessed November 27, 2013.
<http://66.147.244.210/~westerz0/wp-content/uploads/2010/12/sub-regional-map.jpg>

- [71] Yu, Q., Guo, J., and X. Duan. *Dynamic Multi-Stage Transmission network Expansion Planning*. DRPT 2008. Presented April 2008 in Nanjing, China.
- [72] Zhao, Jh.H Dong, Z.H. Lindsay, P. and K.P. Wong. *Flexible Transmission Expansion Planning with Uncertainties in an Electricity Market*. IEEE Transactions on Power Systems. Vol.24, No. 1. November 2009.

Appendix A. Map Generation for Method of Moments

Before the moments can be calculated, the transmission expansion plan must be converted to an image. For this process, the latitude and longitude of each bus in the system must be known in radians. First, the origin is set to the geographic center of the transmission network; the origin may also be set to another point, for example the geographic centroid of the buses. In the image or graph of the transmission expansion plan, all locations are defined in reference to the origin. The units of x and y in Eq. 29 are unique to the image study; for a given image they be pixels, inches, millimeters, etc. The natural units for transmission expansion planning are miles such that an arbitrary unit of one is equal to one mile. To minimize the distortion of translating distances across the continental scale from the spherical latitude and longitudinal coordinates to Cartesian coordinates, the Haversine Formula is used and shown in Table 6-1. The Haversine Formula calculates distances between points on the surface of a sphere rather than straight line distances which pass through the volume of the sphere.

Table 6-1 Transformation of Latitude and Longitude to Cartesian Coordinates

$a = \sin\left(\frac{\Delta l}{2}\right)^2 + \cos(l_0) \cdot \cos(l_1) \cdot \sin\left(\frac{\Delta l}{2}\right)^2$ $b = 2 \arcsin(\min(1, \sqrt[3]{a})) \cdot R_{earth}$ $coordinate = b \cdot sign(\Delta l)$ $\Delta l = l_0 - l_1$ <p><i>R_{earth}</i> = 3965 miles</p> <p><i>l</i>₀: latitude or longitude coordinate of the origin (radians)</p> <p><i>l</i>₁: latitude or longitude coordinate of the bus (radians)</p>
--

For transmission planning, $f(x,y)$ was reinterpreted to represent the capacity of a transmission line. The weighting function for each new line, $f_k(x,y)$, was calculated to be the total capacity of a line divided by its length as shown in Eq. 38. This weighting was used to reflect the physical properties of a transmission network. As will be expanded

upon below, when integrated with the weighting function, moments reflect the quantity of new transmission capacity in geographic areas.

$$\text{Eq. 38} \quad f_k(x, y) = \frac{f_{k,max}}{length_k}$$

With the weighting above, each transmission line can be conceptualized as a rectangular prism. The length of the prism is given by the x and y coordinates and the height of each prism is given by the weighting function given in Eq. 38. Both of these dimensions are indicative of physical quantities – the length of the line and the capacity of the line. The width of the rectangular prism, however, is not representative of a physical quantity and is set to one in order not to distort the physical representations. In the given reference frame, this implies each transmission line (or right of way) is one mile in width, while rights-of-way typically range from 150 to 200 feet [2]. The width values, shown as w in an example calculation of the zeroth moment for a 1,000 MW 100 mile line in Table 6-2, serve to scale the result. As a result, w may be thought more of as a scaling factor.

Table 6-2 Calculation of Zeroth Moment for Arbitrary Line shown in

$$(1) \quad m_{p,q} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q f(x, y) dx dy$$

$$(2) \quad m_{0,0} = \int_0^w \int_0^{100} x^0 y^0 \cdot 10 dx dy$$

$$(3) \quad m_{0,0} = \int_0^w \int_0^{100} 10 dx dy$$

$$(4) \quad m_{0,0} = \int_0^w 1000 dy$$

$$(5) \quad m_{0,0} = 1000w$$

Appendix B. Depth First Search Transmission Plan Generation

A set of sample transmission expansion plans is required to test the dimensionality reduction and interpolation techniques. Ideally, the set of sample transmission plans should be characterized by a diversity of network expansion plans and each plan should be characterized by an expected cost across multiple generation expansion scenarios. Identifying a diversity of viable transmission expansion plans in a time-effective manner, however, is a non-trivial problem. In this context, a viable plan is one with very little or no non-served energy. A wide-area planning problem with 1,000 candidate lines, for example, has $2^{1,000}$ or $10e30$ potential transmission expansion plans; however, due to the networked and non-linear properties of the transmission system only a small subset of these possible expansion plans are viable.

Any method used to generate the sample transmission expansion plans must satisfy the competing demands of identifying a diversity of transmission expansion plans and only identifying viable plans. Instead of a purely random approach, the approach used here combines a depth-first search (DFS) for non-served energy and a traditional integer formulation of the optimal. The DFS avoids repeatedly sampling sets of transmission lines which are cannot produce viable plans and the use of the optimization algorithm ensures that the plans selected will be both viable and have relatively low expected total costs. By selecting only viable plans with low expected costs, the dimensionality reduction techniques can be used to identify features which differentiate good plans from best plans.

Diversity is introduced into the set of sample transmission expansion plans in two ways. First, the set of candidate lines that the optimization algorithm may consider are repeatedly randomly partitioned. The random partitioning introduces diversity by forcing the optimization to select from different sets of lines. Second, the generation expansion scenario considered by the optimization algorithm is varied. Thus, for the same random set of transmission lines, multiple plans may be constructed by optimizing for different future generation scenarios. The full process is outlined in Figure 6-1 and is discussed step-by-step in the following sections.

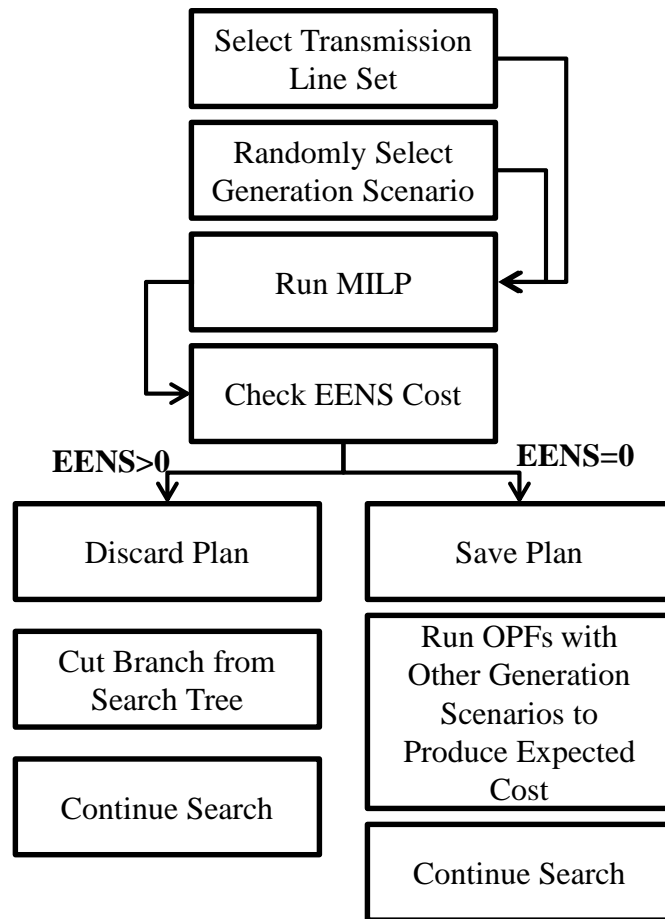


Figure 6-1 Outline of Depth-First Search for Non-Served Energy

The first step of the process is to build a binary search tree using the full set of candidate lines. As shown in Figure 6-2, the tree is constructed by randomly partitioning the set of candidate lines at each branching step. For example, at the first branching point, the full set of candidate transmission lines are randomly divided into two groups. Following the right branch of the tree in Figure 6-2, this set is then again randomly divided into another two sets. This process is repeated a third time. The tree shown here as a depth of three, however, trees may be constructed to other depths depending on the number of candidate transmission lines under consideration. The greater the depth of the tree, the faster the optimization algorithm will be able to solve the lower leaves on the tree and potentially higher diversity will be introduced into the set of sample transmission expansion

sion plans. On the other hand, at higher depths, the number of candidate lines in each set is reduced and the likelihood that no viable plan can be identified with the progressively smaller sets of lines is increased.

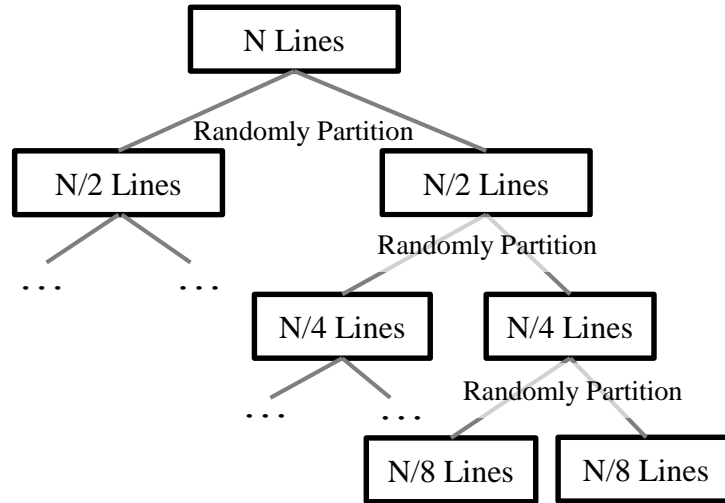


Figure 6-2 Binary Search Tree Composed of Randomly Partitioned Sets of Lines

Once the tree is constructed, a depth-first search is conducted for non-served energy. Within the depth-first search, specific plans are identified and evaluated through a combination of a deterministic and integer transmission-network expansion optimization and an optimal power flow as outlined in Figure 6-1. Each node or leaf in the tree contains a set of candidate transmission lines, not a complete transmission expansion plan. The candidate transmission expansion lines from each set are optimized with a deterministic generation expansion scenario to determine a specific transmission expansion plan. If the plan has no non-served energy, the plan is added to the sample set. If the plan produced by the optimization has non-served energy, the plan is discarded and the tree is pruned to remove the all child nodes of the parent with non-served energy. The pruning prevents the search algorithm from testing subsequent sets of candidate lines which are guaranteed to produce non-served energy as they strictly contain subsets of the same candidate lines. When the tree has been fully explored, a new tree is constructed with freshly randomized sets of candidate lines.

Multiple generation scenarios are used to incorporate stochasticity into the cost of and create diversity in the set of transmission expansion plans. In the second step of Figure 6-1, a random generation scenario is selected to combine with the set of candidate transmission lines. The rationale for this step is two-fold. First, optimizing for different generation scenarios with the same set of lines will produce different transmission expansion plans, inducing diversity in the sample set. Second, solving the MILP is faster for deterministic rather than stochastic problems. The motivation for this work, however, is to solve a stochastic problem and interpolate between the expected value of different plans. In order to incorporate stochasticity, each transmission plan in the set is assessed using an optimal power flow across the full set of generation expansion scenarios. This method produces an expected total cost across many possible future generation scenarios. The cost for each transmission plan accepted is the sum of the expected generation cost and the transmission investment cost.

B.1 Disjunctive Formulation

The transmission network expansion planning problem with a DC load flow is a non-linear mixed or binary integer problem. The non-linearity arises from the flow equation for new lines, shown in Eq. 39, where the voltage angles θ_i and θ_j are continuous variables are multiplied the integer investment variable x . Given the difficulty of solving the mixed-integer non-linear problem, commonly used formulations of the TNEP transform the problem to a binary-integer linear problem using a Big-M disjunctive formulation.

$$\text{Eq. 39} \quad \text{flow} \propto (\theta_i - \theta_j) x$$

The Big-M disjunctive formulation of the transmission linearizes the flow equation using a fictitious variable to effectively turn on or off equations with binary variables. The basic disjunctive formulation is given in equations (Eq. 42 - Eq. 8) and the disjunctive formulation is applied to linearize the flow equation for new lines (Eq. 46 - Eq. 47). When a new line is selected, flow is constrained by the maximum possible flow on that

line and Big-M constants are multiplied by zero and have no impact. When a line is not selected, however, the problem may be artificially constrained. For example, assume the M value of a given candidate is set to zero and the line is not selected by the optimization. The total flow on the line is constrained to zero by (Eq. 5). In this case, the flow equations reduce from Eq. 46 and Eq. 47 to Eq. 40 and Eq. 41. As seen in Eq. 41, a zero M value constrains the voltage angles at the termini of any non-selected line to be equal. In order not to constrain the nodal voltage angles, an M value must be chosen which is larger than the maximum possible difference in voltage angles divided by the reactance.

$$\text{Eq. 40} \quad 0 \leq f_{k,h} + 1/X_k \cdot (\theta_{i(k),h} - \theta_{j(k),h})b \leq 0$$

$$\text{Eq. 41} \quad 0 \leq (\theta_{i(k),h} - \theta_{j(k),h}) \leq 0$$

The basic disjunctive formulation may be tightened by re-formulating the flow in new lines as the sum of its positive and negative components. For these new flow equations, the difference in nodal voltage angles is also divided into its positive and negative. In this formulation, Eq. 46 and Eq. 47 are divided into four new constraints (Eq. 55-Eq. 59). Two additional constraints, Eq. 59 and Eq. 60, are added to constrain the flow and difference angles to be the sum of their positive and negative components. This is the disjunctive formulation used here, coded in GAMS and solved with the base CPLEX Branch and Bound solver.

Basic Disjunctive Formulation

$$\text{Eq. 42} \quad \text{Min } ct \cdot x_k + \sum_{h=1}^H \sum_{i=1}^I (cg \cdot g_{n,h} + c\mu \cdot \mu_{n,h})$$

s.t.

$$\text{Eq. 43} \quad d_{i,h} = \mu_{i,h} + \sum_{k \in \Omega_i} f_{k,h} + \sum_{n \in \sigma_i} g_{n,h} \quad i=1, \dots, I$$

$$\begin{aligned}
\text{Eq. 44} \quad & f_{k,h} - 1/X_k (\theta_{i(k),h} - \theta_{j(k),h})b = 0 && k \in K^0 \\
\text{Eq. 45} \quad & -f_k^{\max} \cdot x_k \leq f_{k,h} \leq f_k^{\max} \cdot x_k && k \in K^+ \\
\text{Eq. 46} \quad & -f_{k,h} + 1/X_k (\theta_{i(k),h} - \theta_{j(k),h})b - M_k(1 - x_x) \leq 0 && k \in K^+ \\
\text{Eq. 47} \quad & f_{k,h} - 1/X_k (\theta_{i(k),h} - \theta_{j(k),h})b - M_k(1 - x_x) \leq 0 && k \in K^+ \\
\text{Eq. 48} \quad & d_{i,h} \geq \mu_{i,h} && i=1, \dots, I \\
\text{Eq. 49} \quad & g_{n,h} \leq g_{n,h}^{\max} && n=1, \dots, N \\
\text{Eq. 50} \quad & f_{k,h} \leq f_k^{\max} && k=1, \dots, K^0
\end{aligned}$$

Complete Disjunctive Formulation

$$\text{Eq. 51} \quad \text{Min } ct \cdot x_k + \sum_{h=1}^H \sum_{i=1}^I (cg \cdot g_{n,h} + c\mu \cdot \mu_{n,h})$$

s.t.

$$\begin{aligned}
\text{Eq. 52} \quad & d_{i,h} = \mu_{i,h} + \sum_{k \in \Omega_i} f_{k,h} + \sum_{n \in \sigma_i} g_{n,h} && i=1, \dots, I \\
\text{Eq. 53} \quad & f_{k,h} - 1/X_k (\theta_{i(k),h} - \theta_{j(k),h})b = 0 && k \in K^0 \\
\text{Eq. 54} \quad & -f_k^{\max} \cdot x_k \leq f_{k,h} \leq f_k^{\max} \cdot x_k && k \in K^+ \\
\text{Eq. 55} \quad & f_{k,h}^- \leq 1/X_k \cdot \theta_{k,h}^- \cdot b - M_k(1 - x_x) && k \in K^+ \\
\text{Eq. 56} \quad & f_{k,h}^+ \leq 1/X_k \cdot \theta_{k,h}^+ \cdot b - M_k(1 - x_x) && k \in K^+ \\
\text{Eq. 57} \quad & f_{k,h}^- \geq 1/X_k \cdot \theta_{k,h}^- \cdot b && k \in K^+ \\
\text{Eq. 58} \quad & f_{k,h}^+ \geq 1/X_k \cdot \theta_{k,h}^+ \cdot b && k \in K^+ \\
\text{Eq. 59} \quad & f_{k,h} = f_{k,h}^+ + f_{k,h}^- && k \in K^+ \\
\text{Eq. 60} \quad & \theta_{i(k),h} - \theta_{j(k),h} = \theta_{k,h}^+ - \theta_{k,h}^- && k \in K^+ \\
\text{Eq. 61} \quad & d_{i,h} \geq \mu_{i,h} && i=1, \dots, I
\end{aligned}$$

$$\text{Eq. 62} \quad g_{n,h} \leq g_{n,h}^{\max} \quad n=1, \dots, N$$

$$\text{Eq. 63} \quad f_{k,h} \leq f_k^{\max} \quad k=1, \dots, K^0$$

Indices and Sets:

i	index of buses
k	index of circuits
$i(k), j(k)$	index of terminal buses of circuit k
h	index of load hours
n	index of generators
K^+	set of candidate circuits
K^0	set of existing circuits
Ω_i	set of all circuits connected to node i
σ_i	set of all generators located at node i
H	number of load hours
I	number of buses
M	number of candidate circuits
N	number of generators

Parameters/Constants

c_g	generator costs [\$/MW]
c_t	annualized cost of candidate circuits [\$]
c_μ	cost of non-served energy [\$/MW]
d	bus demands [MW]
f^{\max}	circuit capacities [MW]
g^{\max}	generator capacities [MW]
X	circuit reactances [pu]
M	Big-M constant [pu]
b	per unit base

Free Variables

$\theta_{i,h}$ bus voltage angle

Positive Variables

$f_{k,h}$ circuit flow [MW]

$f_{k,h}^+$ positive flow component (from i to j) [MW]

$f_{k,h}^-$ negative flow component (from j to i) [MW]

$\theta_{k,h}^+$ positive bus voltage angle component across circuit k

$\theta_{k,h}^-$ negative bus voltage angle component across circuit k

$g_{n,h}$ generator output [MW]

$\mu_{i,h}$ non-served energy [MW]

Binary Variables

x_k investment variable

B.2 Calculation of M Values

The disjunctive formulation of the binary integer linear model requires calculation of appropriate Big-M constants. As demonstrated previously, under-sizing of M values can artificially constrain the planning model; however, over-sizing of is also problematic, potentially ill conditioning the problem and slowing Branch and Bound searches [8]. Optimal M values then are as small as possible without creating artificial constraints. From Eq. 41, M must be greater than the reactance of a candidate line multiplied by any given value of voltage angle difference, θ_i minus θ_j . As the reactance is a constant value for each investment, the minimum value of M is given by the maximum difference in voltage angles across the candidate line's terminuses.

In corridors connecting two nodes in an already well-connected system, the maximum angular difference is constrained by the characteristics of existing lines. For lines in this type of corridor, the provable minimum value of M can be calculated using a shortest

path problem based on the C_k value from Eq. 65 for each existing line. This minimum value can also be tightened as an algorithm proceeds.

$$\text{Eq. 64} \quad \frac{f_k^{max} X_k}{b} = \max(\theta_{i(k),h} - \theta_{j(k),h})$$

$$\text{Eq. 65} \quad C_k = \frac{f_k^{max} X_k}{b}$$

Methods to identify M values for corridors connection new nodes into an existing well-connected system or connecting multiple well-connected systems were developed in the context of reinforcing and expanding small systems. In [8], Binato proved that minimum values in these cases can be calculated by solving a longest path problem. In this context, the cost of each arc is equal to the maximum angular difference possible across the transmission corridor. One common method of calculating a longest path problem is to multiply costs, in this case angles, by negative one and then solve the shortest path problem. Viewed as a graph, however, the transmission network contains many cycles. To avoid infinite loops [39] suggests a heuristic of summing maximum angular differences as between each set of nodes as an upper bound.

This approach implicitly assumes a small system. For stability reasons, differences in voltage angles never exceed 180 degrees in a system and the maximum voltage between two nodes is 90 degrees. Voltage angles, even across long lines, typically do not exceed 40 degrees. In a wide-area system, or a system of realistic size, this heuristic will dramatically overestimate the minimum M value. For example, if a system has 100 candidate lines in unique corridors and 100 existing lines, a 1 degree difference across each line would produce an M value of 200 degrees, well above the absolute maximum value 180. For wide-area planning problems with new generation areas, a more useful heuristic would be to assume a maximum angular distance of 180 degrees.