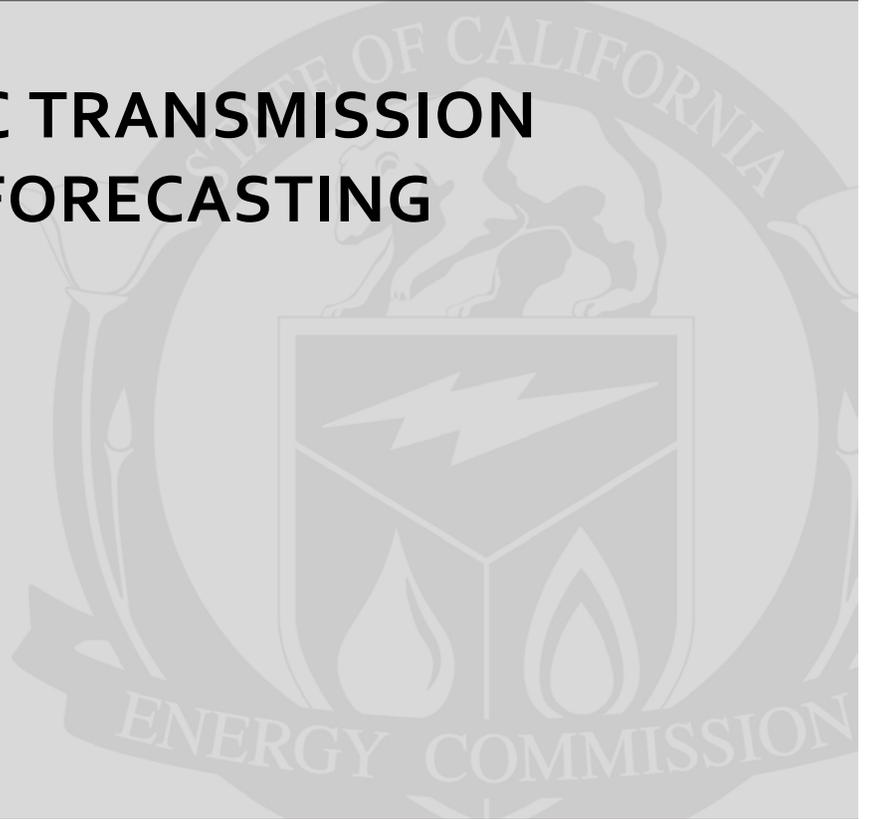


**Energy Research and Development Division
FINAL PROJECT REPORT**

**PROBABILISTIC TRANSMISSION
CONGESTION FORECASTING**



Prepared for: California Energy Commission
Prepared by: Electric Power Research Institute

DECEMBER 2012
CEC-500-2013-120

PREPARED BY:

Primary Author(s):

Steven Lee
Liang Min

Electric Power Research Institute

Contract Number: UC MR-052

Prepared for:

California Energy Commission

Jamie Patterson
Contract Manager

Fernando Pina
Office Manager
Energy Systems Research Office

Laurie ten Hope
Deputy Director
ENERGY RESEARCH AND DEVELOPMENT DIVISION

Robert P. Oglesby
Executive Director

DISCLAIMER

This report was prepared as the result of work sponsored by the California Energy Commission. It does not necessarily represent the views of the Energy Commission, its employees or the State of California. The Energy Commission, the State of California, its employees, contractors and subcontractors make no warranty, express or implied, and assume no legal liability for the information in this report; nor does any party represent that the uses of this information will not infringe upon privately owned rights. This report has not been approved or disapproved by the California Energy Commission nor has the California Energy Commission passed upon the accuracy or adequacy of the information in this report.

ACKNOWLEDGEMENTS

The work reported in this document is a result of collaborative research between EPRI and the California Energy Commission with funding provided from the California Energy Commission.

Participation, technical contribution and review of this project by California ISO are greatly appreciated. It is an important factor for the success of this research project.

The authors wish to thank Jim Detmers, Vice President of Operations and Armando Perez, Vice President of Planning, both of the California ISO, for the support of this project. The vision and research challenges of Virgil Rose, University of California/CIEE project manager, have been instrumental in forming and delivering this research scope.

PREFACE

The California Energy Commission Energy Research and Development Division supports public interest energy research and development that will help improve the quality of life in California by bringing environmentally safe, affordable, and reliable energy services and products to the marketplace.

The Energy Research and Development Division conducts public interest research, development, and demonstration (RD&D) projects to benefit California.

The Energy Research and Development Division strives to conduct the most promising public interest energy research by partnering with RD&D entities, including individuals, businesses, utilities, and public or private research institutions.

Energy Research and Development Division funding efforts are focused on the following RD&D program areas:

- Buildings End-Use Energy Efficiency
- Energy Innovations Small Grants
- Energy-Related Environmental Research
- Energy Systems Integration
- Environmentally Preferred Advanced Generation
- Industrial/Agricultural/Water End-Use Energy Efficiency
- Renewable Energy Technologies
- Transportation

Probabilistic Transmission Congestion Forecasting is the final report for the Probabilistic Transmission Congestion Forecasting project (contract number 500-02-005, work authorization number MR-052) conducted by the Electric Power Research Institute. The information from this project contributes to Energy Research and Development Division's Transmission Research Program and Energy Systems Integration Program.

For more information about the Energy Research and Development Division, please visit the Energy Commission's website at www.energy.ca.gov/research/ or contact the Energy Commission at 916-327-1551.

ABSTRACT

Improved forecasting of transmission congestion in both the short term and the long term can increase the efficiency and reliability of the California electricity system. Improved forecasting requires an approach that accounts for significant uncertainty caused by load and generation forecasts as well as random unplanned equipment outages. California's Renewable Portfolio Standard requires that 20 percent of the state's power be generated by renewable energy by 2020, which will increase the volatility of congestion due to wind generation. Annual congestion costs on the California-Oregon Intertie path increased to \$12 million in 2006 compared to \$6.7 million in 2005, so improvements in congestion management could yield significant cost savings and help achieve California's energy goals.

This report introduced two probabilistic methods for long-term and short-term transmission congestion forecasting that were developed by the Electric Power Research Institute. The proposed method applied sequential Monte Carlo simulation in a probabilistic load flow as the conceptual framework, added all the significant uncertainties and their probability distributions that needed to be modeled, developed the models, and specified how to accurately model the key input assumptions in order to derive valid confidence levels of the forecasted congestion variables. The developed probabilistic method was successfully applied to the four-area Western Electricity Coordinating Council equivalent system. The focus was on the confidence levels of making such forecasts so that a window of forecastability was defined beyond which any forecast would be considered to contain little actionable information. The probabilistic forecasts of congestion would provide confidence limits and information for ranking the potential benefits of alleviating congestion at the various transmission bottlenecks within the window of forecastability.

Keywords: Transmission congestion forecasting, Monte Carlo Simulation (MCS), probability distribution, probabilistic load flow, transmission bottlenecks, Western Electricity Coordinating Council (WECC).

TABLE OF CONTENTS

Acknowledgements	i
PREFACE	ii
ABSTRACT	iii
TABLE OF CONTENTS	iv
LIST OF FIGURES	v
LIST OF TABLES	vii
EXECUTIVE SUMMARY	5
Introduction	5
Project Purpose.....	5
Project Results.....	5
Project Benefits	7
CHAPTER 1: Introduction	8
CHAPTER 2: Long Term Probabilistic Transmission Congestion Forecasting Methodology	10
2.1 Probabilistic Model of Load Demand	11
2.1.1 Model of the Load Increase Uncertainty.....	12
2.1.2 Model Correlations among Load Areas.....	17
2.1.3 Correlation Sampling.....	18
2.2 Probabilistic Model of Generation Capacity	19
2.2.1 Model Generation Units’ In-Service and Retirement Date Uncertainty	19
2.2.2 Model of Wind Generation Peak Condition Uncertainty.....	21
2.3 Economic Dispatch	26
2.4 Summary of Probabilistic Models in Long Term PCF	27
CHAPTER 3: Short Term Probabilistic Transmission Congestion Forecasting Methodology	28
3.1 Probabilistic Model of Load Demand	29
3.1.1 Model of the Day-Ahead Load Forecasting Uncertainty	30
3.1.2 Model Correlations among Load Areas.....	31
3.2 Probabilistic Model of Generation Capacity	32

3.2.1	Generation Forced Outage Rate	32
3.2.2	Model of the Day-Ahead Wind Forecasting Uncertainty	33
3.3	Probabilistic Model of Transmission Capacity	35
3.4	Summary of Probabilistic Models in Short Term PCF and Implementation.....	36
3.4.1	Monte Carlo Simulation in Short Term PCF	37
3.4.2	Monte Carlo Simulation in Long Term PCF.....	37
CHAPTER 4: Introduction of Study Case and Simulation Results.....		38
4.1	Introduction of Study Case.....	38
4.2	Long Term PCF Data Preparation and Simulation Results	41
4.2.1	Load Demands Data Preparation	41
4.2.2	Generation Resources Data Preparation.....	42
4.2.3	Long Term PCF Simulation Results	45
4.2.4	Key Caveats and a Sensitivity Analysis based on Long Term PCF	51
4.3	Short Term PCF Data Preparations and Simulation Results	52
4.3.1	Load Demand Data Preparation.....	52
4.3.2	Generation Resources Data Preparation.....	53
4.3.3	Short Term PCF Simulation Results	54
4.3.4	Key Caveats and a Sensitivity Analysis based on Short Term PCF.....	60
CHAPTER 5: Conclusions and Future Research		64
REFERENCES		68
Appendix A: Literature Survey		A-1
Appendix B: Summary of Probabilistic Models		B-1

LIST OF FIGURES

Figure 2-1: The Conceptual Output of Long Term PCF	10
Figure 2-2: Historical Annual Peak Load Demand Growth Rates (1994 - 2004)	11
Figure 2-3: Forecasted Annual Peak Load Demand Growth Rates (2005 - 2014)	11
Figure 2-4: Historical and Projected Load Demand in WECC	12
Figure 2-5: Probabilistic Increase Exponent Subjected to a Normal Distribution.....	14
Figure 2-6: PDF of next 10 Years Forecasted Load Demand.....	15
Figure 2-7: Projected peak demand annual growth rate	16

Figure 2-8: WECC Actual Summer Peak Demand 1994-2006 Compared with 1998 Projection ...	16
Figure 2-9: Histogram of Actual Summer Peak Demand Growth Rates 1994-2006 Compared with Normal Distribution	17
Figure 2-10: PDF of in-service date for a new generator or retirement date for an old generator	20
Figure 2-11: Wind Speed Data in Four Stations: Oakland (CA), Vandenberg (CA), Denver (CO), and Albuquerque (NM).	22
Figure 2-12: PDF of Wind Speed in Four Stations: Oakland (CA), Vandenberg (CA), Denver (CO), and Albuquerque (NM).....	23
Figure 2-13: Sample Wind Power Curve vs. Wind Speed.....	24
Figure 2-14: PDF of Wind Generation Outputs	25
Figure 3-1: The Conceptual Output of Short Term PCF	28
Figure 3-2: CAISO DA Load Forecasting Error	30
Figure 3-3: Estimated Mean of Hourly Load with Uncertainty Bands.....	31
Figure 3-4: Forecasted Area Loads in CAISO	31
Figure 3-5: Typical Variation of Total and Regional One-Minute Wind Generation in California on a Summer Day (CA ISO, March 2005).....	34
Figure 4-1: WECC Test System Diagram	39
Figure 4-2: Monitored WECC Congestion Paths.....	41
Figure 4-3: Existing and additional non-Hydro Renewable Resource Mix	42
Figure 4-4: Generation Portfolios in the first Year.....	43
Figure 4-5: Generation Portfolios in the 15th Year	44
Figure 4-6: System Load vs. Generation Capacity (Scenario 1) for the CA/MX Area (Top curve is generation, bottom curve with wide ranges is load)	45
Figure 4-7.a: Long Term PCF on COI path without Transmission Upgrade (Scenario 1).....	46
Figure 4-7.b: Long Term PCF on COI path with Hypothetical Transmission Upgrade (Scenario 1)	47
Figure 4-8: System Load vs. Generation Capacity (Scenario 2).....	48
Figure 4-9.a: Long Term PCF on COI path without Transmission Upgrade (Scenario 2).....	49
Figure 4-9.b: Long Term PCF on COI path with Hypothetical Transmission Upgrade (Scenario 2)	50
Figure 4-10: PDF and CDF of Congestion Path Flow on COI path in the 3 rd Year	50
Figure 4-11: The Relationship between Congestion Probability and the Congestion Path Capacity (using the 15 th year congestion forecast of Scenario 2).....	51
Figure 4-12: System Load vs. fixed Generation Capacity.....	53
Figure 4-13: Generation Portfolios in Short Term PCF study	53
Figure 4-14: CA/MX Area System Load vs. Generation Capacity with Typical Variation of Wind Power	55
Figure 4-15: CA/MX Area System Load vs. Generation Capacity with 3370 MW Must-take Generation.....	56
Figure 4-16: Short Term PCF on COI Path (Scenario 1: Generation Capacity with Typical Variation of Wind Generation)	57
Figure 4-17: Short Term PCF on COI Path (Scenario 2: Generation Capacity with Assumed 3370 MW Must-Take Generation).....	57

Figure 4-18: Short Term PCF on WOR Path (Scenario 1: Generation Capacity with Typical Variation of Wind Generation)	58
Figure 4-19: Short Term PCF on WOR Path (Scenario 2: Generation Capacity with Assumed 3370 MW Must-Take Generation).....	59
Figure 4-20: PDF and CDF of congestion Path Flow on COI path at 4:00 p.m.....	60
Figure 4-21: The Relationship between Congestion Probability and Congestion Cost on the COI path at 4:00 p.m. of Scenario 1	61
Figure 4-22: The Relationship between Congestion Probability and the Congestion Path Capacity from the long term PCF Results (same as Figure 4-10).....	62
Figure 4-23: Putting the Long Term and the Short Term PCF Analysis Together for a Financial and Reliability Risk Tradeoff.....	63
Figure 5-1: Putting the Long Term and the Short Term PCF Analysis Together for a Financial and Reliability Risk Tradeoff.....	67

LIST OF TABLES

Table 3-1: Generation Unit FOR.....	33
Table 3-2: Generic Transmission Line FORs.....	35
Table 4-1: Comparison between the Equivalent System and the Actual WECC 2008 Data.....	39
Table 4-2: Assumed Equivalent System Congestion Path Limits	40
Table 4-3: Assumed Mean Values of Load Increase Rates in four WECC Areas.....	41
Table 4-4: Assumed Standard Deviations from Mean Values of Load Increase Rates in four WECC Areas	41
Table 4-5: Future Years' Resource Assumptions of the Study System.....	43
Table 4-6: Summary of Future Years' Resource Assumptions	44

EXECUTIVE SUMMARY

Introduction

Significantly more power transfers are taking place across the Western region, including California due to the restructuring of the power markets in the Western region after 1999. Transmission investment has lagged behind the growth of electricity demand and the increasing amount of power transfers since approximately 1998. As a result, transmission congestion has increased significantly causing very high congestion costs. There are many factors that combine to determine how electricity flows on an interconnected power grid and the uncertainties of these factors compound the uncertainty and the difficulty of forecasting transmission congestion for any particular transmission path. It is important to forecast transmission congestion for various transmission paths to improve the reliability and efficiency of the electricity delivery system in California, both in the short term (e.g., the next 24 hours) and for the long term (e.g., the next 10 to 20 years).

Project Purpose

The goal of this project was to develop new probabilistic methods for forecasting short-term and long-term transmission congestion in the California electric transmission system, which is an integral part of the Western Electric Coordinating Council (WECC) interconnection. The researchers intended to place particular emphasis on identifying congestion and practical considerations about the forward-looking time limit of any forecasting approach due to the inherent uncertainties. The researchers intended that the findings would be used as a building block to develop a comprehensive congestion planning process for California. It was also important to have a methodology and a tool for doing sensitivity studies that would assess the implications of different factors on the overall uncertainty of the congestion level. This would enable policy decisions to be made with informed analyses. A key component of the project's success was the advice and data provided by the California Independent System Operator (ISO) technical staff.

Project Results

The method proposed for this project utilized the Performing Institution's practical experience in power system modeling. It combined the use of analytical functions with regression methods to provide accurate models of the uncertainties, including the effect of correlation. It used a Monte Carlo simulation method to accurately model the physical relationships between generation dispatch, load demand, and the configuration of the transmission grid in order to mathematically predict the key operating constraints of line loading along critical transmission paths in the WECC system, with a focus on the impact of such congestion on the California power grid and consumers. The mathematical models and the time frames of the simulation differ between the short term (24 hours) and the long term (10-20 years), and therefore two computer models were developed to address the two time frames. Each Monte Carlo simulation computed the power flow uncertainties under one particular scenario. Thousands of Monte Carlo simulations were conducted to gain confidence about the variability of the forecasted results of transmission congestion.

The mathematical models of uncertainties in Long-Term Probabilistic Congestion Forecasting (PCF) could be summarized as:

- A load demand model was a chronological exponential expression with time measured in years. The load increase exponent followed normal distribution. The load demand model also considered the correlations among loads.
- The generation capacity was modeled with the uncertainties of generators' installation and retirement time following the discrete exponential distribution.
- Wind generation's spatial correlation and statistical expectation at peak load condition were considered in the long-term planning model. Wind generation outputs were assumed to be subject to a probability distribution.
- Economic dispatch was used to balance the generation and load.

Mathematical models of uncertainties in the short-term PCF could be summarized as:

- Load demand was assumed to follow the normal distribution with time dependent mean value and standard deviation. The hourly load demand model considered the correlation between loads.
- Forced Outage Rate (FOR) was used to express the outage rate of a generation unit to model generation status uncertainties. Each generation unit's available capacity followed the binomial distribution.
- An autoregressive integrated moving average (ARMA) method was used to provide an accurate model of wind uncertainties in short-term planning.
- FOR was also used to express the outage rate of a transmission line to model transmission uncertainties. Transmission line status followed the binomial distribution.
- Economic dispatch was used to balance the generation and load.

Some key caveats from the long-term PCF simulation were:

- The selected approach of probabilistic or deterministic wind generation model in long-term planning will affect the forecasting results.
- Forecasted congestion was dependent on long-term load forecasting.
- Forecasted congestion was also highly dependent on the location and timing of future resources' installation and retirements.
- The forecasted congestion provided significant information regarding incremental improvements and timing of required future transmission upgrade solutions.

Some key caveats from the short term PCF simulation were:

- Integration of wind power into the system had a great influence on the forecasted congestion.

- The highly variable daily patterns of wind power may cause serious congestion and spinning reserve and load following requirements.
- More accurate day-ahead, hour-ahead and minutes-ahead wind forecasts will lead to less uncertainty regarding congestion.
- Forecasted congestion was also highly dependent on load forecasting and generation dispatch.

Project Benefits

Improved forecasting of transmission congestion in both the short term and the long term can increase the reliability and efficiency of the California electricity system. This project developed an approach that accounts for significant uncertainty caused by load and generation forecasts as well as random unplanned equipment outages. California's Renewable Portfolio Standard requires that 20 percent of the state's power be generated by renewable energy by 2020, which will increase the volatility of congestion due to wind generation and present tremendous challenges for the California ISO in managing the power grid. The California ISO reported that annual congestion costs on the California-Oregon Intertie (COI) path increased to \$12 million in 2006 compared to \$6.7 million in 2005. Congestion costs will likely increase even more with the higher penetration of wind power, so improvements in congestion management could yield significant cost savings and help achieve California's goals for reducing greenhouse gas emissions and providing reliable and low-cost electricity to California consumers with minimum environmental impacts.

CHAPTER 1:

Introduction

Congestion occurs when actual or scheduled flows of electricity on a transmission line or a related piece of equipment are restricted below desired levels— either by the physical or electrical capacity of the line or by operational restrictions created and enforced to protect the security and reliability of the grid (DOE 2006). Transmission congestion threatens the system reliability and results in a higher price of energy in the constrained area than in the unconstrained area because of the combined effects of transmission limitations and the cost of local generation.

Transmission congestion occurs when available, least-cost energy cannot be delivered to loads for a period because transmission facilities are not adequate to deliver that energy to some loads. In the short term, when one transmission path to the local electricity demand center is congested, additional electricity demand must be served by other, higher-cost generation sources. To relieve the congestion, system operators need to re-dispatch available generation at higher cost or reduce wholesale transactions to serve local electricity demand. In the long run, when one transmission path to the local electricity demand center is always congested, it means that the local generation sources are either not enough or are too expensive to serve the local electricity demand. To relieve the congestion in this type of situation, system planners need to suggest some new generation sources close to the local electricity demand, distributed resources, demand-side management, or upgrades to the transmission capacity to enable remote generation to serve the local electricity demand.

Congestion cost can be estimated by summing the value of low-cost transactions that cannot be completed due to transmission constraints, and comparing those to the more expensive value of the generation or imports forced by the constraint (DOE 2006). It could be very high. For example, the total annual congestion cost for the California-Oregon Intertie (COI) path were \$6.7 million in 2005 and \$12 million in 2006 (CAISO 2007). Long term transmission and resource planning are very important to alleviate congestion. Besides that, managing the short-term load, generation and transmission outages are also effective ways to manage congestion. Therefore, it is necessary for both system planners and operators to forecast congestion as part of their planning and operation for system reliability and efficiency. Also, system users (generators, transmission owners, and load serving entities) need to forecast congestion to ensure efficient market transactions.

Conventional transmission congestion studies typically use deterministic information, such as the forecasted hourly (or shorter, e.g., 15 minutes) load values of the next few hours, committed generators, scheduled imports and exports, and the system topology in that time frame. With industry restructuring and open transmission access by buyers and sellers of wholesale electricity, this information is not as certain as it used to be when the power system was a vertically integrated system. Therefore, Probabilistic Transmission Congestion Forecasting (PCF) is essential for both long term and short term power system planning. Review of the

presented approaches and discussion of their advantages and drawbacks will help to identify existing or potential new approaches and criteria for PCF.

There are two categories of congestion forecasting techniques. One category is to extend the electricity price forecasting to congestion forecasting. The electricity price forecasting methods include price simulation methods (Hamoud and Bradley 2001; Bastian et al. 1999) and stochastic methods such as artificial neural network (ANN) methods (Hippert et al. 2001; Szkuta et al. 1999; Nogales et al. 2002), time series methods (Contreras et al. 2003), etc. Another category is to adopt probabilistic approach to study load flow (or power flow) problems (Borkowska 1973; Allan et al. 1981; Zhang and Lee 2004; Meliopoulos et al. 1990; Liete da Silva and Arienti 1990), so that system operators are able to forecast the load-flow situation at least for the following day and to identify possible congestion.

In this project, a literature survey was conducted and the results are shown in the appendix of this report. It reviews many papers and reports about new methods that have been published in the technical literature mostly due to the improvement of the long-term generation and transmission planning introduced by the electric power sector deregulation. The literature survey focuses on the long-term generation and transmission planning techniques. It also discusses some related topics, including the energy price forecasting, probabilistic production costing, probabilistic load flow, and security-constrained optimal power flow. In addition, discussion about the available commercial programs for congestion identification is included.

This report introduces a probabilistic method for both long term and short term transmission congestion forecasting, which is recently developed by EPRI. The proposed method applies the sequential Monte Carlo simulation in a probabilistic load flow as the conceptual framework, adds all the significant uncertainties and their probability distributions to be modeled, develops the models, and most importantly specifies how to accurately model the key input assumptions in order to derive valid confidence levels of the forecasted congestion variables. The developed probabilistic method is successfully applied to the four-area WECC equivalent system. Focus is on the confidence levels of making such forecasts, so that a window of forecastability is defined, beyond which any forecast would be considered to contain little actionable information. Within the window of forecastability, the probabilistic forecasts of congestion would provide confidence limits and information for ranking the potential benefits of alleviating congestion at the various transmission bottlenecks.

CHAPTER 2: Long Term Probabilistic Transmission Congestion Forecasting Methodology

As transmission congestion has increased in recent years, so have efforts to improve data collection, analysis and modeling, and methods to improve the efficiency of transmission systems. For a long term time frame (more than a ten-year horizon), FERC rules that regional/state authorities are responsible for the adequacy of generation and transmission resource. However, there are significant uncertainties in the long term congestion forecasting, driven by generation and transmission resource acquisition and regulatory uncertainties. Understanding the influences of these uncertainties on transmission congestion is vital to the reliability and adequacy assessment of bulk power systems.

Long term probabilistic transmission congestion forecast (PCF) anticipates if the transmission path capacity will be sufficient for meeting the demands in the next ten to twenty years, based on forecasted load demand increase, available generator capacity and transmission path capacity, etc. Figure 2-1 gives the conceptual output of long term PCF for planning. The time frame for planning could be from months to years. The solid red curve is the mean value of forecasted transmission path flow. Due to the uncertainties of load increase and generation availability, the actual transmission path flow is also uncertain but still can be predicted. The two dashed red curves are the lower and upper prediction bounds with a certain confidence level of the future transmission path flow.

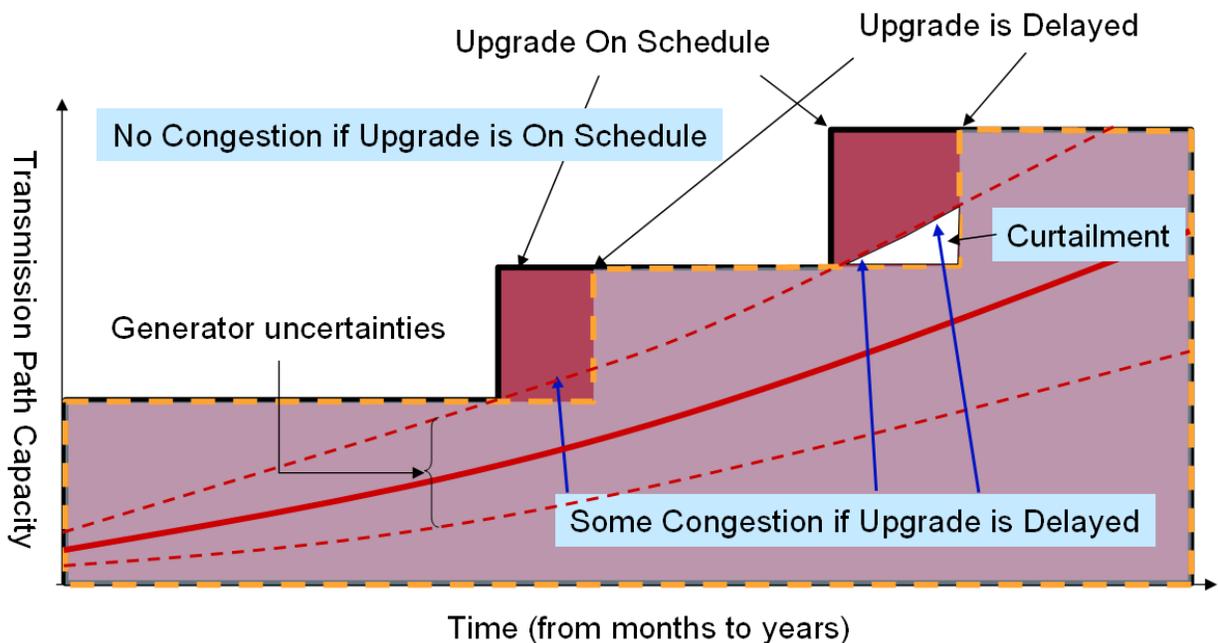


Figure 2-1: The Conceptual Output of Long Term PCF

The dashed yellow curve indicates the transmission path capacity. If the transmission upgrade is on schedule and the available transmission path capacity is larger than the transmission path

flow, congestion can be avoided. Otherwise, if the transmission upgrade is delayed and there is a curtailment between the transmission path capacity and the transmission path flow, congestion occurs.

2.1 Probabilistic Model of Load Demand

Long term PCF analysis depends on numerous probabilistic parameters/inputs. Long term load forecasting is one of the inputs. The U.S. Energy Information Administration’s (EIA) peak demand forecast ten-year compound annual growth rate is 1.7 percent for the summer and 1.5 percent for the winter during 2007 to 2016 in the U.S.¹ WECC’s next 10-Year coordinated plan summary illustrated the 1994-2004 actual load peak demand growth rate and forecasted 2005-2014 load peak demand growth rate, as shown in Figure 2-2 and Figure 2-3².

	<u>1994</u>	<u>1995</u>	<u>1996</u>	<u>1997</u>	<u>1998</u>	<u>1999</u>	<u>2000</u>	<u>2001</u>	<u>2002</u>	<u>2003</u>	<u>2004</u>
WECC – Total	4.4	1.3	5.1	1.3	5.4	-2.0	1.4	-4.5	8.9	2.8	0.8
Northwest Power Pool Area	2.0	8.7	-2.5	-0.4	8.4	-6.6	2.1	-8.1	-0.4	13.5	0.5
Rocky Mountain Power Area	3.4	4.5	1.9	7.1	0.6	-4.2	12.4	8.6	6.0	6.0	-0.8
Arizona-New Mexico-So. Nevada Power Area*	7.3	4.4	4.5	1.8	7.4	-2.3	8.9	7.5	3.7	5.5	0.3
California-Mexico Power Area*	5.9	-0.7	4.2	3.8	4.2	-4.1	-3.6	-5.6	8.0	1.6	5.4

*The 1994 through 1996 projected peak demand growth rate percentages include the Southern Nevada reporting area data in the California-Mexico Power Area data.

Figure 2-2: Historical Annual Peak Load Demand Growth Rates (1994 - 2004)

		<u>2005</u>	<u>2006</u>	<u>2007</u>	<u>2008</u>	<u>2009</u>	<u>2010</u>	<u>2011</u>	<u>2012</u>	<u>2013</u>	<u>2014</u>	<u>2004-2014</u>
WECC - Total	Summer	3.6	2.5	2.7	2.5	2.4	2.4	2.3	2.2	2.3	2.3	2.5
	Winter	3.0	2.3	2.1	2.3	2.0	2.2	2.2	2.0	2.1	1.7	2.3
Northwest Power Pool Area	Summer	-0.8	2.1	2.7	2.2	1.8	1.9	1.9	1.8	2.0	1.9	1.7
	Winter	2.1	2.3	1.8	1.4	1.5	1.7	1.4	1.6	1.5	1.6	1.8
Rocky Mountain Power Area	Summer	4.4	2.7	2.6	2.5	2.4	2.3	2.3	2.4	2.3	2.6	2.6
	Winter	3.6	2.3	2.5	2.2	2.2	2.1	2.1	2.3	2.2	2.3	2.4
Arizona-New Mexico-So. Nevada Power Area	Summer	5.2	3.4	3.2	3.2	3.2	3.2	2.8	2.5	2.7	2.5	3.2
	Winter	2.9	2.9	2.8	3.0	3.0	2.7	2.4	2.4	2.4	2.3	3.0
California-Mexico Power Area	Summer	4.0	1.8	2.3	2.5	2.5	2.5	2.5	2.4	2.4	2.4	2.5
	Winter	2.5	2.1	2.4	2.4	2.4	2.4	2.4	2.4	2.4	2.4	2.4

Figure 2-3: Forecasted Annual Peak Load Demand Growth Rates (2005 - 2014)

¹. The data are from EIA’s report “Annual Energy Outlook 2007, with Projection to 2030”.

². The data and figures are from the WECC 10-Year Coordinated Plan Summary, June 2005.

However, long term load forecasts cannot precisely predict the future and actual load demand will be affected by future weather and economy, etc. Therefore, instead of using a specific forecasted number, a statistical way to represent probabilities of a range of possible future load demand is necessary. Figure 2-4 illustrates the historical and project future load demand. A future year's actual demand may deviate from the mean value of projections due to the inherent variability of the key factors that drive electrical usage. For long term planning purpose, it is useful to have an estimate not only of the expected mean value of possible future outcomes, but also of the distribution of probabilities around the projection³.

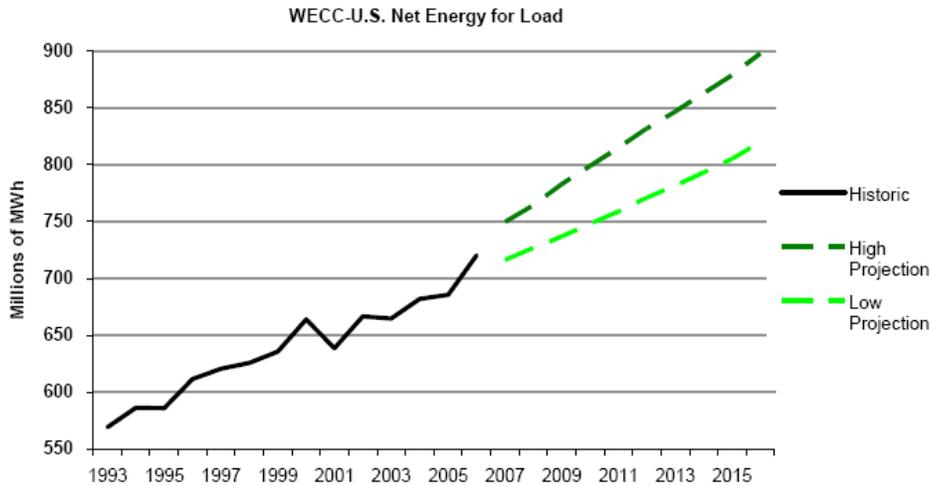


Figure 2-4: Historical and Projected Load Demand in WECC⁴

2.1.1 Model of the Load Increase Uncertainty

The entire power system is separated into several load areas. The annual peak load demand of every area is assumed in this study to change with the same proportion to WECC, and peak load increase exponent is subjected to probabilistic distribution. This report decomposes the WECC system according to the definition of report areas in the WECC, which are:

- Northwest Power Pool Area (NWPP)
- Rocky Mountain Power Area (RMPA)
- Arizona-New Mexico-Southern Nevada Power Area (AZ/MX/SNV)
- California-Mexico Area (CA/MX)

³ Accordingly, the Load Forecast Working Group (LFWG) develops upper and lower 10 percent confidence bands around the NERC region demand and energy projection. This means that there is an 80% probability that future demand and energy will occur within these bands. Concurrently, there is a 10% chance that future outcome could be less than the lower band and a 10 percent chance that future outcomes could be higher than the upper band.

⁴ The figure is from the NERC 2007 Long Term Reliability Assessment, October 2007.

The area projection of load demand is modeled as a function of its last year's peak load demand and the load demand growth rate. It is assumed that the annual peak load of one area is increased with a time series exponential expression as the following equation:

$$L_i(t) = L_i(t-1)e^{\lambda_i(t)} \quad (2-1)$$

where $L_i(t-1)$ is the value of the last year's peak load of the i^{th} area;
and $\lambda_i(t)$ is the increase exponent of this year which is decided by the load demand growth rate.

Load growth is a complex process which basically is the summation of a very large number of variables involving electricity customers' addition of things that use electricity and their electricity usage patterns which may respond to uncertain factors such as weather, etc. Thus it is simplistic to assume a constant growth rate. In Equation (2-1), a growth rate that may change each year is assumed instead. It should be recognized that the growth rate may be a negative value as well as a positive value. The challenge is in estimating these growth rates. Econometric models are typically used to make such forecasts.

Because the load demand growth rate is uncertain and cannot be precisely predicted, for the purpose of this research project, the exponent λ_i is treated as a probabilistic variable which is assumed to follow the normal distribution. (Note that there is a reasonable argument for the normal distribution because of the Central Limit Theorem⁵, which states that the sum of a large number of independent random variables, no matter what their own probability distributions are like, will approach the normal distribution.) The probability distribution of λ_i can be obtained from the statistical analysis of annual peak load growth rates. The mathematical relationship between exponent λ_i and annual peak load demand growth rate α_i is shown in the following equation:

$$\lambda_i = \ln(1 + \alpha_i) \quad (2-2)$$

The Figure 2-5 shows an example of probabilistic increase exponent subjected to normal distribution ($\mu = 0.0204, \sigma = 0.0326$).

⁵ "Central Limit Theorem." From MathWorld--A Wolfram Web Resource.
<http://mathworld.wolfram.com/CentralLimitTheorem.html>

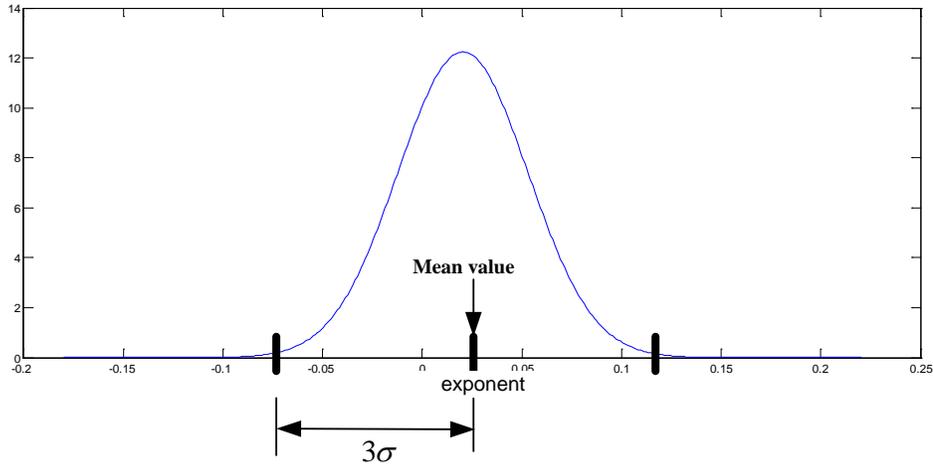


Figure 2-5: Probabilistic Increase Exponent Subjected to a Normal Distribution

Figure 2-6 shows the probability density of the forecasted load demand based on the annual increase exponent shown in Figure 2-5. To build this probabilistic model, some variables need to be inputted.

- $L_i(0)$ - The current load peak demand value;
- μ_i - The mean value of increased exponent of the future forecasted load;
- σ_i^2 - The variance of increased exponent of the future forecasted load.

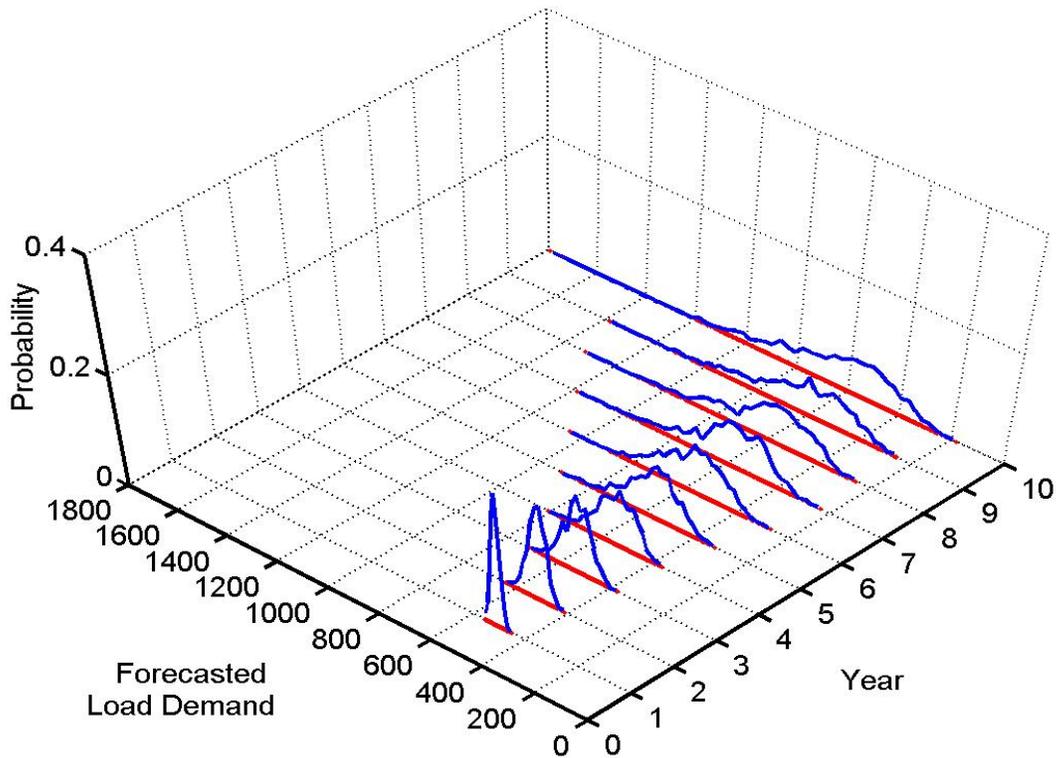


Figure 2-6: PDF of next 10 Years Forecasted Load Demand

To check how well the normal distribution fits the annual load demand growth rate, an analysis of the WECC historical summer peak demand between 1998 and 2006 was conducted.

Figure 2-7 illustrates the projected 1998-2007 peak demand annual growth rate made by WECC in its 10-year plan summary published in 1998. In Figure 2-8, the actual peak demand growth from 1998 to 2006 are shown, as well as the upper and lower range, the two dashed black curves representing the lower and upper prediction bounds with 95 percent confidence level. The solid blue curve is the actual WECC summer peak demand⁶. From this figure, we can see that the actual WECC summer peak demands are within the lower and upper prediction bounds, which were based on the assumption of a normal distribution for the annual growth rate with a standard deviation of 0.03. This standard deviation was obtained from the limited samples of the annual growth rates from 1998 to 2007.

⁶ The data are from the 2007-2016 Regional & National Peak Demand and Energy Projection Bandwidths.

	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
WSCC-Total	1.7	1.9	1.5	1.9	1.5	1.8	1.8	1.6	1.4	1.6

Figure 2-7: Projected peak demand annual growth rate⁷

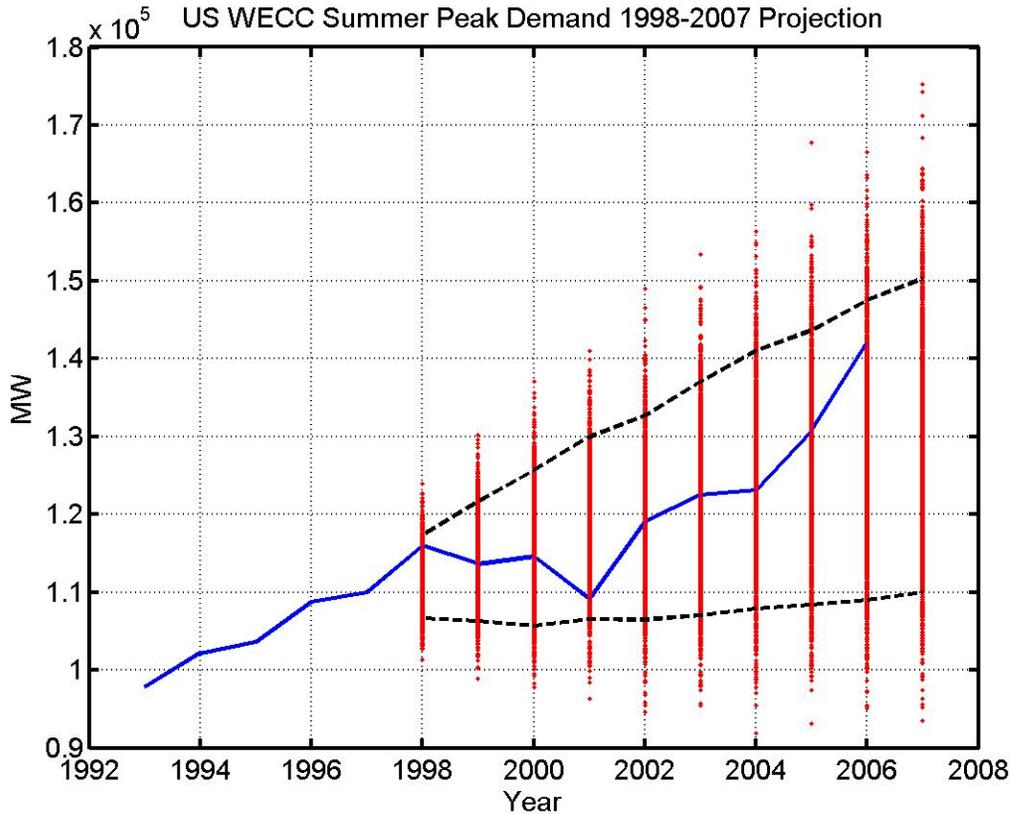


Figure 2-8: WECC Actual Summer Peak Demand 1994-2006 Compared with 1998 Projection

Figure 2-8 does not show the shape of the probability distribution. In order to verify how good the normal distribution is for modeling the actual peak demand growth, the historical actual growth rates between 1994 and 2006 are plotted as a histogram in Figure 2-9. It can be seen that even with the limited number of samples, the general shape of the histogram is approximately symmetrical and centered on a mean value, which fits a bell-shape curve reasonably well.

⁷The data are from the WECC 10-year plan summary 1999-2008.

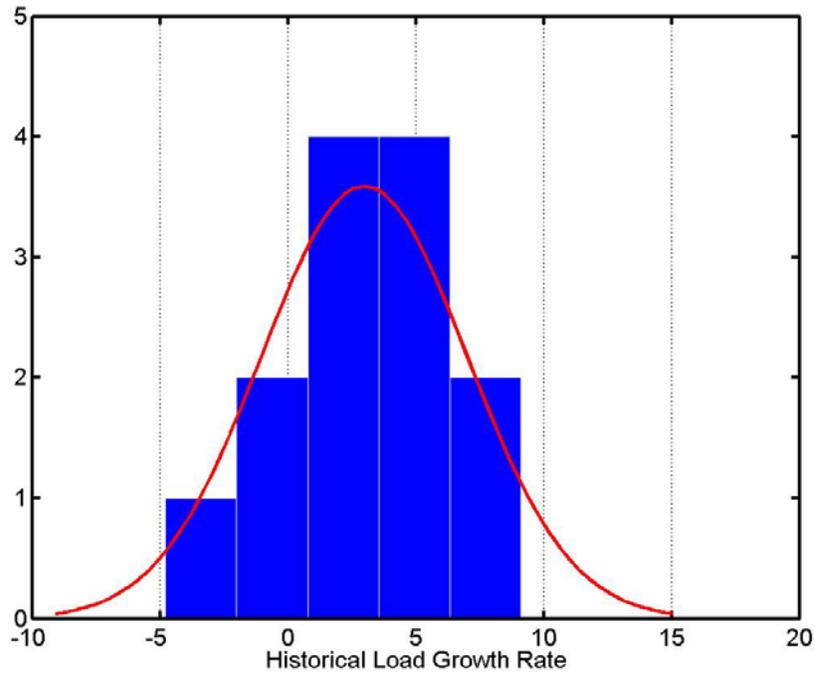


Figure 2-9: Histogram of Actual Summer Peak Demand Growth Rates 1994-2006 Compared with Normal Distribution

2.1.2 Model Correlations among Load Areas

The annual peak load demands are significantly affected by the economic and weather factors, and obviously these factors are correlated among neighboring areas. It is necessary to assume there are correlations among the load demand of neighboring areas, which should be reflected on the probabilistic load demand model in the equation 2-1. These correlations are expressed in a covariance matrix, where each covariance component reflects the correlation between the peak load demand increase rates of two areas. In the long term PCF, we define that $\text{cov}(\lambda_i, \lambda_j)$ is the covariance between i^{th} area increase exponent λ_i and j^{th} area increase exponent λ_j .

Given n sets of variables denoted $\{\{X_1\}, \{X_2\}, \dots, \{X_n\}\}$, the covariance matrix is defined by

$$\text{Cov}_{ij} = \text{cov}(X_i, X_j) = \langle (X_i - \mu_i)(X_j - \mu_j) \rangle \quad (2-3)$$

where $\mu_i = \langle X_i \rangle$ and $\mu_j = \langle X_j \rangle$ are the respective means of $\{X_i\}$ and $\{X_j\}$; Cov_{ij} is called the covariance of $\{X_i\}$ and $\{X_j\}$.

The covariance between peak load demand increase rates can be calculated by the historical peak load demand increase information. Each area's historical peak load demand increase rate in Figure 2-2 is used as a set of variables denoted in the equation 2-3. Then the covariance matrix can be calculated. For example, the covariance matrix for the four report areas in WECC according to the Figure 2-2 data is:

$$Cov = \begin{matrix} NPPA \\ RMPA \\ ANMS \\ CMPA \end{matrix} \begin{bmatrix} 20.8068 & 1.4455 & 3.1267 & 4.4551 \\ 1.4455 & 10.6680 & 5.6490 & -2.2666 \\ 3.1267 & 5.6490 & 6.1043 & -0.6668 \\ 4.4551 & -2.2666 & -0.6668 & 10.6061 \end{bmatrix} \times 10^{-3} \quad (2-4)$$

2.1.3 Correlation Sampling

Mathematically, if four areas' load increase exponents are completely dependent, only one normally distributed random number is required to be drawn in order to determine the load increase exponents for all areas. If four areas' load increase exponents are independent of each other, it is necessary to independently draw the different normally distributed random numbers for each area's load increase exponents. From a practical operation point of view, sets of variables are neither completely dependent nor independent but there is some correlation between each other. In this case, it is necessary to generate a normally distributed random vector in which each component corresponds to a particular area's load increase rate and whose components satisfy the specified correlation.

We adopted the correlation sampling technique in (Li and Billinton 1999) in this study. Let R be an N dimensional normally distributed random vector with mean B and covariance matrix C . Let G be an N dimensional normally distributed random vector whose components are independent of each other, and each component has a mean of zero and a variance of unity. Linear combination of normal distributions is still a normal distribution. There exists therefore a matrix A which can create the following transformation relationship between R and G :

$$R = AG+B \quad (2-5)$$

The mean value vector and the covariance matrix of R can be calculated from the following equation:

$$E(R) = A E(G) + B = A 0 + B = B \quad (2-6)$$

$$E[(R-B)(R-B)^T] = E[AGG^T A^T] = E(AA^T) = AA^T = C \quad (2-7)$$

The equation (2-7) gives the relation between matrix A and matrix C . On the other hand, covariance matrix C is a non-negative definite symmetry matrix and the matrix C can be triangularized into the unique lower triangular matrix multiplied by its transposed matrix. Consequently, in the following equation,

$$C = AA^T \quad (2-8)$$

where A is a lower triangular matrix.

Bus load correlation sampling technique can be summarized in the following steps:

- 1) Generate an independent random vector G , which follows the standard normal distribution and has N dimension. Each component of vector G corresponds to a bus load.

- 2) Calculate the lower triangular matrix A by Cholesky factorization.
- 3) Create a correlative N dimension random vector R from the equation (2-5). Each component of R still corresponds to a bus load and can have different mean values and variances. There exists the correlation between all bus loads that is defined by matrix C .

2.2 Probabilistic Model of Generation Capacity

Generation resource forecasting is another important input for the long term PCF analysis. WECC's Loads and Resources Subcommittee categorize additional generation resource according to their status:

- Committed resources: all resources presently under active construction with expected in-service dates;
- Uncommitted resources: all resources are currently undergoing regulatory approval;
- All resources that generally have identified in-service dates and locations but do not fall into the previous two categories.

However, there are significant uncertainties in the in-service dates and even the retirement dates, driven by construction and regulatory uncertainties.

To meet the Renewable Portfolio Standard (RPS) goal, wind generation is becoming a large part of the generation resource mix. Wind generation resources are non-bid and must-take generation when integrating with system grid. However, besides the construction and regulatory uncertainties, the maximum wind power output is uncertain due to its intermittent characteristic and it will never reach its maximum output simultaneously.

2.2.1 Model Generation Units' In-Service and Retirement Date Uncertainty

All the existing, new and retired generators in the system are considered for long term generation capacity. Similar to the impact and range of possible demand, the magnitude of the total available generation resources can be expected to fall within a range of uncertainty due to the uncertainties of in-service and retirement dates. These uncertainties, such as the exact time of new generation station's installation and old generation station's retirement, lead to the uncertainty of generation capacity in a specific future year. To consider and model these uncertainties in the long term transmission congestion forecast, we assume that the in-service and retirement dates are subject to an exponential distribution.

The probability density function of the in-service date for an installed new generator or the retirement date of an old generator has the following form:

$$f(x) = \begin{cases} \lambda e^{-\lambda(x-t_0)}, & x \geq t_0 \\ 0, & x < t_0 \end{cases} \quad (2-9)$$

where t_0 is the planned in-service date or the planned retirement date; λ is the delay rate.

Since this is a discrete annual assessment and the exponential distribution is a continuous analog of the [geometric distribution](#), the [geometric distribution](#) is used in this report. The

probability density function of the in-service date for an installed new generator or the retirement date of an old generator has the following form:

$$f(x) = \begin{cases} \lambda(1-\lambda)^{(x-t_0)}, & x \geq t_0 \\ 0, & x < t_0 \end{cases} \quad (2-10)$$

By sampling the in-service date and retirement date in Monte Carlo simulation, the generation capacity for a specific year can be calculated in a probabilistic way.

In order to build this model, the following variables need to be decided:

- t_0 : The planned in-service date or the planned retirement date.
- λ : The delay rate parameter of the in-service date or the retirement date of one specific generator

Delay rate λ is the key parameter of this exponential distribution model for in-service or retirement date. The value of delay rate should be decided by the user based on the historical data or reasonable assumptions. For example, if there is a 12 percent probability of delay in this project, the value of λ will be set as 0.88, The PDF of the in-service date for a new generator or the retirement date for an old generator is as shown in Figure 2-10.

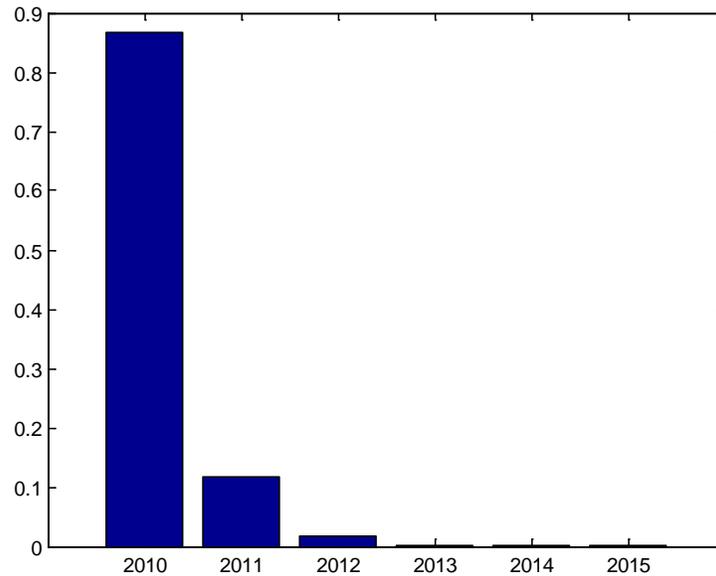


Figure 2-10: PDF of in-service date for a new generator or retirement date for an old generator

2.2.2 Model of Wind Generation Peak Condition Uncertainty

To meet the Renewable Portfolio Standard (RPS) goal, the transmission infrastructure has to be expanded to accommodate the renewable penetration levels defined in the state's renewable energy policy. Otherwise, the power grid will have congestion problems if transmission infrastructure changes are behind the construction schedule of renewable resources.

Wind generation is becoming a large part of the generation resource mix. However, the intermittent characteristic of wind power output is a major challenge to power system planning and operating with a high penetration of wind generation. Practically, the combined output will never reach their maximum output simultaneously due to the uncertainty over time and geographical areas (question of time and spatial correlation). Thus, long term transmission congestion forecast should evaluate wind generation's spatial correlation and statistical expectation of their peak conditions.

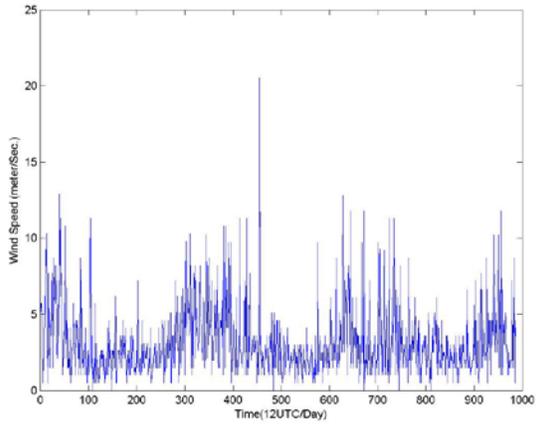
- Statistical Expectation of Peak Condition

The importance of statistical expectation motivates the estimation of the Probability Density Function (PDF) of peak condition for wind power output, rather than just a single point forecast (peak value). The PDF of peak condition for wind power output can be calculated by using the historical data.

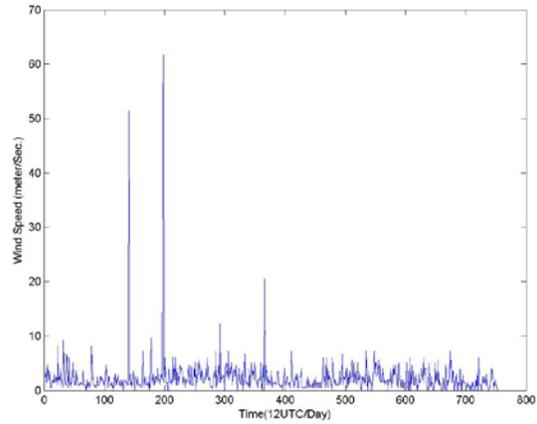
For this study, wind speed data (from year 2001 to year 2006) from Forecast Systems Laboratory (FSL)⁸ were used as historical data. For each station, the data were collected at 0:00 and 12:00 UTC of each day. In the west coast of U.S., wind speed builds in the afternoon as the marine-layer spreads from the high-pressure region over the cool Pacific Ocean to the low pressure regions over the hot interior valleys. It then reaches a peak in early evening, and begins to fall off during the early morning hours reaching a minimum. So, we picked the data collected at 12:00 UTC of each day, which is 4:00 a.m. standard Pacific Time to investigate the PDF of peak wind speed. *This limited approach is to test the idea of using wind speed statistics to investigate the uncertainty of wind speed over a long term of several years, i.e., seasonal uncertainty. It is also intended to provide some data to look into the spatial diversity over the long distances between Northern and Southern California and other adjacent regions to California. This limited analysis is not intended to be a definitive study.*

Given the wind speed of those stations, we performed statistical analysis and calculated PDF of wind speed at those stations. For example, Figure 2-11.a shows the actual wind speed measured at Oakland, CA every day 12:00 UTC from January 1, 2001 to December 31, 2006. Figures 2-11 b, c and d show the actual wind speed over the same time period and the same measuring hour for Vandenberg, S. California, Denver and Albuquerque.

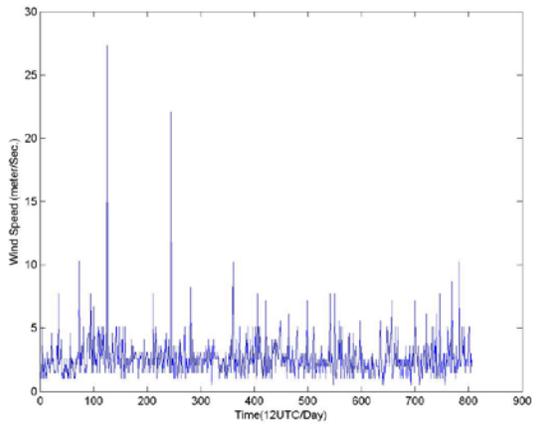
⁸ The data are from the FSL, Radiosonde Data Archive, [Online]. Available: <http://raob.fsl.noaa.gov/>.



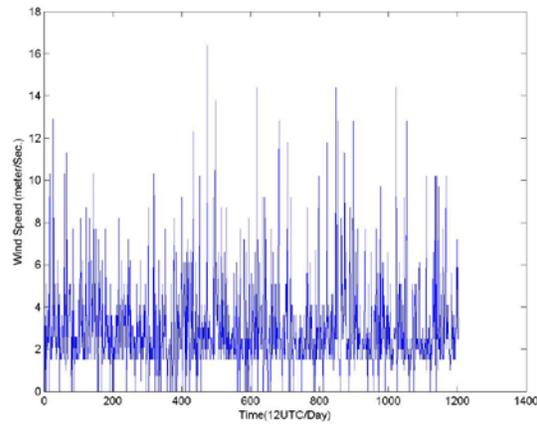
a. Oakland (CA) 2001-2006



b. Vandenberg (CA) 2001-2006



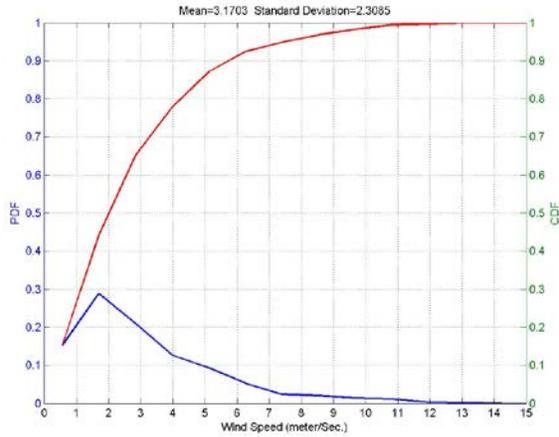
c. Denver (CO) 2001-2006



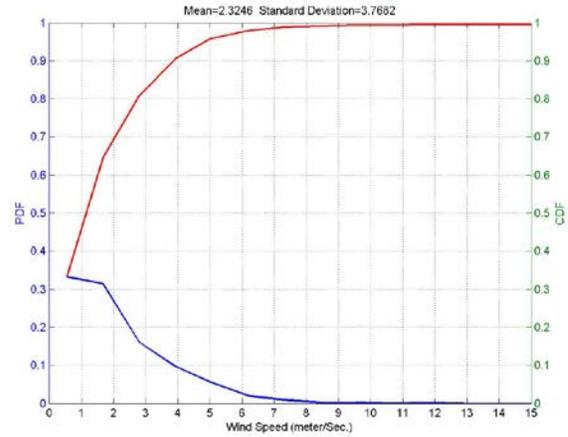
d. Albuquerque (NM) 2001-2006

Figure 2-11: Wind Speed Data in Four Stations: Oakland (CA), Vandenberg (CA), Denver (CO), and Albuquerque (NM).

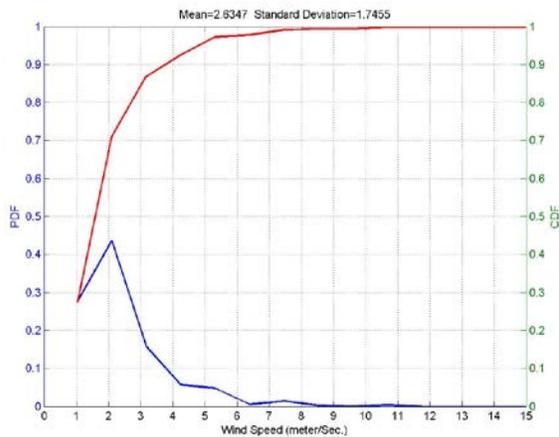
Note that the height above sea level of the wind speed measuring point in Oakland (CA), Vandenberg (CA), Denver (CO) and Albuquerque (NM) is 140 meters, 100 meters, 1,611 meters and 1,619 meters, respectively.



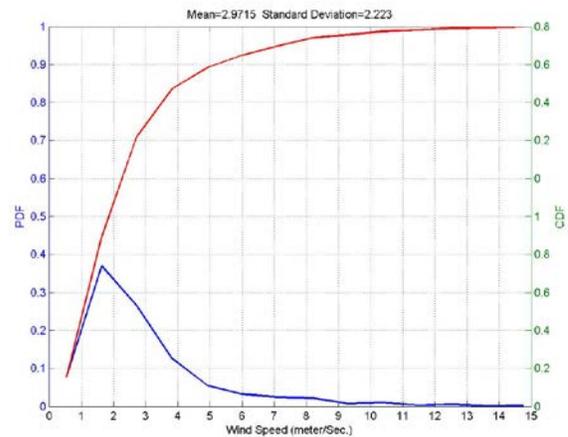
a. Oakland (CA) Wind Speed PDF



b. Vandenberg (CA) Wind Speed PDF



c. Denver (CO) Wind Speed PDF



d. Albuquerque (NM) Wind Speed PDF

Figure 2-12: PDF of Wind Speed in Four Stations: Oakland (CA), Vandenberg (CA), Denver (CO), and Albuquerque (NM).

Figure 2-12.a shows the probabilistic density function of wind speed in Oakland, CA. Figures 2-12 b, c and d show the probabilistic density functions of wind speed over the same time period and the same measuring hour for Vandenberg, S. California, Denver and Albuquerque. The mean wind speed in Oakland (CA), Vandenberg (CA), Denver (CO) and Albuquerque (NM) is 3.17m/s, 2.32m/s, 2.63m/s and 2.97m/s respectively⁹. The distribution functions of wind speed at these four locations are close to the Rayleigh Distribution, which is generally applied to model the annual or monthly wind speed distribution. For a typical Rayleigh distribution, the probability density function of wind speed has the following form:

⁹ 1m/s = 2.2369mph.

$$f(x, \sigma) = \frac{x}{\sigma^2} e^{\left(\frac{-x^2}{2\sigma^2}\right)} \quad (2-11)$$

where σ is the parameter; x is the wind speed.

This statistical analysis is just based on the actual wind speed data measured at 12:00UTC every day instead of hourly wind speeds data. Because there may be some anomalous wind conditions during the 24 hours measurement gap, the mean and modal wind speed of 24 hours could be quite different if it were measured hourly.

Generally, most manufacturer nameplate outputs are taken at around 28-30 mph (12.5-13.5 m/s). However, we cannot always assume that wind generation output is at its nameplate output. Figure 2-13 shows a typical wind power output curve of a 1000 MW wind turbine from a manufacturer. We project the statistical characteristic of wind speed in Oakland (CA) onto the wind power output. Figure 2-14 shows the PDF of the wind generation output, which is close to exponential distribution. To consider and model wind power output uncertainty in the long term transmission congestion forecast, it could be assumed that the wind generation outputs in long term planning model follow exponential distribution.

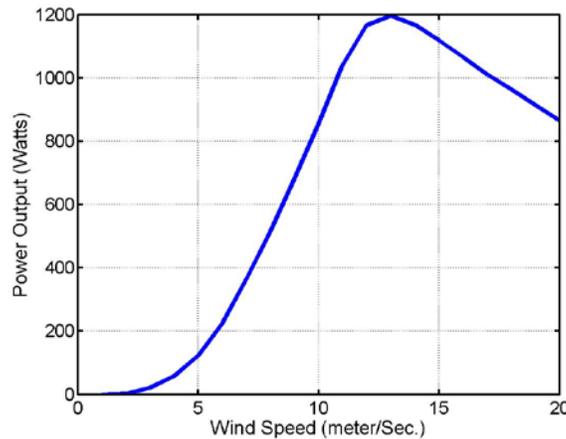


Figure 2-13: Sample Wind Power Curve vs. Wind Speed¹⁰

¹⁰ <http://www.bergey.com/Technical/XL.1.R.xls>

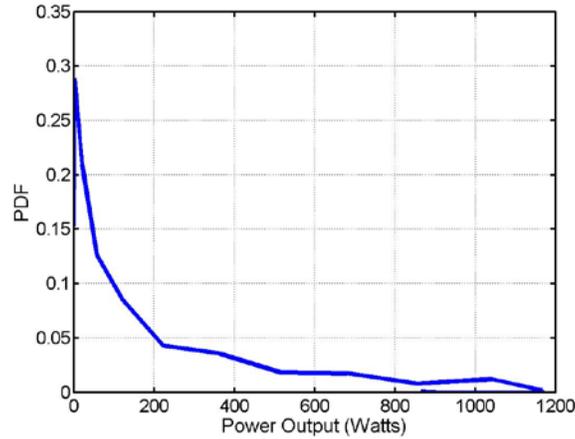


Figure 2-14: PDF of Wind Generation Outputs

- Spatial Correlation

Wind resource variability cannot be considered random and uncorrelated to each other. Besides the above statistical analysis of the peak condition of wind speed, it is also possible to investigate how wind speed correlations change with geographical location.

The covariance between wind speeds at those four stations can be calculated by the actual wind speed measurements at 12:00UTC every day from January 1, 2001 to December 31, 2006. The covariance matrix for the four stations analyzed is:

$$Cov = \begin{matrix} ABQ \\ DEN \\ OAK \\ VBG \end{matrix} \begin{bmatrix} 0.4996 & -0.0145 & -0.0030 & -0.0619 \\ -0.0145 & 0.3056 & -0.0066 & 0.0334 \\ -0.0030 & -0.0066 & 0.2840 & -0.0258 \\ -0.0619 & 0.0334 & -0.0258 & 1.43701 \end{bmatrix} \times 10^3$$

The normalized covariance matrix called correlation coefficients is:

$$corr = \begin{matrix} ABQ \\ DEN \\ OAK \\ VBG \end{matrix} \begin{bmatrix} 1.0000 & -0.3905 & -0.2278 & -0.4597 \\ -0.3905 & 1.0000 & -0.3615 & -0.1522 \\ -0.2278 & -0.3615 & 1.0000 & -0.4024 \\ -0.4597 & -0.1522 & -0.4024 & 1.0000 \end{bmatrix}$$

The correlation coefficient is a normalized measure of the strength of the linear relationship between two variable sets. Uncorrelated data results in a correlation coefficient of 0; equivalent data sets have a perfect correlation coefficient of 1. So, from this limited analysis, there are no strong correlations among wind speeds of these four locations.

2.3 Economic Dispatch

With the probabilistic models of load demand and generation available capacities already defined, the actual scheduled output power of generators need to be determined by simulation to balance the load. In this model, economic dispatch is used to balance load and generation. Economic dispatch is a reasonably valid approach to approximate a market operation, wherein generators are free to bid at any price. In the long term, no generator can sustain a bid that is below its real variable operating cost, and competition will induce generators to bid close to their real variable operating costs in order not to lose market shares. Therefore, assuming that generators bid at their real variable operating costs is a reasonable approach at least for simulating the effect of the dispatch on the power flows of the grid.

The basis for the economic dispatch is a table of lambda (variable energy plus variable O&M cost) versus unit generation output. When the system lambda is less than the lowest lambda of a thermal unit, the output of that unit is set to zero or to its minimum capacity. When the system lambda is greater than the highest lambda of a unit, the output of that unit is set to its maximum. By summing across a row corresponding to a value of system lambda, the total economic generation capacity that can serve load at that system lambda cost can be found. If the total generation at that system lambda is less than the system load at that time period, then the estimated value of the system lambda will be raised, and the sum of generation will be re-determined. If the total generation at that system lambda is greater than the system load, then the estimated value of the system lambda will be lowered, and the sum of generation will be re-determined. This iterative process will converge when the sum of generation at the final system lambda equals the system load. The procedure of searching for the system lambda (optimal dispatch point) is called Equal Lambda Method of Economic Dispatch (EPRI 2001). For example, the following table lists six generators with their capacities, types and energy costs (variable O&M costs assumed negligible in this example.)

Unit	Capacity (MW)	Type	Energy Cost (\$/MWh)
G1	200	Hydro	25
G2	300	Nuclear	20
G3	500	Coal	40
G4	60	Natural Gas	150
G5	30	Solar	400
G6	40	Wind	46

The economic dispatch order is listed in the following table according to the energy cost. If all these generators are available and the total system load is 1100MW, the system energy cost is 150\$/MWH, which is equal to the energy cost (\$/MWH) of the last generator (G4) used to meet the load prevailing at this time. So, the outputs of G1, G2, G3, G4 and G6 are set to their own capacity; the output of G5 is set to zero.

Unit	Capacity (MW)	Type	Energy Cost (\$/MWh)
G2	300	Nuclear	20
G1	200	Hydro	25
G3	500	Coal	40
G6	40	Wind	46
G4	60	Natural Gas	150
G5	30	Solar	400

Practically, all the nuclear generators are in a must-run status and at their maximum production. Slow-starting thermal units are also in a must-run status and operating at their minimum output levels because they are required for future operating hours. Wind generation is also in non-bid and must-take status and is with its intermittent characteristic. Hydro generation is at high production levels.

2.4 Summary of Probabilistic Models in Long Term PCF

Probabilistic models of uncertainties in Long Term Probabilistic Congestion Forecasting (PCF) can be summarized as:

- Load demand model can be described as a chronological exponential expression. The load increase exponent follows normal distribution. The load demand model also considers the correlations among loads.
- The generation capacity is modeled with the uncertainties of generators installation and retirement time following the discrete exponential distribution.
- Wind generation's spatial correlation and statistical expectation of peak condition will be considered in long term planning model. Wind generation outputs at the system peak load are assumed to follow a probability distribution.
- Economic dispatch is used to balance the generation and load.

CHAPTER 3: Short Term Probabilistic Transmission Congestion Forecasting Methodology

Congestion problems are growing in the operational planning and real-time market operation. The total annual congestion costs for the California-Oregon Intertie (COI) path were \$6.7 million in 2005 and \$12 million in 2006 (CAISO 2007). In a short term time frame (from minutes to weeks), when a remote generation center transfers electricity to a load center to meet its load demand, such as the transfer of power from the Pacific Northwest to California. Because the transmission capacity into the load center is limited, when additional contingencies occur, e.g. losing major transmission lines or power plants in the load center during peak load periods, it may cause congestion problem and a need to re-dispatch generation and curtail market transactions.

For the short term congestion forecast purpose, the known information is the forecasted hourly (or shorter, say, 15 minutes) load values of the next few hours, the committed generators, and the system topology during that time frame. However, there are significant uncertainties, driven by load forecasting errors, unplanned forced outages of generation units and transmission lines. Short term probabilistic transmission congestion forecast (PCF) is to forecast the system congestion over the time frame from the next 24 hours to next few weeks, to assess present congestion condition, to provide a leading indicator of the near future congestion margin, and to guide decision-making associated with congestion management, thus helping the operation planner to adjust the generation dispatch or scheduled transactions with external market areas to relieve the congestion. Under extreme conditions, demand reduction through public appeals, brown outs or rotating blackouts may be necessary.

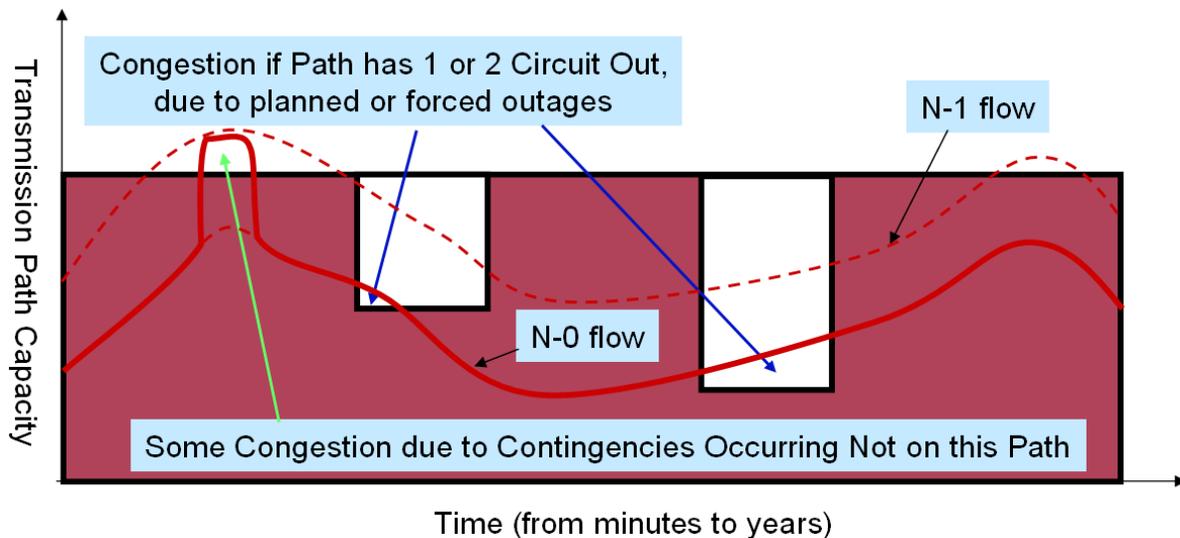


Figure 3-1: The Conceptual Output of Short Term PCF

Figure 3-1 gives the conceptual output of short term PCF for operational planning. The time frame for operation planning and real-time market could be from minutes to days. Just like the long term PCF, the solid red curve is the mean value of forecasted transmission path flow. Due to the uncertainties of load, transmission and generation availability, the actual transmission path flow is also uncertain but still can be predicted. The most important additional uncertainty modeled in the short term PCF compared to the long term PCF is the random forced outages of transmission lines. These outages can happen at any time but only last for hours or days at most, i.e., right in the time frame of the short term PCF. Congestion could occur if a specific line on the congestion path is out of service due to forced outage. Congestion may also occur due to contingencies happening on other transmission lines not on this congestion path, which through the effect of the parallel flows of the transmission grid causes flows to increase on other paths.

3.1 Probabilistic Model of Load Demand

In the real-time mode, the hourly (or half-hourly) values of the load for the specified forecast period are predicted. These forecast data are used to drive the basic scheduling functions of the EMS or to provide dispatcher information. Real-time mode execution of the forecasting procedure uses the historical load and weather data files, and forecast weather data, automatically received or entered by the forecast engineer, and real-time tele-metered data, if applicable. The CAISO utilizes an Automated Load Forecasting System to calculate its Day-Ahead (DA) Hourly Forecasted Demand approximately 14 hours prior to the next operating day. Multiple weather forecasting data sources are used to determine the weather forecast.

The accuracy of the forecasted load depends on the lead time and the explanatory variables. Normally, it is expected that the load used in this model is forecasted just before doing the short-term congestion forecasting. It should be very accurate because it is forecasted with a much shorter lead time. But there are still a number of uncertainties associated with the load forecasting such as substantial change of weather forecasts, occurrence of abnormal events, etc. Figure 3-2 that illustrates that approximately 10 percent of the DA hourly forecasting error was greater than 1,000 MW in 2006.

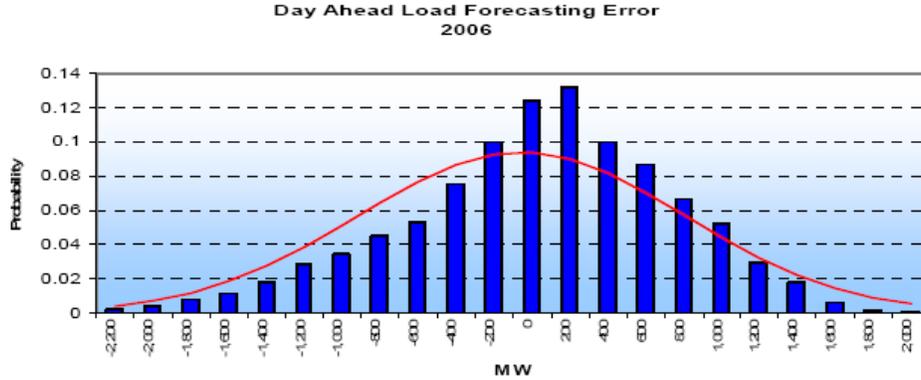


Figure 3-2: CAISO DA Load Forecasting Error¹¹

3.1.1 Model of the Day-Ahead Load Forecasting Uncertainty

The actual load demand is uncertain and might be different from the DA forecasted value, but it still can be predicted. Based on our observation, load deviation to the forecasted value is not an unchanged value at different time. So, it is assumed that the forecasted load follows the normal distribution at a certain time. Mean value and standard deviation of the hourly load are functions of time:

$$L(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu(t))^2}{2\sigma(t)^2}} \quad (3-1)$$

where

$\mu(t)$ is the mean value of the day-ahead (DA) hourly forecasted load at time t ;

$\sigma(t)^2$ is the variance of the hourly forecasted load at time t that is introduced to consider the load forecasting errors.

If we generate Monte Carlo samples using the distribution in equation (3-1), the forecasted load values are uncertain but fall within the range of $[\mu(t) - 1.96\sigma(t), \mu(t) + 1.96\sigma(t)]$ with probability 0.95. So we can draw three curves as shown in Figure 3-2. The solid curve shows the mean value of hourly load, while the two dashed curves indicate the upper and lower ranges of the 95 percent confidence level. So at each time T , we will have three points in these three curves, M is the mean value point, U is the upper range point and L is the lower range point. The load at time T has 0.95 probability of falling within points U and L .

¹¹ This figure is from the CAISO Integration of Renewable Resource Report.

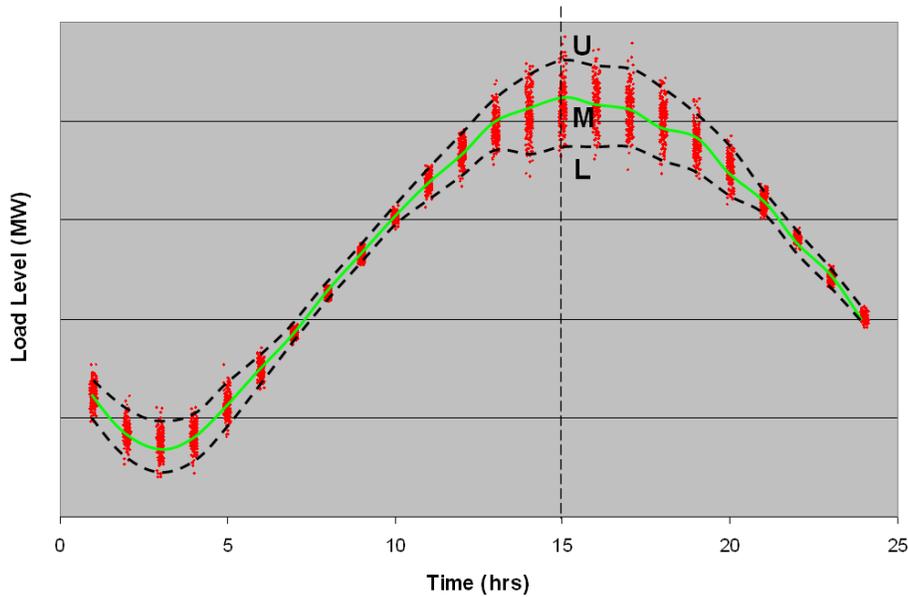


Figure 3-3: Estimated Mean of Hourly Load with Uncertainty Bands

3.1.2 Model Correlations among Load Areas

In the operation environment, the correlations among load demands will be strong due to common environmental factors such as temperature, daylight hours, weather fronts, etc., and to social factors such as sporting events, meal time, end-use habits, etc. As these factors are likely to affect all loads of a similar nature in a like manner, a degree of correlation will exist.

Generally, the correlation among loads can be calculated by the consecutive hourly loads data. Each load's hourly load data is treated as a set of variables denoted in Equation (2-3). Equation (2-3) then can be used to calculate the covariance matrix. Bus load correlation sampling technique described in section 2.1.3 can be used to generate the random variables for Monte Carlo simulation. For example, Figure 3-2 shows the forecasted next 24-hour loads of SCE, PG&E, and SDGE areas in CAISO, which was forecasted on August 1, 2007.

Forecasted UDC Direct Access Load																								
Hour Ending	HEB1	HEB2	HEB3	HEB4	HEB5	HEB6	HEB7	HEB8	HEB9	HE10	HE11	HE12	HE13	HE14	HE15	HE16	HE17	HE18	HE19	HE20	HE21	HE22	HE23	HE24
Utility Distribution Co	Megawatts (MW)																							
08/01/2007																								
SCE	615.65	605.21	599.16	608.46	635.44	671.41	734.91	778.64	813.18	839.75	866.36	877.56	887.70	888.51	886.13	876.80	863.94	844.19	827.54	822.16	798.97	739.75	672.90	640.86
PG&E	1146.07	1110.14	1086.77	1084.82	1108.22	1159.09	1246.62	1323.33	1394.24	1449.27	1503.58	1539.21	1566.12	1593.99	1616.44	1624.92	1613.37	1574.49	1523.98	1484.41	1437.61	1386.82	1277.68	1202.68
SDGE	162.37	158.96	156.62	159.47	166.09	178.89	197.85	214.59	229.86	240.42	249.12	251.06	254.41	255.08	252.88	248.11	240.45	232.83	223.90	218.26	211.32	198.04	182.26	170.94

Figure 3-4: Forecasted Area Loads in CAISO¹²

The covariance matrix for these three area's load is:

¹² The figure is captured from the OASIS of CAISO; [Online]. Available: <http://oasis.caiso.com/>.

$$Cov = \begin{bmatrix} 1.1465 & 1.9998 & 0.3793 \\ 1.9998 & 3.7125 & 0.6569 \\ 0.3793 & 0.6569 & 0.1275 \end{bmatrix} \times 10^4$$

The normalized covariance matrix called correlation coefficients is:

$$Corr = \begin{bmatrix} 1 & 0.9693 & 0.9923 \\ 0.9693 & 1 & 0.9549 \\ 0.9923 & 0.9549 & 1 \end{bmatrix}$$

The correlation coefficient is a normalized measure of the strength of the linear relationship between two variable sets. Uncorrelated data results in a correlation coefficient of 0; equivalent data sets have a correlation coefficient of 1. Clearly, if there is a strong linear correlation among these reporting areas' peak load demand increases, the results are close to 1. So, the correlations among SCE, PG&E, and SDGE's load demands are very strong due to common environmental factors and social factors.

3.2 Probabilistic Model of Generation Capacity

Different from the long term congestion forecasting, un-planned generation forced outages and un-predictable patterns of wind variability bring significant uncertainties into the short term congestion forecasting. Similar to the impact and range of possible load demand, the magnitude of the available generation resources can be expected to fall within a range of uncertainty due to the forced outages of generation units and the variation of wind generation outputs.

3.2.1 Generation Forced Outage Rate

To model generation uncertainties, a forced outage rate (FOR) is used to express the outage rate of a generation unit. Forced Outage Rate controls [generator](#) forced outages, a partial or complete loss of generating capability for a certain period of time. The method of convolution is well known and used in generation reliability models, e.g., for Loss Of Load Probability (LOLP) models, which uses the binomial distribution for individual generator outages and combine them into a discrete probability mass function of the total available generation capacity.

We use the open access information from CEC and WECC to define generation units' FORs in WECC 179-bus system. WECC Power Supply Assessment¹³ mentioned that an 8 percent FOR was applied to the non-hydro generation units in the zones associated with the CAISO and LADWP due to a high proportion of older thermal units and stringent environmental levels. In all other zones the seasonal non-hydro capacities were reduced by 5 percent. This difference is

¹³ WECC Power Supply Assessment, June 14, 2004. [Online]. Available: http://www.wecc.biz/documents/library/publications/PowerAssessment/PowerSupplyAssessment_BoardApproved_8-04.pdf.

supported by historical loads and resources data for actual year forced outage capabilities. As more of the older plants retire in California, the average forced outage rate will likely be decreased. The average forced outage rate (FOR) of hydro generation units in the Northwest Power Pool area was 2.5 percent. Therefore, the FORs used in this case study are assumed in Table 3-1.

Table 3-1: Generation Unit FOR

	Hydro	Non-Hydro
NWPP	2.5%	5%
CA/MX	3%	8%
RMPA	3%	5%
AZ/NM/SNV	3%	5%

3.2.2 Model of the Day-Ahead Wind Forecasting Uncertainty

Figure 3-5 illustrates the fluctuations of one-minute wind power delivered to the California grid during a single 24-hour period on a summer day in California (CA ISO, 2005). The California daily wind generation profile in Figure 3-5 is typical of a summer day in the areas affected by the coastal marine-layer. Wind speed and generation build up during the afternoon as the marine-layer spreads from the high-pressure region over the cool Pacific Ocean to the low pressure regions over the hot interior valleys. It then reaches a peak in early evening, and begins to fall off during the early morning hours reaching a minimum between about 10:00 a.m. and 2:00 p.m. (EPRI 2006).

Unfortunately, the daily patterns of wind power are opposite to the typical daily system load curve in Figure 3-1. During the evening, wind generation is reaching its peak when system load is reaching its trough. During the day time, the reverse situation exists. When load is increasing rapidly in the morning, wind generation is diminishing to its minimum.

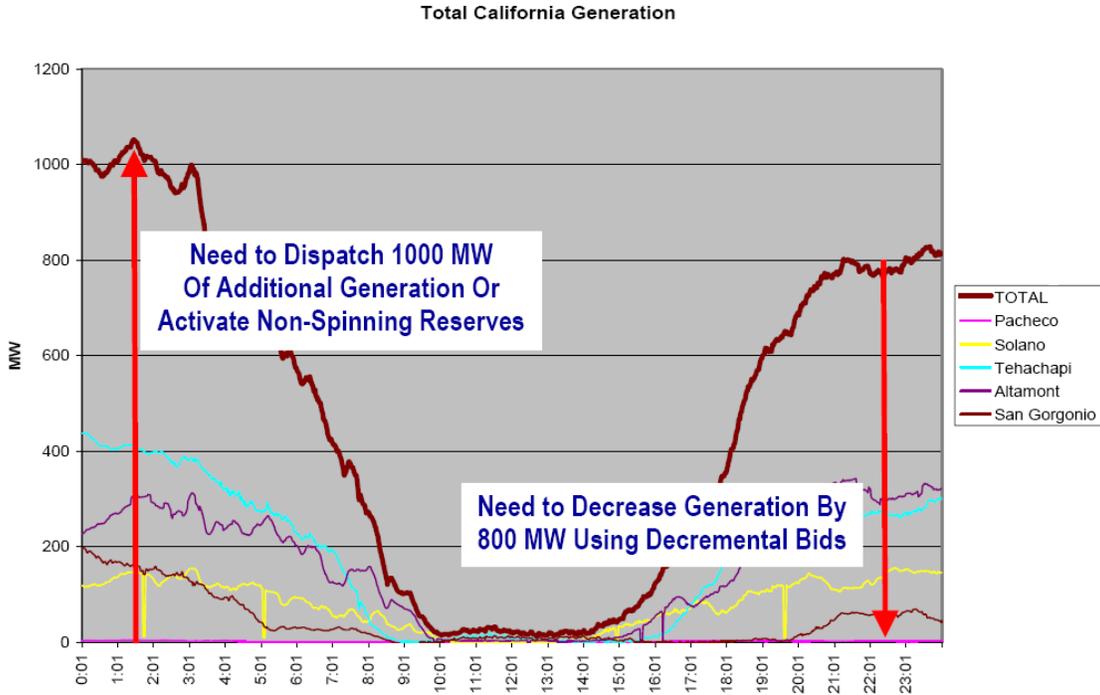


Figure 3-5: Typical Variation of Total and Regional One-Minute Wind Generation in California on a Summer Day (CA ISO, March 2005)¹⁴

The above wind power curves are based on measurements. For the research project, we need to develop a model. For the congestion forecasting study, we will also consider the wind forecasting error in this study¹⁵. Day-Ahead (DA) wind forecasting error could be 15 percent of the total installed wind generation capacity or 50 percent of the forecasted wind power production.

Another factor to consider is that the wind power outputs from different locations cannot be assumed to be independent of one another. It will be affected by common weather conditions. Some wind power output prediction methods also use statistical time-series models. For example, Autoregressive integrated moving average (ARMA) model in (Milligan et al. 2003) is shown as:

$$X_t = \sum_{j=1}^p a_j X_{t-j} + \sum_{k=0}^q b_k e_{t-k} \quad (3-2)$$

¹⁴ This figure is from the CEC report "California Regional Wind Energy Forecasting System Development".

¹⁵ This is not development of wind forecasting methods, but rather considering the forecasting error in the study and specifying the uncertain impact on system operation and planning.

where the current time t 's observation X_t depends on a linear combination of past observations of X_{t-j} plus a moving average of series e_{t-k} , which is a white-noise process characterized by zero mean and variance as function of time. p is the order of the autoregressive process of X on itself and q is the order of the moving-average error term.

We will make use of these considerations in the short term PCF model for wind generation uncertainty.

3.3 Probabilistic Model of Transmission Capacity

Similar to the probabilistic model of generation with the consideration of FORs, FOR is also used to express the outage rate of a transmission line to model transmission uncertainties.

Generic outage probability data are used for this study. The generic probability file contains the parameters that enable the model to compute unavailability rates of various power system components.

As an example, the FOR of a line is estimated using the following equation:

$$FOR = Outage\ Freq \times Repair\ Time / 8760 \quad (3-3)$$

In the above equation, the outage frequency is estimated using the following equation:

$$Outage\ Freq = a + b \times (Z / ZpuPerMile) \quad (3-4)$$

where a (1/year) is the constant parameter of the forced outage frequency, b (1/year/mile) is the proportional parameter of the forced outage frequency, $ZpuPerMile$ (pu/mile) is the average impedance (p.u.) per mile used to estimate the line length, $RepairTime$ (hour) is the average repair time (hour) after a forced outage.

The meaning of FOR is that if the FOR of a transmission line is 0.5 percent, it implies that on the average this line will be out of service (OOS) $0.005 \times 8760 = 52.8$ hours every year. And, the repair time is 8 hrs, meaning that there will be on the average $52.8 / 8 = 7$ random outage events every year.

The required input data are the FORs for transmission facilities. Utilities can calculate the FORs based on the historical transmission outage data. Table 3-2 illustrates a general FOR for transmission lines which was developed by EPRI's Reliability Initiative in 2001.

Table 3-2: Generic Transmission Line FORs

[VnomInf(kV)	VnomSup(kV)]	Zpu/L (pu/mile)	a (1/year)	b (1/year/mile)	RepairTime (hour)
765	9999	0.0001	0.4	0.01	8
735	765	0.000117	0.39434	0.011132	8

500	735	0.00025	0.35	0.02	8
345	500	0.0005	0.3	0.03	8
315	345	0.000761	0.286957	0.032609	8
287	315	0.001004	0.274783	0.035043	8
230	287	0.0015	0.25	0.04	8
220	230	0.001717	0.244203	0.040725	8
161	220	0.003	0.21	0.045	8
144	161	0.003739	0.202609	0.048696	8
138	144	0.004	0.2	0.05	8
121	138	0.004944	0.181111	0.054722	8
118	121	0.005	0.18	0.055	8
115	118	0.0055	0.18	0.055	8
0	115	0.005	0.18	0.055	8

3.4 Summary of Probabilistic Models in Short Term PCF and Implementation

Probabilistic models of uncertainties in short term PCF can be summarized as:

- Load demand is assumed to follow the normal distribution with time dependent mean value and time dependent standard deviation. The load demand model considers the correlation between loads.
- Forced Outage Rate (FOR) is used to express the outage rate of a generation unit to model generation availability uncertainties. Each generation unit's available capacity follows the binomial distribution.
- Autoregressive integrated Moving Average (ARMA) method is used to provide an accurate model of wind uncertainties in short term planning.
- FOR is also used to express the outage rate of a transmission line to model transmission uncertainties. Transmission line status follows the binomial distribution.
- Economic dispatch is used to balance the generation and load, as described in the long term PCF.

Appendix B summarizes the probabilistic models and provides comments about the reasonableness of such models.

3.4.1 Monte Carlo Simulation in Short Term PCF

For the short term PCF, by far we have already defined the probabilistic models of load demand, generation capacity, and transmission network. A sequential Monte Carlo method based load flow will be used to simulate the probabilistic problem.

For every hour, perform the following steps (1-5):

- 1) Generate a Monte Carlo sample-set for load, generation, and transmission status according to the probabilistic models from section 3.1 to section 3.3.
- 2) Define the outputs of all generators using economic dispatch in section 2.3.
- 3) AC power flow computation is performed to get the power flow conditions of entire power system.
- 4) The program will aggregate the line flows on the specific flow path to evaluate whether there is congestion or not by comparing it with path limit.
- 5) Return to Step 1; begin a new trial until 5000 trials are finished.

For every hour and every Monte Carlo trial, there will be a “Yes-or-No” result for the congestion on a specific path. By repeating the trials enough times such as 5000 trials in each hour, a probabilistic distribution of congestion on each path is obtained.

3.4.2 Monte Carlo Simulation in Long Term PCF

The differences between long term and short term PCF lie in the mathematical models and the time frames of the simulation. For the long term PCF simulation, we use the probabilistic models of load demand and generation capacity that have been defined in Chapter 2. The key simulation engine is the same as the short term PCF, which is a sequential Monte Carlo method based load flow, but the Monte Carlo based load flow for the long term PCF is performed for the peak load condition of each year.

It should be noted that the objective of the Monte Carlo simulation is not only to obtain an expected value of congestion. The main objective is to obtain the probability distribution of the degree of congestion on each path. With the full probability distributions, much more can be learned about the confidence level of various congestion levels. It should also be noted that the probability distributions are functions of time. This method will provide the ability to characterize when in the long term future would various confidence levels of certain degree of congestion take place.

CHAPTER 4:

Introduction of Study Case and Simulation Results

The developed probabilistic forecasting methods have been applied to the equivalent California electric transmission system, operating as an integral part of the Western Electric Coordinating Council interconnection. This equivalent power grid model of the WECC was developed by EPRI in the past for research purposes, and has been used by a number of researchers in the academic research community. In this study, particular emphasis has been placed on identifying congestion and practical considerations about the forward-looking time limit of any forecasting approach due to the inherent uncertainties.

The mathematical models and the time frames of the simulation differ between the short term (24 hours) and the long term (10-20 years). The simulations of the two models, long term and short term PCF, will be presented in this chapter. The first section is to demonstrate the long term PCF method with the assumed generation and transmission infrastructure construction plans. The second section is to demonstrate the short term PCF method when the system status is assumed to be in the 15th year with 15 percent penetration of wind power. Both methods use Monte Carlo simulation to accurately model the probabilistic and physical relationships among generation dispatch, load demand and configuration of the transmission grid in order to mathematically predict the key operating constraints of line loading along critical transmission paths in the WECC system.

4.1 Introduction of Study Case

The methods described in the preceding sections are applied to an equivalent WECC test system, which consists of 179 buses and 263 lines. This system consists of 4 reporting areas, Northwest Power Pool Area (NWPP), Rocky Mountain Power Area (RMPA), Arizona-New Mexico-Southern Nevada Power Area (AZ/NM/SNV) and California-Mexico Power Area (CA/MX). The reduced WECC 179-bus system is decomposed as shown in Figure 4-1.

Because the equivalent is different from the actual system in terms of size, Table 4-1 lists the comparison between the equivalent and the actual 2008 WECC data. The assumed equivalent WECC system congestion path limits are listed in Table 4-2. Figure 4-2 displays the congestion paths on the WECC map.

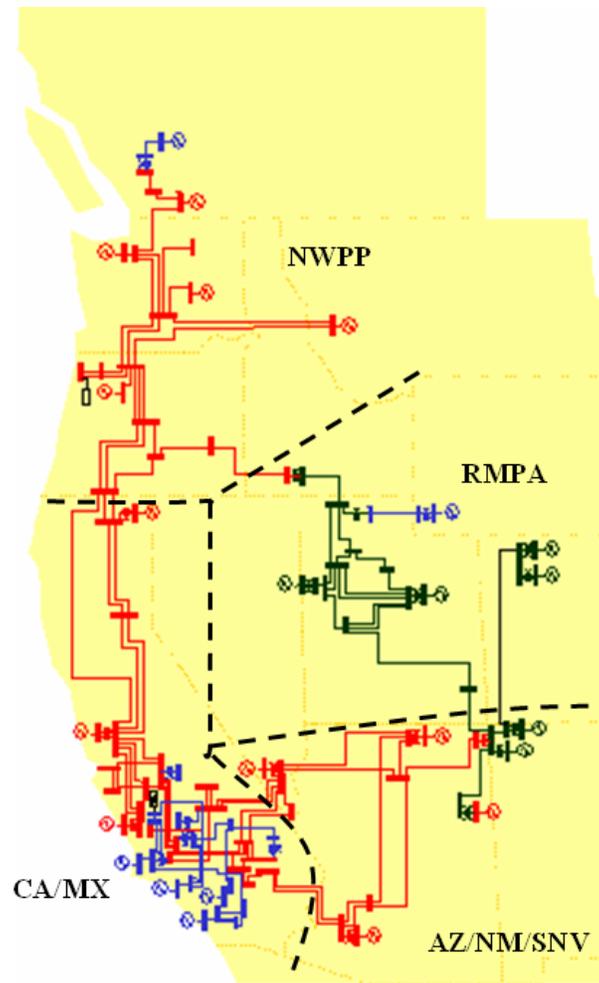


Figure 4-1: WECC Test System Diagram

Table 4-1: Comparison between the Equivalent System and the Actual WECC 2008 Data

	System Load (MW)		System Resource (MW)
	Summer Peak	Winter Peak	
WECC 2008 Model ¹⁶	160,237	136,926	209,211
Equivalent System	60,677		68,004

¹⁶ The data are from Western Interconnection 2006 Congestion Study, Prepared by the Western Congestion Analysis Task Force, May 08, 2006. [Online]. Available: http://nietc.anl.gov/documents/docs/DOE_Congestion_Study_2006_Western_Analysis.pdf

Ratio	38%	44%	33%
-------	-----	-----	-----

Table 4-2: Assumed Equivalent System Congestion Path Limits

Congestion Area	Path Name	Assumed Forward Limit
NWPP to Canada	NW to Canada	2000 MW
NWPP to CA/MX	COI	4800 MW
AZ/NM/SNV to CA/MX	WOR	8000 MW
N-S/S-N CA	Midway to Vincent	3600 MW
WY to UT	Ben Lomond to Midpoint	800 MW
RMPA to AZ/NM/SNV	TOT2A	750 MW

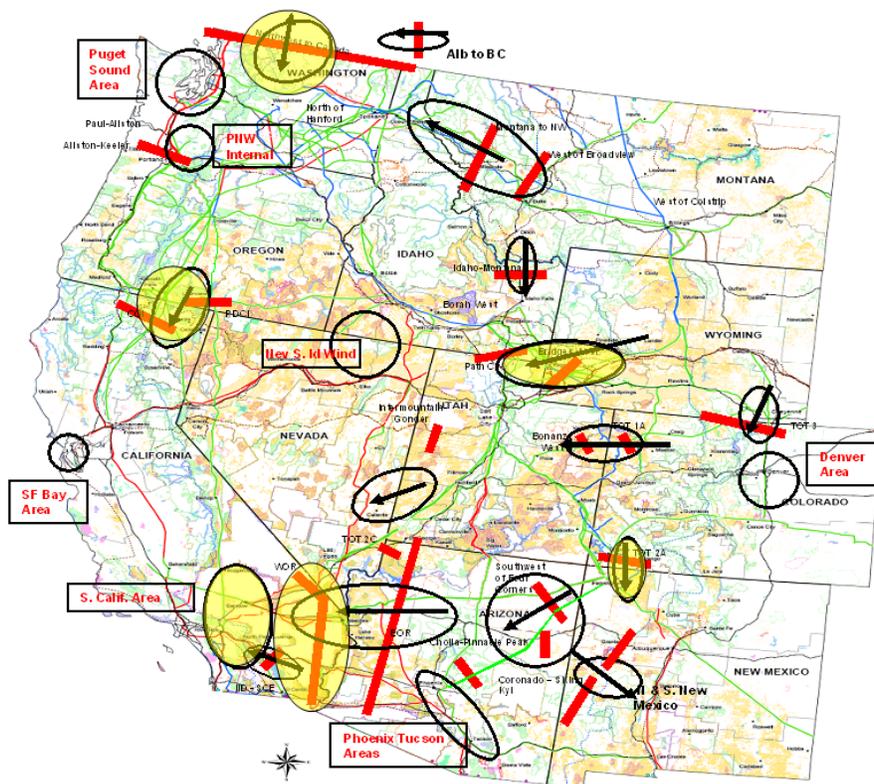


Figure 4-2: Monitored WECC Congestion Paths¹⁷

4.2 Long Term PCF Data Preparation and Simulation Results

4.2.1 Load Demands Data Preparation

The four reporting areas’ mean values of load increase rates are illustrated in Table 4-3. It is assumed that the standard deviation of load increase rates of different years is the same, which is illustrated in Table 4-3. The covariance matrix for the four reporting areas in WECC is:

$$Cov = \begin{matrix} NPPA \\ RMPA \\ ANMS \\ CMPA \end{matrix} \begin{bmatrix} 20.8068 & 1.4455 & 3.1267 & 4.4551 \\ 1.4455 & 10.6680 & 5.6490 & -2.2666 \\ 3.1267 & 5.6490 & 6.1043 & -0.6668 \\ 4.4551 & -2.2666 & -0.6668 & 10.6061 \end{bmatrix} \times 10^{-3}$$

After the annual peak load demand of one area is decided, the annual peak load demand on each bus is assigned. The entire load demand will be distributed to each bus in this area, based on their original proportions in the base case.

Table 4-3: Assumed Mean Values of Load Increase Rates in four WECC Areas

Year	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
CAMX	2.5	2.5	2.5	2.5	2.4	2.4	2.4	2.4	2.4	2.4	1.2	1.2	1.2	1.2	1.2
AZNM	3.2	3.2	3.2	2.8	2.5	2.7	2.5	2.5	2.5	2.5	1.5	1.5	1.5	1.5	1.5
NWPP	2.2	1.8	1.9	1.9	1.8	2	1.9	1.9	1.9	1.9	1	1	1	1	1
RMPP	2.5	2.4	2.3	2.3	2.4	2.3	2.6	2.6	2.6	2.6	1.3	1.3	1.3	1.3	1.3

Table 4-4: Assumed Standard Deviations from Mean Values of Load Increase Rates in four WECC Areas

	Standard Deviation
NWPP	0.0456

¹⁷ This figure is from the U.S. Department of Energy’s report “National Electric Transmission Congestion Study”.

RMPP	0.0327
AZNM	0.0247
CAMX	0.0326

4.2.2 Generation Resources Data Preparation

To simulate the activities of meeting the 20 percent CA Renewable Portfolio Standard (RPS) goal as illustrated in Figure 4-3, it is assumed that the wind generation penetration in CA/MX area is 5 percent in the beginning year. (This is approximately the amount of wind power penetration in California in 2008, so year 1 in this study may be considered as 2008.) Then, some key wind plants will be built and interconnected into CA/MX at Northern California (Altamont area) and at Southern California (Tehachapi area). So, by the last simulated year (the 15th year), the wind generation penetration in CA/MX area will have reached 15 percent of the total generation capacity in this area.

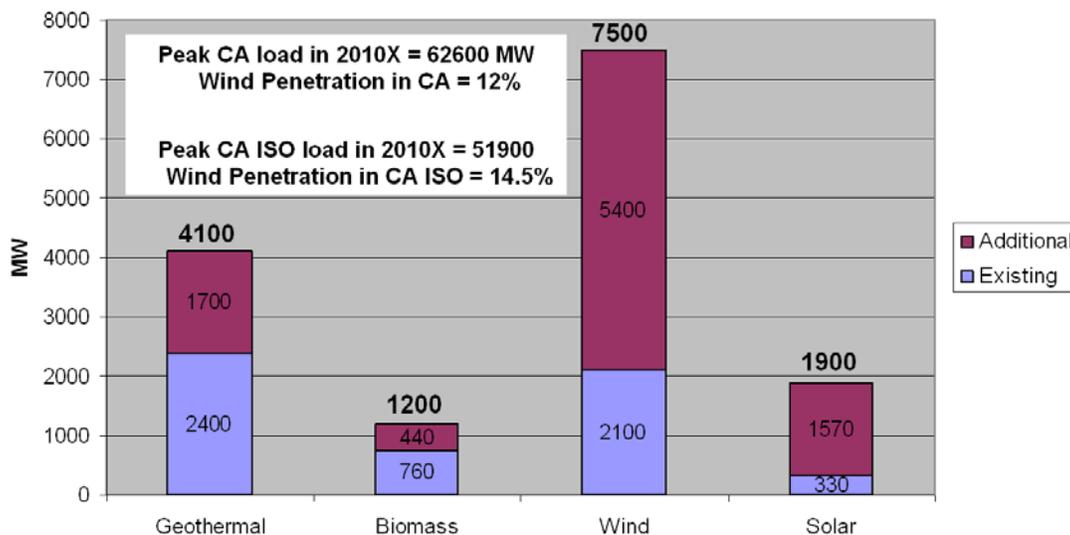


Figure 4-3: Existing and additional non-Hydro Renewable Resource Mix

Figure 4-4 shows the current year generation portfolios of four reporting areas. A significant portion of generation resource in the NWPP area is derived from hydro generation. Coal-fired generation in the RMPP and the AZ/NM/NV areas is very significant. The CA/MX area is dominated by gas-fired plants and wind generation is 5 percent of the total generation capacity in that area. Table 4-3 and Table 4-4 show the future generation resource construction plans. We can see that wind generation increases rapidly in Northern California (Altamont area) and Southern California (Tehachapi area) during the first four years. Figure 4-5 shows the 15th year

generation portfolios of four reporting areas. The wind generation penetration in CA/MX area will reach 15 percent of the total generation capacity in this area.

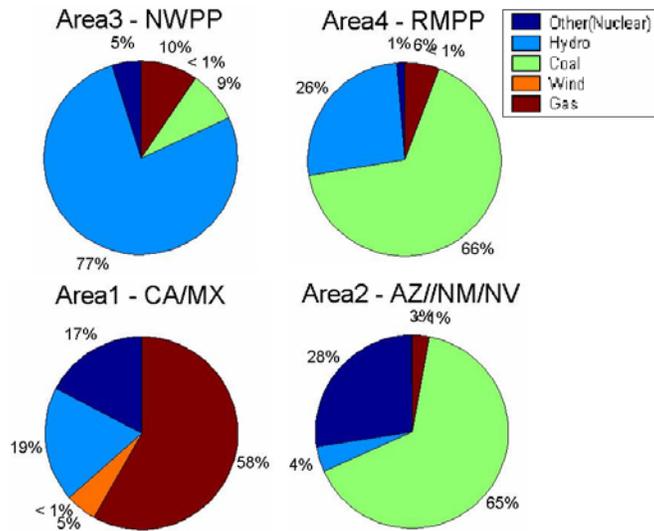


Figure 4-4: Generation Portfolios in the first Year

Table 4-5: Future Years' Resource Assumptions of the Study System

Scheduled Year	Station	Status	Type	Amount (MW)
2	MOHAV1CC (Bus148)	Retired	Coal	1580
2	PARDEE (Bus149)	Added	Wind	500
2	ROUND MT (Bus112)	Added	Wind	500
3	DIABLO1 (Bus103)	Added	Wind	500
3	HAYNES3G (Bus43)	Added	Gas	800
3	PARDEE (Bus149)	Added	Wind	500
4	PARDEE (Bus149)	Added	Wind	500
4	PARDEE (Bus149)	Added	Gas	750
4	ROUND MT (Bus112)	Added	Gas	250
4	TEVATR (Bus118)	Added	Gas	250
5	CASTAI4G (Bus40)	Added	Gas	1000
5	CORONADO (Bus4)	Added	Gas	500
6	FCNGN4CC (Bus9)	Added	Coal	750
6	DALLES21 (Bus70)	Added	Gas	1000
7	MOHAV1CC (Bus148)	Added	Gas	800
7	JOHN DAY (Bus77)	Added	Hydro	800
8	ELDORADO (Bus138)	Added	Coal	800
9	NAVAJO 2 (Bus13)	Added	Gas	800
10	ELDORADO (Bus138)	Added	Wind	500
10	JOHN DAY (Bus77)	Added	Hydro	800
11	BRIDGER2 (Bus36)	Added	Coal	750
12	EMERY (Bus159)	Added	Coal	750
13	HAYDEN (Bus11)	Added	Coal	750
14	INTERM1G (Bus45)	Added	Wind	500
15	NAUGHTON (Bus162)	Added	Wind	500

Table 4-6: Summary of Future Years' Resource Assumptions

Difference with the current year "Existing" base case (MW)						
Region	Comment	Other (Nuclear)	Hydro	Coal	Wind	Gas
CAMX	Added				2500	3050
	Retired					
AZNMNV	Added			1550	500	2400
	Retired			1580		
NWPP	Added		1600			1000
	Retired					
RMPP	Added			2250	1000	
	Retired					
Total Net Change to the current Year		0	1600	2220	4000	6450
Total Additions Only		0	1600	3800	4000	6450

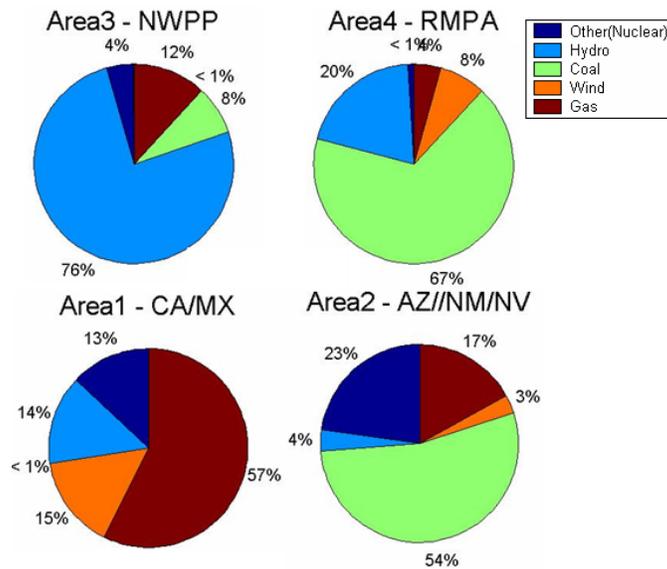


Figure 4-5: Generation Portfolios in the 15th Year

4.2.3 Long Term PCF Simulation Results

The generation economic dispatch order follows the sequence of variable energy cost: nuclear, hydro, coal, and gas. Nuclear and wind generations in all areas are treated as non-bids and must-take generations in this study. Coal-fired generation has the slow start characteristic so it is often operated in a base-load mode with a minimum operating capacity. Also, in the CA/MX area, it is assumed that some generating units are in a must-run status due to local reliability reasons.

Since the wind plants exhibit wide variations in output, to investigate the impact of wind power uncertainty on long term system congestion, two scenarios are illustrated. The first scenario is to assume that at the peak load of each year, all existing and future planned wind plants are running at their rated capacities. Another scenario is to replace the rated wind generation output model with the probabilistic model that is described in section 2.2.2. The contentious debate in the wind power and power grid industries is about the dependable wind capacity for capacity credit. What we attempt to model in these two approaches is to investigate the effect of the uncertainty of that dependable capacity. As this research project is mainly to develop models and not yet at the point of applying the models to perform definitive studies, the results from this research project are not to be taken for drawing any conclusions for that debate.

- Scenario 1: Constant wind generation output at rated capacity

In Figure 4-6, the results of simulating the load and generation uncertainties are shown.

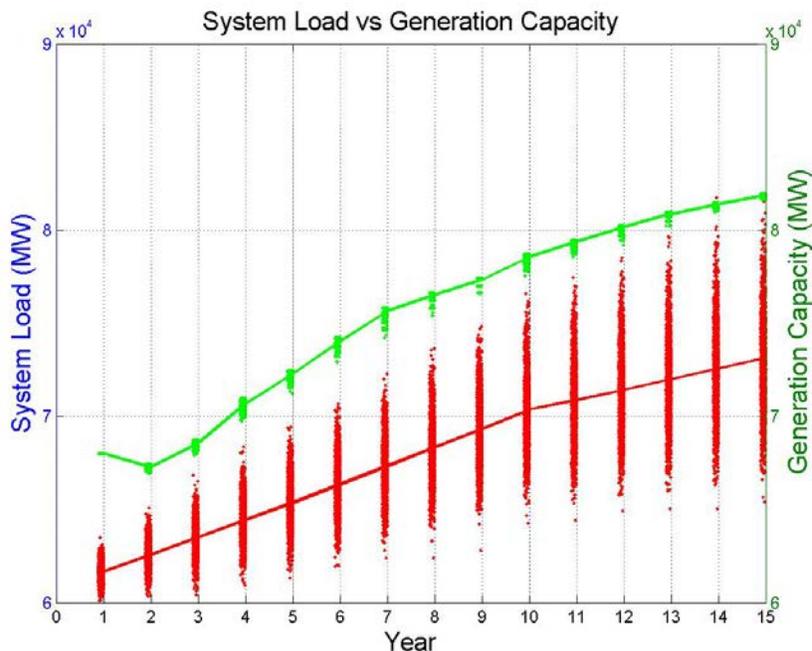


Figure 4-6: System Load vs. Generation Capacity (Scenario 1) for the CA/MX Area (Top curve is generation, bottom curve with wide ranges is load)

The generation capacity curve is the green curve at the top of the graph. It varies within some range of uncertainty due to generation in-service and retirement uncertainties. The lower red curves represent the uncertain ranges of the system load due to forecast uncertainties. They show much greater uncertainty than the generation availability curve. From this graph, we can observe that for the CA/MX area the planning reserve margin is around 10 percent. The least margin (8 percent) is in the 2nd year because a 1,580 MW capacity coal plant retires in the Mohave Desert area. At the end of the 15th year, the planning reserve margin is 11.96 percent and the mean value of total generation capacity is 81,887 MW.

The PCF results on the COI path are shown in Figure 4-7a and Figure 4-7b. The red points are the congestion forecasting results of sequential Monte Carlo trials, which are expressed as percentages of congestion path capacity limits. The blue curve is the mean value of congestion forecasting results of all Monte Carlo trials. The two dashed black curves are the lower and upper prediction bounds with 95 percent confidence level. This interval indicates that we have a 95 percent chance that the congestion is actually contained within the lower and upper prediction bounds and we take a 5 percent chance of being incorrect about forecasting congestion. The green curve is the probability of congestion plotted on a logarithmic scale, to be read on the vertical scale to the right of the graph. It is a numerical translation of the likelihood of the red dots of power flows exceeding the solid blue line of the transmission path capacity limit. So when you see the range of the red dots spreading over and above the solid blue line, the green curve will rise to provide a very useful index of the probability of transmission congestion.

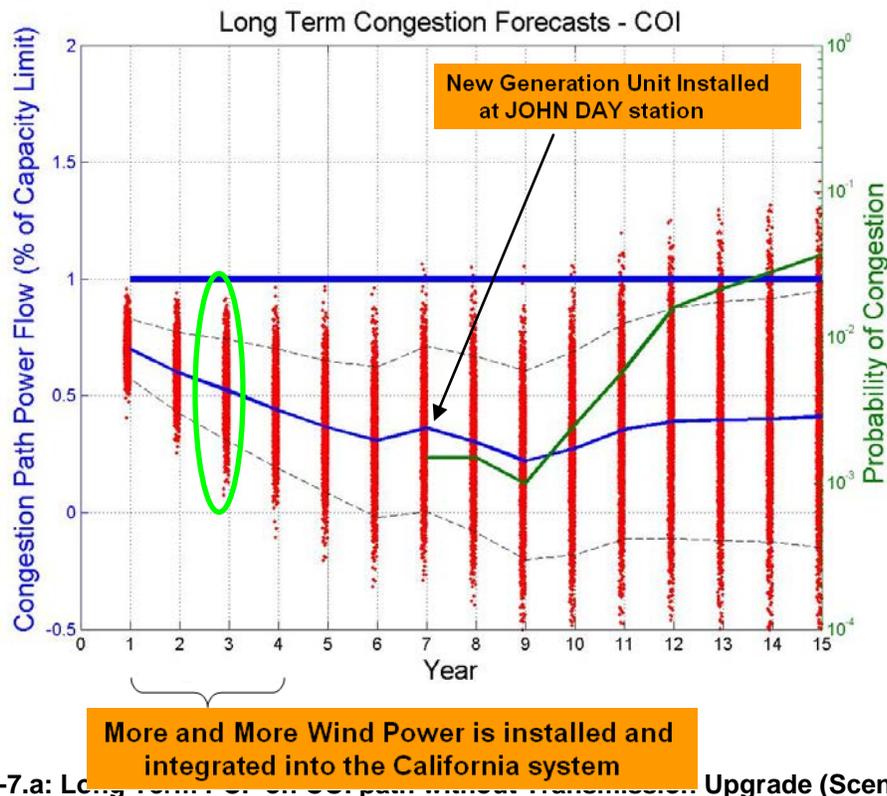


Figure 4-7.a: Long Term Congestion Forecasts - COI Path (Scenario 1) Upgrade (Scenario 1)

In Figure 4-7.a, it shows that the mean value of congestion path flow on the COI path goes down in the first four years. In the beginning year, the CA/MX area has to import a large amount of hydro generation from the NWPP area to serve the local load demand, which results in the heavy loading on the COI path. Then, when 2,500 MW wind plants are built and interconnected into the CA/MX area in the following four years, the CA/MX area can import less hydro generation from the NWPP area and use these additional must-take wind generations to serve the local load demand. This situation will help to relieve the heavy loading on the COI path. On the other hand, we can see that the mean value of congestion path flow on the COI path takes a leap in the 7th year. This is because a new hydro generation unit is built at John Day station in the NWPP area. Its economic utilization pushes more hydro generation into the load demand area in CA/MX and causes heavier loading on the COI path.

In Figure 4-7.b, the green curve which is the probability of congestion disappears after the 10th year, compared with Figure 4-7.a, because there is a hypothetical transmission upgrade in the 10th year. The transmission capacity on the COI path is increased to 130 percent of its existing capacity, which helps relieve the congestion on the COI path. The probability of congestion emerges again after the 14th year.

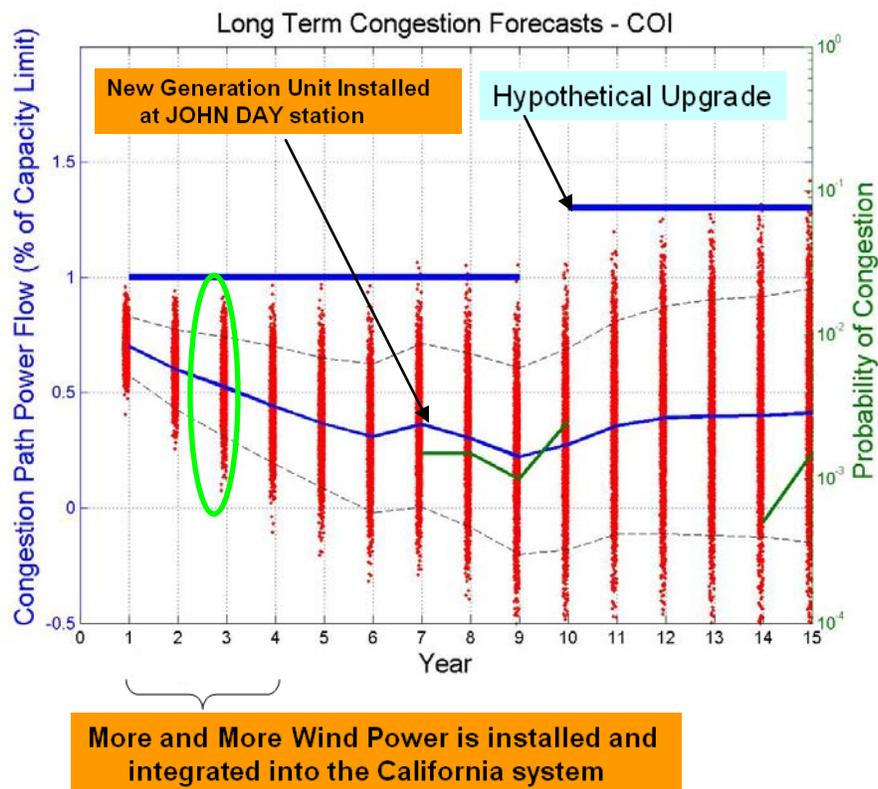


Figure 4-7.b: Long Term PCF on COI path with Hypothetical Transmission Upgrade (Scenario 1)

- Scenario 2: Probabilistic wind generation output

The simulation assumptions are the same as in Scenario 1 except for the wind generation model. All existing and future wind generations are represented by the probabilistic model that is described in section 2.2.2. Figure 4-8 shows that at the end of the 15th year, the planning reserve margin is 9.8 percent and the mean value of total generation capacity is 80,289 MW. This is because the wide fluctuation of wind output and the probabilistic dependable capacity affect the wind model in long term planning. Using the probabilistic model of wind generation in the study, the planning reserve margin in scenario 2 will be less than the margin in scenario 1. It tells us that the traditional deterministic planning analysis gives an overly optimistic assessment of the reliability contribution of wind plants.

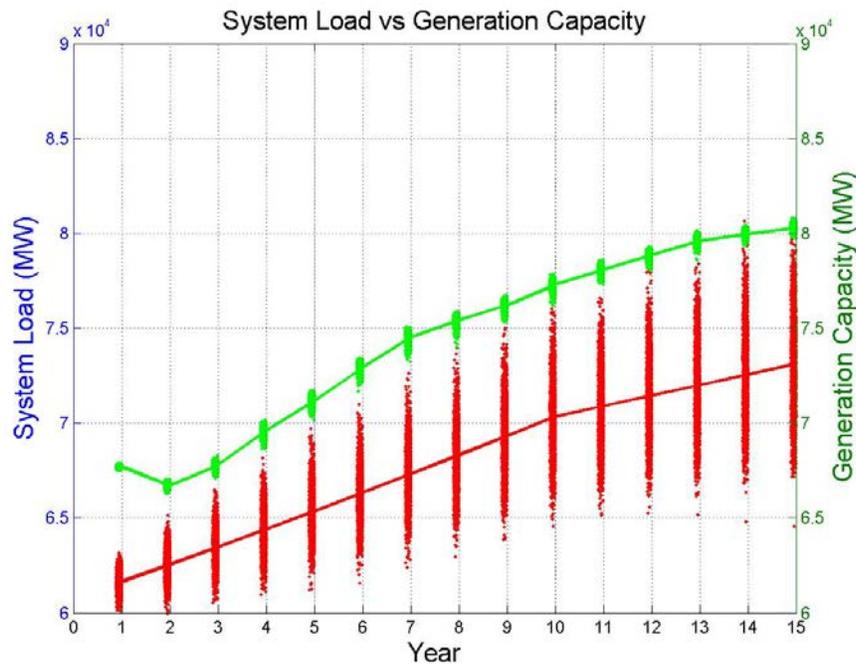


Figure 4-8: System Load vs. Generation Capacity (Scenario 2)

In Figure 4-8.a, it shows that the mean value of congestion path flow on the COI path goes down in the first four years. The factors affecting Scenario 1 similarly affect Scenario 2. In the beginning year, the CA/MX area has to import a large amount of hydro generation from the NWPP area to serve the local load demand, which results in the heavy loading on the COI path. Then, when 2,500 MW wind plants are built and interconnected into the CA/MX area in the following four years, the CA/MX area can import less hydro generation from the NWPP area and use these additional must-take wind generations to serve the local load demand. This situation will help to relieve the heavy loading on the COI path. On the other hand, we can see that the mean value of congestion path flow on the COI path takes a leap in the 7th year. This is because a new hydro generation unit is built at John Day station in the NWPP area. Its

economic utilization pushes more hydro generation into the load demand area in CA/MX and causes heavier loading on the COI path.

In Figure 4-8.b, the green curve which is the probability of congestion disappears after the 10th year, compared with Figure 4-8.a., because there is a hypothetical transmission upgrade in the 10th year. The transmission capacity on the COI path is increased to 130 percent of its existing capacity, which helps relieve the congestion on the COI path. The probability of congestion emerges again after the 13th year.

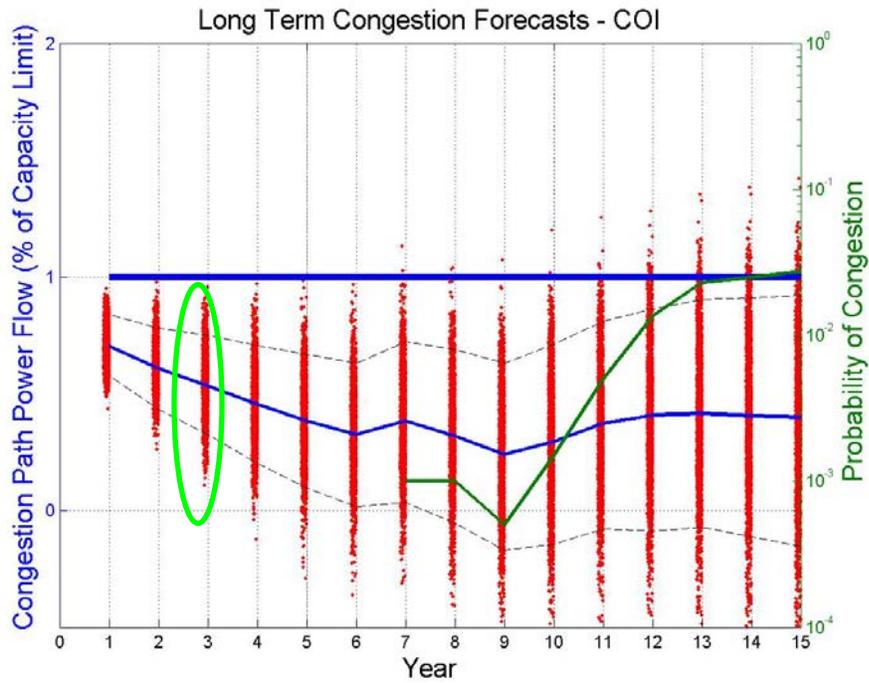


Figure 4-9.a: Long Term PCF on COI path without Transmission Upgrade (Scenario 2)

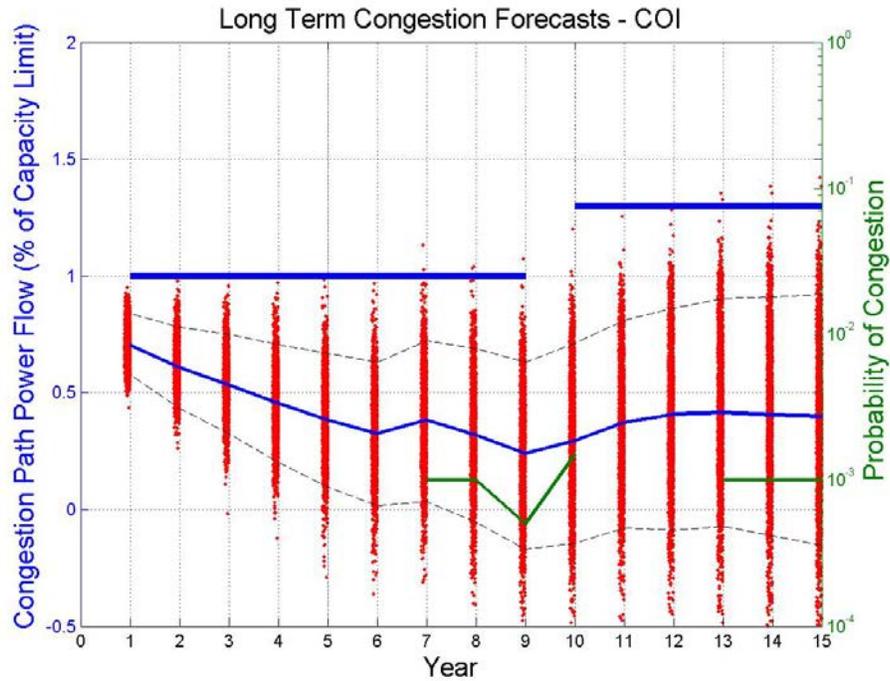


Figure 4-9.b: Long Term PCF on COI path with Hypothetical Transmission Upgrade (Scenario 2)

- Comparison between two Scenarios

Figure 4-9 shows a comparison between Scenario 1 and Scenario 2 on the PDF and Cumulative Density Function (CDF) of congestion path flow on the COI path in the 3rd Year.

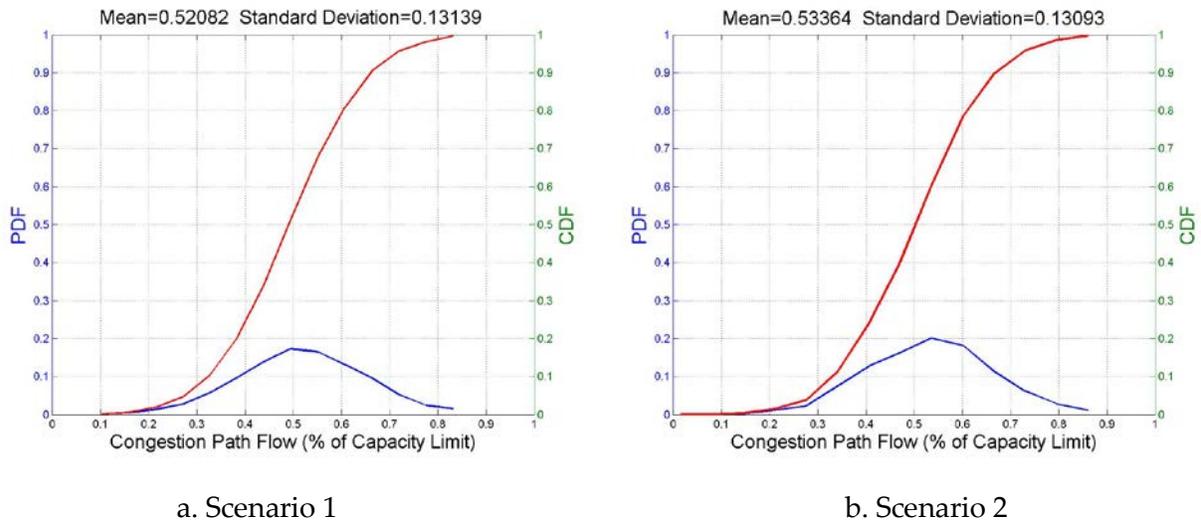


Figure 4-10: PDF and CDF of Congestion Path Flow on COI path in the 3rd Year

We can see that the mean value of congestion path flow of the Scenario 1 is smaller than the Scenario 2. In Scenario 2, the actual additional available wind generations in CA/MX area are less than 2,500 MW if we consider the uncertainty of wind and use the probabilistic model to represent it. This requires importing more hydro generation from the NWPP area, which results in the increased loading on the COI path compared with Scenario 1.

4.2.4 Key Caveats and a Sensitivity Analysis based on Long Term PCF

Some key caveats from the long term PCF simulation are:

- The choice of using a probabilistic or a deterministic wind generation model in long term planning will affect the forecasting results; ignoring the probabilistic nature of dependable wind power capacity is overly optimistic about the reliability impact of wind plants;
- Forecasted congestions are highly dependent upon the long term load forecasting;
- Forecasted congestions are also highly dependent upon the location and timing of future resources' installation and retirement;
- The forecasted congestions give significant information regarding incremental improvements and timing of future transmission upgrade requirements.

Figure 4-10 shows the relationship between the probability of congestion and the congestion path capacity. When the congestion path capacity is at 4,800 MW, the congestion probability is 2.9 percent. It is obvious that the congestion probability will decrease with the increase of the congestion path capacity/transmission investment.

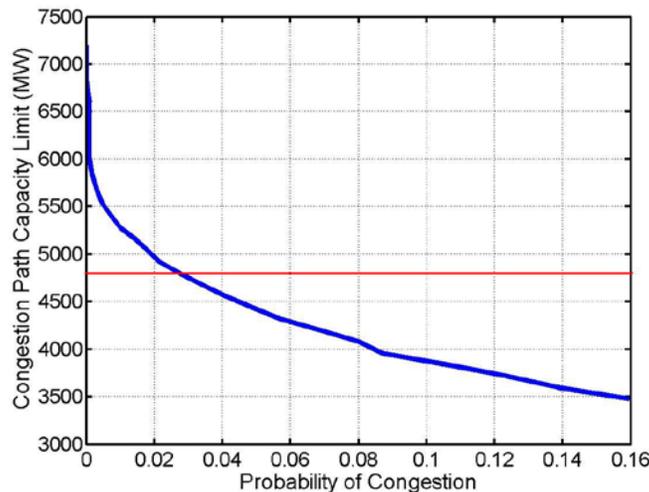


Figure 4-11: The Relationship between Congestion Probability and the Congestion Path Capacity (using the 15th year congestion forecast of Scenario 2)

Because more transmission investment will be needed to increase the transmission capacity, the vertical Y-axis can also represent a measure for the transmission investment. On the other side, the horizontal X-axis is also a measure that can represent the utilization of the transmission investment. Congestion equates to 100 percent instantaneous capacity utilization or greater. There is a tradeoff between investment in additional transmission capacity and the subsequent reduction in the utilization of the total transmission path. *Thus, a graph like the one shown in Figure 4-10, which can be produced by the long term PCF analysis, will be an extremely valuable output for facilitating the financial and risk analysis of future transmission investment for managing long term transmission congestion, through either a market mechanism or a regulatory approach. In short, the long term PCF analysis is potentially capable of providing decision-makers a way to consider the transmission congestion cost, the transmission investment requirements, and the balance between congestion risk with the financial and investment risks together.*

4.3 Short Term PCF Data Preparations and Simulation Results

In order to see the wind generation's impact on a power system in the short term time frame, the short term transmission PCF study is performed in the 15th year system condition in which the wind generation penetration in CA/MX area is 15 percent of the total generation capacity in this area.

4.3.1 Load Demand Data Preparation

To investigate the daily load variation, four season load curves are derived from the 2006 CAISO system load data which are available on-line: <http://oasis.caiso.com/>. Figure 4-11 shows the typical four season load curves in the CA/MX area and the simulated available generation capacity. The model in equation (3-1) with time dependent mean value and variance is applied to the summer load demands. In Figure 4-11, the Summer Load curve is listed as an example that follows the normal distribution with time dependent mean values and standard deviations. The red curve is the mean value. Each red point is the aggregated system load of a Monte Carlo sample. In Figure 4-11, the load curves for the other three seasons are only plotted as the mean values.

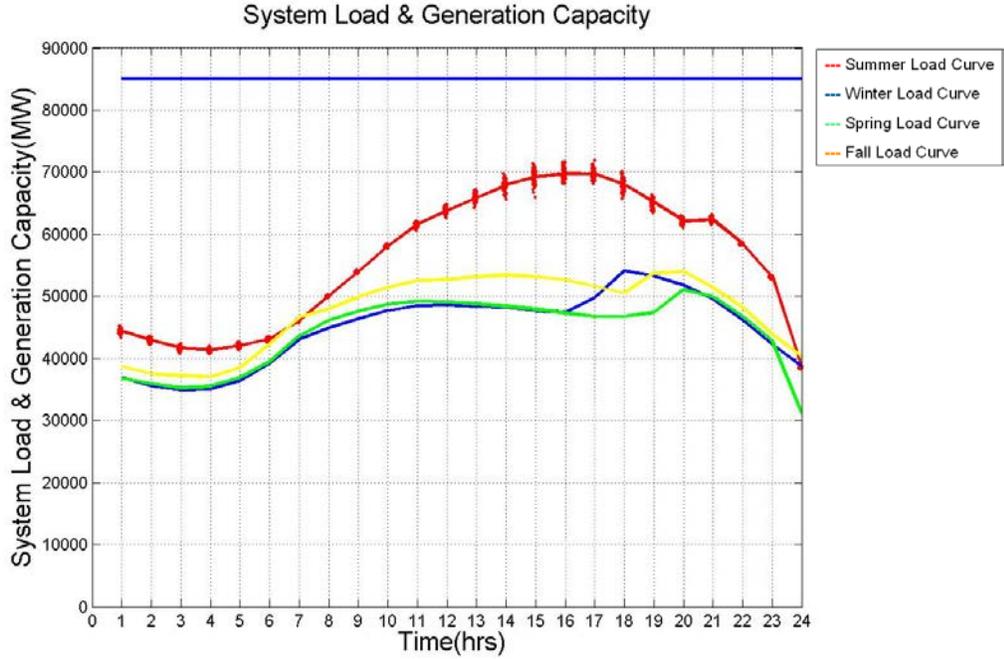


Figure 4-12: System Load vs. fixed Generation Capacity

4.3.2 Generation Resources Data Preparation

Figure 4-12 shows the generation portfolios of these four areas in the short term PCF study. The wind power penetration in the CA/MX area is 15 percent of the total generation capacity in this area.

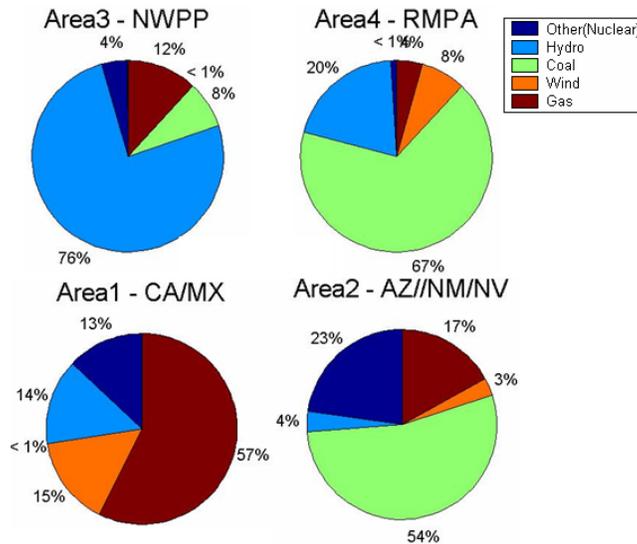


Figure 4-13: Generation Portfolios in Short Term PCF study

In general, if an uncertainty is truly random in nature and short term forecasting cannot significantly reduce its uncertainty, the probabilistic model is a fixed mathematical model even if the model parameters are time-varying. This applies to load models and generator outage models. In the case of wind plants output, this assumption is not correct. Experience from CAISO proves that wind uncertainty can be significantly reduced by having short term wind forecast down to the seconds and the minutes.

With this special nature of wind forecast uncertainty, the role of the short term PCF should be carefully re-examined for its applications. If the short term PCF is to be used as a policy analysis tool, or as an operational planning tool (applied in the day ahead or week ahead time frame), then the effect of the immediate real-time wind forecast has no relevance because the period of interest is not the very next hour, but rather some future realizations of a 24-hour day. However, the short term PCF may be also applied to the near real-time operation arena to help the grid operator to predict the transmission congestion for the rest of the day. This is the objective of another EPRI research project for the PIER TRP under the name of Critical Operating Constraints Forecasting [Commission Contract No. 500-02-004, Commission Work Authorization No: MR-050]. It is in this other context that we tested a new probabilistic model for wind power generation on a forward-looking basis as described below. This model recognizes that the band of uncertainty of the wind forecast starts from a very narrow band for the immediate next hour, and widening over time until it completes a whole 24-hour period.

To model the forward-looking wind generation uncertainty in this sample study, the ARMA (1, 24)¹⁸ model in equation (3-2) is used in this study where 1 is the order of the autoregressive process on itself and 24 is the order of the moving-average error term. It means that day-ahead (DA) wind forecasting will consider the last hour's forecasting result and the past 24 hours' forecasting errors. It is assumed that the simulation begins at 8:00 a.m. The first hour 8:00AM's forecasted error in the first hour of 8:00 a.m. is approximately 7 percent, increasing to approximately 24 percent at 7:00 a.m. of next day. The average forecasted error is 15 percent of the total wind generation capacity.

If the short term PCF is to be used for the policy or planning oriented operational planning studies, the model for the wind generation will not have the same widening forecasted errors over time as used in the forward-looking model. Rather, it will utilize statistics on the wind plant outputs over a 24-hour day, based on the measured wind plant outputs, instead of the forecasted wind plant outputs. Of course, the results of the short term PCF would be different, dependable on the context and the input assumptions for the wind model.

4.3.3 Short Term PCF Simulation Results

To investigate the impact of wind power uncertainty on short term forward-looking system congestion, two scenarios are illustrated.

¹⁸ According to the reference (Milligan et al. 2003), the best overall model specification is the ARMA (1, 24) model.

- Scenario 1: Add typical variation of wind generation that is described in section 3.2.2. The maximum mean value of total wind power outputs in the CA/MX area is 3370 MW at 1:00 a.m.
- Scenario 2: Replace all wind generation units in the CA/MX area with assumed 3370 MW constant must-take generation.

Figure 4-12 and Figure 4-13 show the CA/MX area system load and generation capacity. Compared with Scenario 2, the CA/MX area has significant energy deficiency due to the loss of major wind generation during day time in Scenario 1. When must-run generation outputs are larger than loads, export energy is equal to must-run generation outputs minus load demands. The blue points in both Figures show the export energy from the CA/MX area because the wind plants reach their maximum at mid-night.

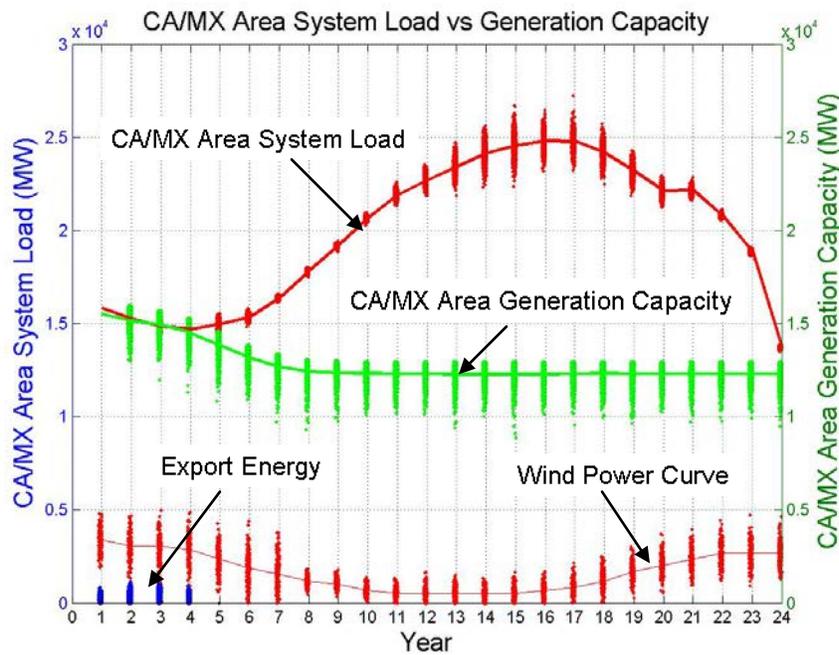


Figure 4-14: CA/MX Area System Load vs. Generation Capacity with Typical Variation of Wind Power

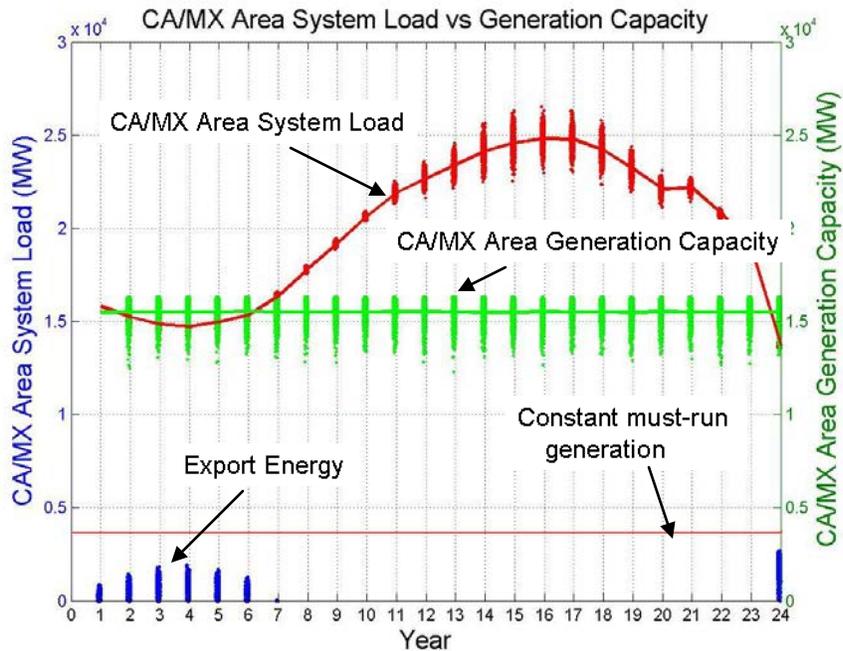


Figure 4-15: CA/MX Area System Load vs. Generation Capacity with 3370 MW Must-take Generation

COI Path - The PCF results on the COI path of these two scenarios are shown in Figure 4-14 and Figure 4-15, respectively. Just like the simulation results in long term PCF, the red points are the congestion forecasting results of sequential Monte Carlo trials, which are expressed as percentages of congestion path capacity limits. The blue curves are the mean values of congestion forecasting results of all Monte Carlo trials. The two dashed black curves are the lower and upper prediction bounds with 95 percent confidence level.

As shown in Figure 4-14, there are heavy congestions happening in scenario 1 during the day time. The reason is that the CA/MX area has to handle a large wind power loss in the day time due to the typical decline of wind power. Therefore, system operators must dispatch additional generation or purchases to balance the system. In this study, according to the economic dispatch rule, the NWPP area delivers to the CA/MX area additional hydro generation, which aggravates the congestion on COI. Compared with scenario 1, there is no congestion happening in scenario 2 and the green curve disappears in Figure 4-15 because the CA/MX area has a constant 3370 MW of must-take generation (modeling the alternative of having a firm 3370 MW of power instead of the uncertain wind power rated at the maximum 3370 MW capacity). Another observation is that the congestion window in Figure 4-14 is broader than Figure 4-15. This is because the wind forecasts error is considered only in scenario 1 of this study. The wind uncertainty creates not only a problem in the morning drop-off period but also into the daytime high load period.

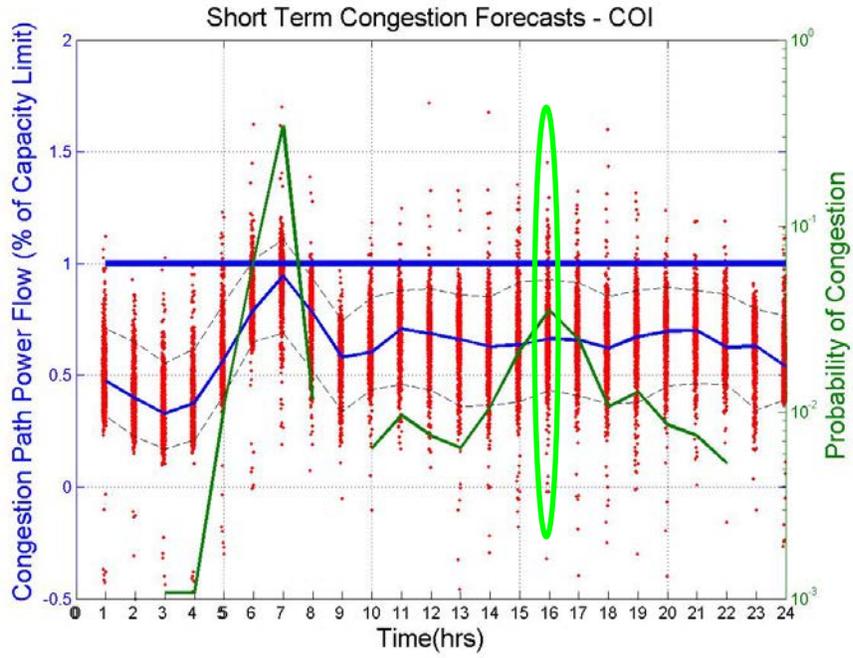


Figure 4-16: Short Term PCF on COI Path (Scenario 1: Generation Capacity with Typical Variation of Wind Generation)

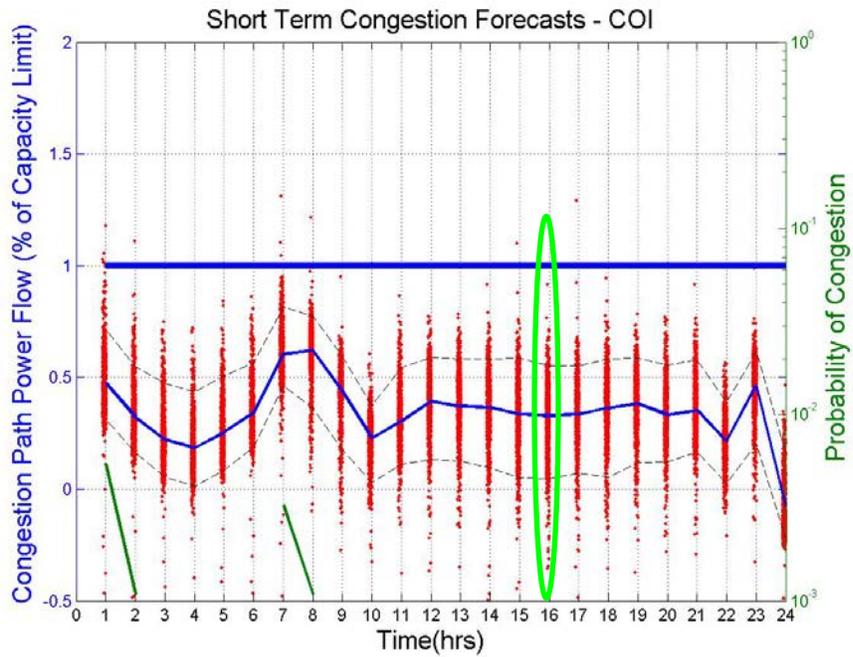


Figure 4-17: Short Term PCF on COI Path (Scenario 2: Generation Capacity with Assumed 3370 MW Must-Take Generation)

WOR Path - The PCF results on the West of River (WOR) path of these two scenarios are shown in Figure 4-16 and Figure 4-17, respectively. Similar to the observations from Figure 4-14 and Figure 4-15, heavier congestions happen in Figure 4-16 during the day time because the CA/MX area imports additional coal generation from the Arizona area to handle the large loss of wind power in the day time. Another observation is that there is heavier congestion happening in the day time in both Figure 4-16 and Figure 4-17. When the total system load climbs, the system energy cost will increase according to the economic dispatch rule. When the system energy cost is greater than the energy cost of coal plants in Arizona area in the day time heavy load condition, these coal plants will be dispatched to their maximum output and start to be imported in greater amount to serve the California load. This results in more coal generation transferring from Arizona to California which then aggravates the congestion in the WOR path.

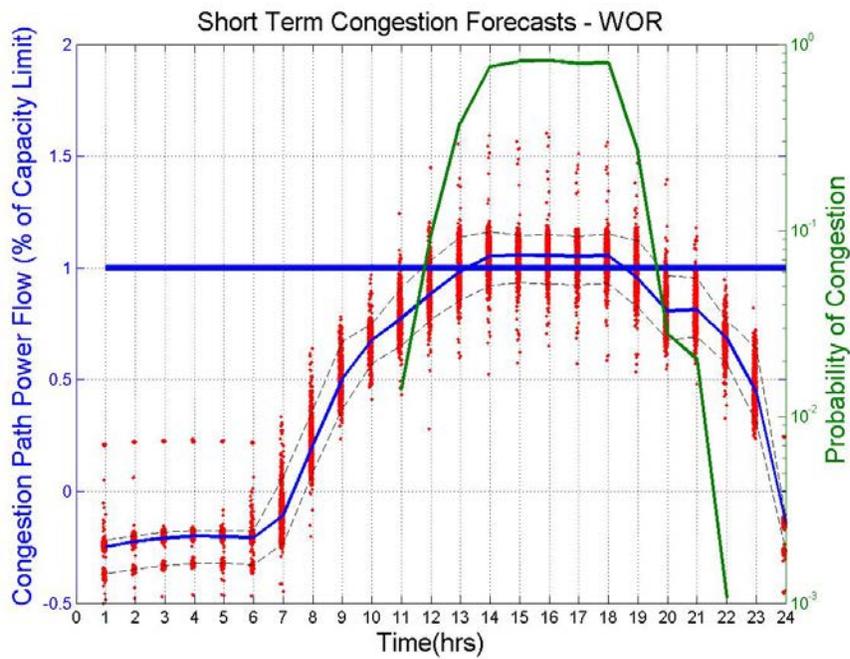


Figure 4-18: Short Term PCF on WOR Path (Scenario 1: Generation Capacity with Typical Variation of Wind Generation)

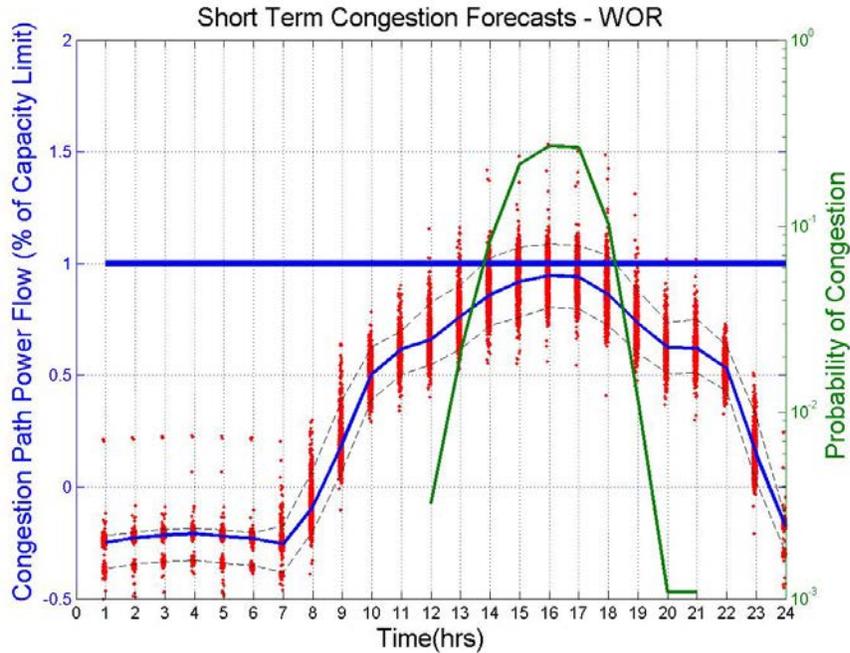


Figure 4-19: Short Term PCF on WOR Path (Scenario 2: Generation Capacity with Assumed 3370 MW Must-Take Generation)

- Comparison between two Scenarios

Figure 4-18 shows a comparison between Scenario 1 and Scenario 2 on the PDF and CDF of congestion path flow on the COI path at 4:00 p.m. The mean value of congestion path flow of Scenario 1 is larger than that in Scenario 2. The reason is that the CA/MX area has to handle a large wind power loss at 4:00 p.m. due to the typical variation of wind power as shown in Figure 4-12. Therefore, system operators must dispatch additional generation or external purchases to balance the system. In this study, the CA/MX area imports additional hydro generation from NWPP according to the economic dispatch rule, which aggravates the congestion on COI.

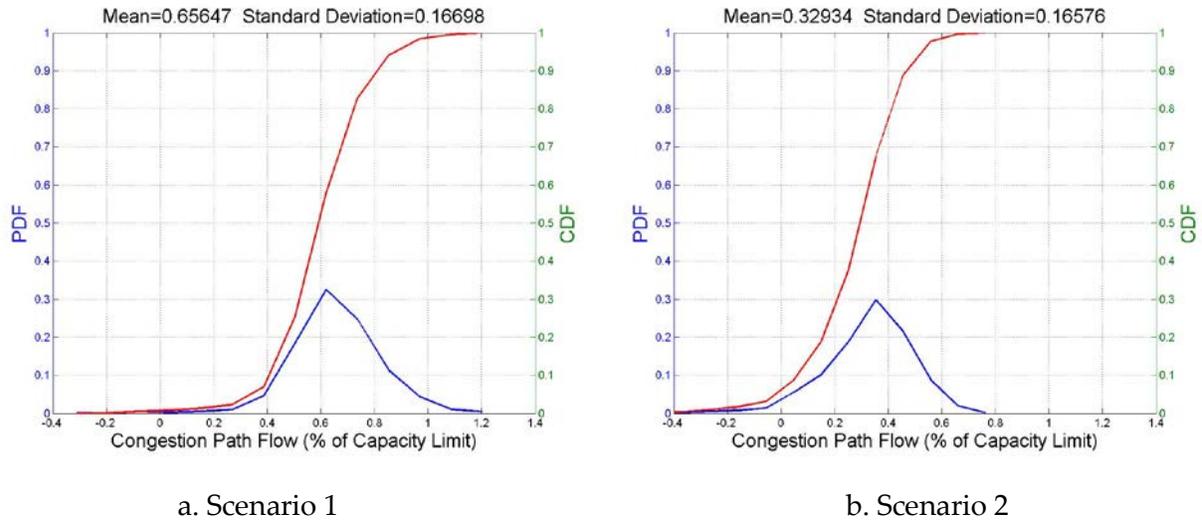


Figure 4-20: PDF and CDF of congestion Path Flow on COI path at 4:00 p.m.

4.3.4 Key Caveats and a Sensitivity Analysis based on Short Term PCF

Some key caveats from the short term PCF simulation are:

- Integration of wind power into system has a great influence on the forecasted congestions;
- Daily patterns of wind power which exhibit a high degree of variability and uncertainty will likely cause more serious congestions with greater uncertainty;
- More accurate minutes-ahead, hour-ahead and day-ahead wind forecasts will lead to less uncertainty of congestion;
- Forecasted congestion is also highly dependent upon load forecasting and the dispatch of generation and external purchases.
- The effect of transmission forced outages on short term PCF has not been observed to be significant. However, this aspect of research has not been studied with sufficient scope to draw any conclusion at this time.

A more detailed look at a particular hour (4:00 pm) in Scenario 1 was taken in order to gain some insight about the potential congestion cost. In a simulation when the COI path is observed to be constrained, the congestion probability is 4.3 percent. The congestion cost would be zero if no dispatch has been made to relieve the congestion. If we re-dispatch 1050 MW of gas-fired generation in the CA/MX area to relieve the congestion, the congestion probability will decrease to 3.2 percent. Further, the congestion probability will keep decreasing when more and more gas-fired generation in the CA/MX area is dispatched to serve the local load demand. Figure 4-19 shows the relationship between the probability of congestion and the congestion cost.

Congestion cost can be estimated by summing the values of least-cost transactions from cheap generation that cannot be completed due to transmission constraints, and comparing the sum to the more expensive value of the generation or imports forced by the constraint (DOE 2006).

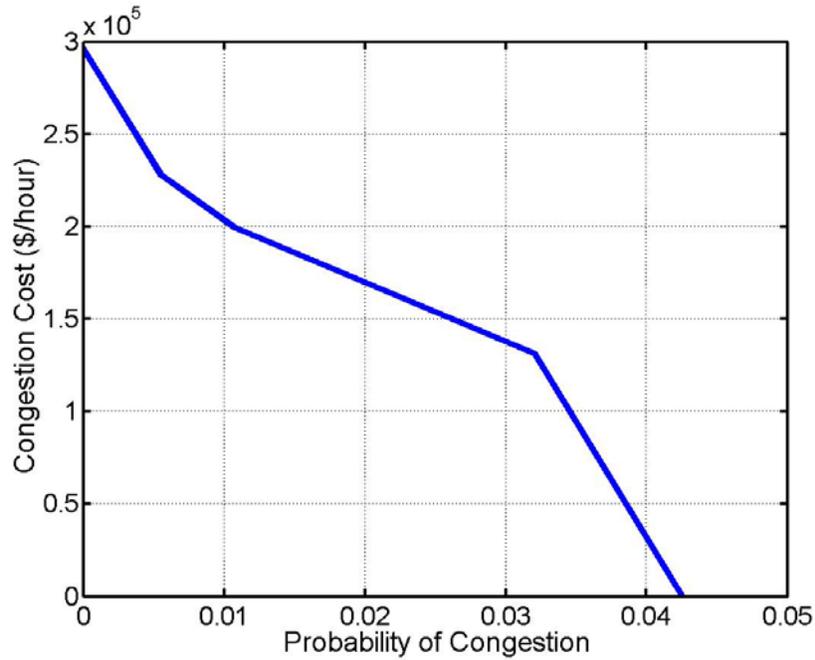


Figure 4-21: The Relationship between Congestion Probability and Congestion Cost on the COI path at 4:00 p.m. of Scenario 1

Recall that in the long term PCF discussions, we had Figure 4-10 showing the relationship between the probability of congestion and the congestion path capacity. It is repeated here as Figure 4-20 for comparison and discussion together. When the congestion path capacity is at 4,800 MW, the congestion probability is 2.9 percent. We pointed out there that the congestion probability will decrease with the increase of the congestion path capacity/transmission investment. With the analysis results from Figure 4-19, we can now relate the probability of congestion to the congestion cost.

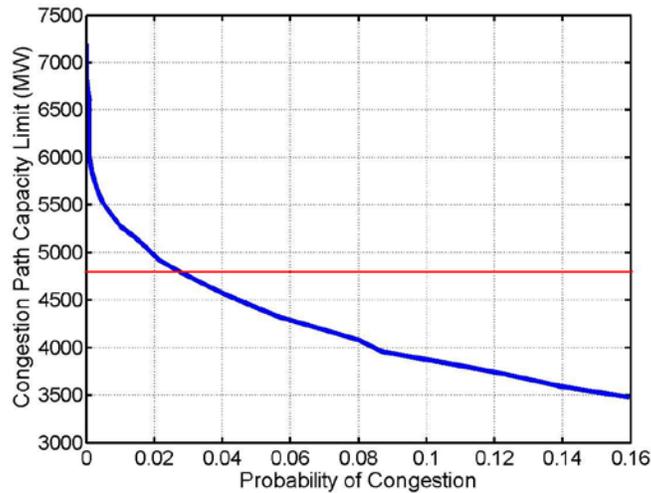


Figure 4-22: The Relationship between Congestion Probability and the Congestion Path Capacity from the long term PCF Results (same as Figure 4-10)

Because more transmission investment will be needed to increase the transmission capacity, the vertical Y-axis in Figure 4-20 can also represent a measure for the transmission investment required to reduce the probability of congestion to a given level, e.g., as shown by the red line in Figure 4-20.. On the other side, the horizontal X-axis is also a measure that can represent the utilization of the transmission investment, as pointed out in the long term PCF discussion. In addition, we now show that the horizontal X-axis is also related to the congestion cost. It is now possible to put these pieces together to show a tradeoff between investment in additional transmission capacity, the reduction of the congestion probability, and the subsequent reduction in both congestion cost and the utilization of the total transmission path. This is shown in Figure 4-21.

By putting the two graphs together, top and bottom, we can first choose a congestion probability, e.g., 0.025 percent, and notice that a transmission investment is needed to bring the transmission path capacity to 4800 MW. When we take that 0.025 percent probability to the bottom graph, we observe that there is a congestion cost during a certain time interval (in this example, this is only for one week of short term PCF simulation). The numerical values in the graphs of Figure 4-21 should be ignored for the time being because the two studies are not integrated. However, the concept is here. It is possible to relate both the capital cost of transmission investment and the operation costs of congestion (including potential extension to include the cost of reliability) to the single measure of the congestion probability of a particular transmission path.

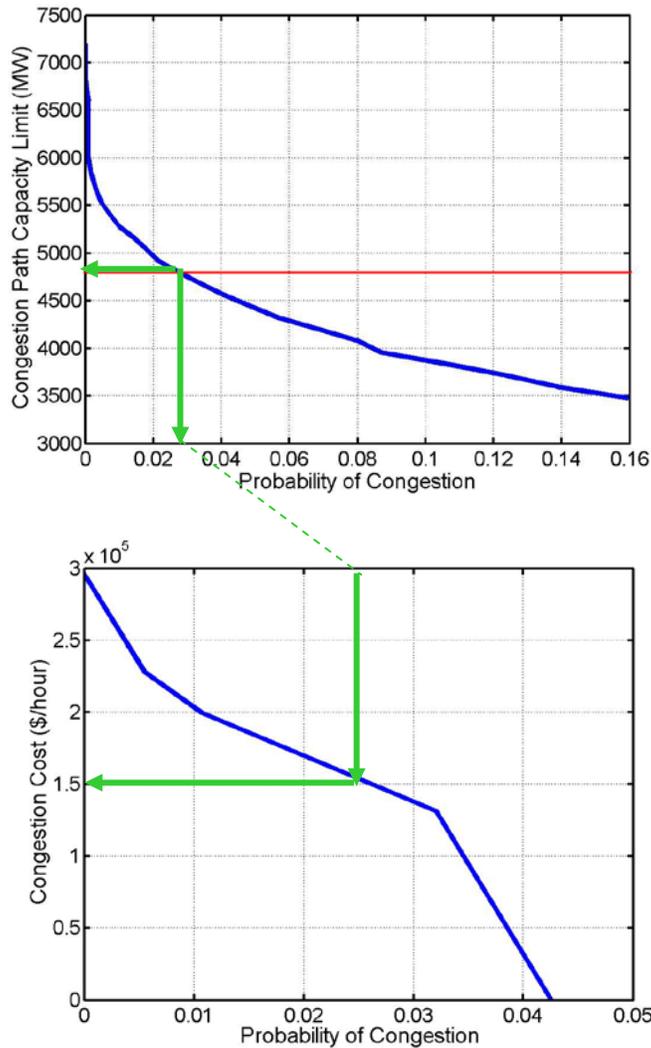


Figure 4-23: Putting the Long Term and the Short Term PCF Analysis Together for a Financial and Reliability Risk Tradeoff

Thus, a graph like the one shown in Figure 4-21, which can be produced by the long term PCF analysis integrated with the short term PCF analysis, will be an extremely valuable output for facilitating the financial and risk analysis of future transmission investment for managing long term transmission congestion, through either a market mechanism or a regulatory approach. In short, the combined long term PCF and the short term PCF analysis is potentially capable of providing decision-makers a way to consider the transmission congestion cost, the transmission investment requirements, and the balance between congestion risk with the financial and investment risks together.

CHAPTER 5: Conclusions and Future Research

This project has developed probabilistic models and most importantly specified how to accurately model the key input assumptions in order to derive valid confidence levels of the forecasted congestion variables. The method proposed by this project combines the use of analytical functions with regression methods to provide accurate models of the uncertainties, including the effect of correlation. It uses a Monte Carlo simulation method to accurately model the probabilistic and physical relationships between generation dispatch, load demand, and the configuration of the transmission grid in order to mathematically predict the key operating constraints of line loading along critical transmission paths. It demonstrated the methodology using the equivalent model of the WECC system, with focus on the impact of such congestion on the California power grid and consumers. The mathematical models and the time frames of the simulation differ between the short term (24 hours) and the long term (10-20 years), and therefore two computer models were developed to address the two time frames. With these computer programs, each Monte Carlo simulation computes the power flow under one particular scenario about the uncertainties. Thus, thousands of Monte Carlo simulations are conducted in order to gain confidence about the variability of the forecasted results of transmission congestion.

Mathematical models of uncertainties in Long Term Probabilistic Congestion Forecasting (PCF) can be summarized as:

- Load demand model can be described as a chronological exponential expression. The load increase exponent follows normal distribution. The load demand model also considers the correlations among loads.
- The generation capacity is modeled with the uncertainties of generators' installation and retirement dates following the discrete exponential distribution.
- Wind generation's spatial correlation and statistical expectation at the system peak load condition are considered in the long term planning model. Wind generation outputs are assumed to be subject to a probability distribution.
- Economic dispatch is used to balance the generation and load.

Mathematical models of uncertainties in the short term PCF can be summarized as:

- Load demand is assumed to follow the normal distribution with time dependent mean value and standard deviation. The load demand model considers the correlation between loads.
- Forced Outage Rate (FOR) is used to express the outage rate of a generation unit to model generation uncertainties. Each generation unit's output follows the binomial distribution.

- Autoregressive integrated Moving Average (ARMA) method is used to provide an accurate model of wind uncertainties in short term planning.
- FOR is also used to express the outage rate of a transmission line to model transmission uncertainties. Transmission line status follows the binomial distribution.
- Economic dispatch is used to balance the generation and load.

Some key caveats from the long term PCF simulation are:

- The choice of using a probabilistic or a deterministic wind generation model in long term planning will affect the forecasting results; ignoring the probabilistic nature of dependable wind power capacity is overly optimistic about the reliability impact of wind plants;
- Forecasted congestions are highly dependent upon the long term load forecasting;
- Forecasted congestions are also highly dependent upon the location and timing of future resources' installation and retirement;
- The forecasted congestions give significant information regarding incremental improvements and timing of future transmission upgrade requirements.

Some key caveats from the short term PCF simulation are:

- Integration of wind power into system has a great influence on the forecasted congestions;
- Daily patterns of wind power which exhibit a high degree of variability and uncertainty will likely cause more serious congestions with greater uncertainty;
- More accurate minutes-ahead, hour-ahead and day-ahead wind forecasts will lead to less uncertainty of congestion;
- Forecasted congestion is also highly dependent upon load forecasting and the dispatch of generation and external purchases.
- The effect of transmission forced outages on short term PCF has not been observed to be significant. However, this aspect of research has not been studied with sufficient scope to draw any conclusion at this time.

The project team would like to recommend that additional work be done to investigate what implications the probabilistic congestion forecasting results may have for energy policies aimed at encouraging reliable and economic electric power delivery in California. Future work should extend this study to look into the relationship between the confidence levels of forecasting congestion and the specific uncertainties, such as generator outages, higher penetration of renewable generation, demand response or management, generator siting, load growth, transmission forced outages and market prices, etc. The extended study results will provide information on how public policies may be improved so as to increase the reliability, economic effectiveness and environment contributions of the California electric power system that is in California's public interest. It fits into the strategic goal of the CEC to improve the reliability of

the electric power system in California and to assure that California's economy will be supported by a reliable and economic electric power system.

In addition, the potential application of this project's methodology to provide a useful framework to enable a linkage between transmission congestion forecasting with the financial community to manage congestion through a risk management approach is extremely exciting and promising. As shown in the following Figure 5-1, which can be produced by the long term PCF analysis integrated with the short term PCF analysis, will be an extremely valuable output for facilitating the financial and risk analysis of future transmission investment for managing long term transmission congestion, through either a market mechanism or a regulatory approach. In short, the combined long term PCF and the short term PCF analysis is potentially capable of providing decision-makers a way to consider the transmission congestion cost, the transmission investment requirements, and the balance between congestion risk with the financial and investment risks together.

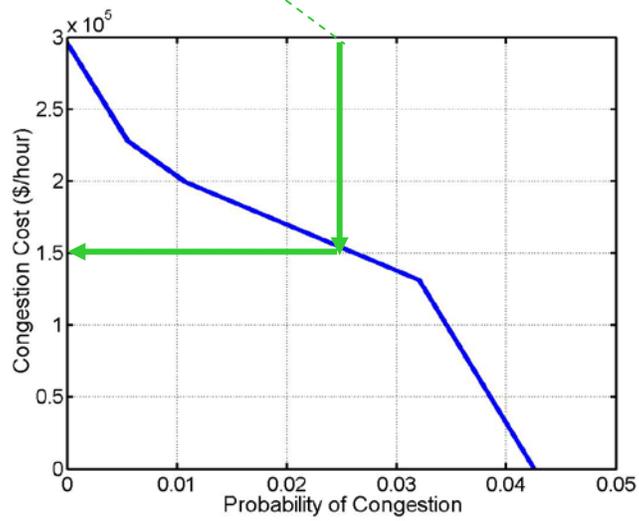
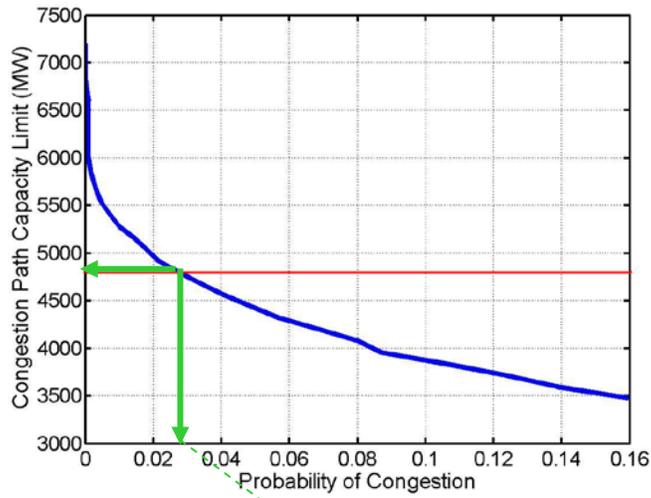


Figure 5-1: Putting the Long Term and the Short Term PCF Analysis Together for a Financial and Reliability Risk Tradeoff

REFERENCES

- Allan, R. N., A. M. L. Da Silva, and R. C. Burchett 1981. "Evaluation methods and accuracy in probabilistic load flow solutions," *IEEE Trans. on Power Apparatus and Systems*, Vol. PAS-100, No. 5, May, pp. 2539-2546.
- Bastian, J., J. Zhu, V. Banunarayanan and R. Mukerji 1999. "Forecasting energy prices in a competitive market," *IEEE Computer Applications in Power*, vol. 12, pp. 40-45, July.
- Borkowska, B. 1973. "Probabilistic load flow," *IEEE Trans. on Power Apparatus and Systems*, Vol. PAS-93, No. 3, May/June, pp. 752- 759.
- CAISO 2007. *Annual Report Market Issues and Performance, Annual Report, Department of Market Monitoring*, CAISO, April. [Online]. Available: <http://www.caiso.com/1bb7/1bb776216f9b0.pdf>.
- Contreras, J., R. Espínola, F. J. Nogales, and A. J. Conejo 2003. "ARIMA models to predict next-day electricity prices," *IEEE Trans. Power Systems*, vol. 18, August, pp. 1014–1020.
- DOE 2006. *National Electric Transmission Congestion Study*, Annual Report, U.S. Department of Energy, August. [Online]. Available: http://www.oe.energy.gov/DocumentsandMedia/Congestion_Study_2006-9MB.pdf.
- EPRI 2001. *DYNATRAN: Dynamic analysis of interconnected systems with energy storage*, Technical Report-1006531, EPRI.
- EPRI 2006. *California Regional Wind Energy Forecasting System Development, Volume 1: Executive Summary*, Technical Report-1013262, EPRI. [Online]. Available: <http://www.energy.ca.gov/2006publications/CEC-500-2006-089/CEC-500-2006-089-V1.PDF>.
- Hamoud, G. and L. Bradley 2001. "Assessment of transmission congestion cost and locational marginal pricing in a competitive electricity market," *2001 IEEE, Power Engineering Society Summer Meeting*, pp. 1468 -1472.
- Hippert, H. S., C. E. Pedreira, and R. C. Souza 2001. "Neural networks for short-term load forecasting: A review and evaluation," *IEEE Trans. Power Systems*, vol. 16, February, pp. 44–55.
- Li, W. and R. Billinton 1999. "Effect of bus load uncertainty and correlation in composite system adequacy evaluation," *IEEE Trans. on Power Systems*, Vol. 6, No. 4, November, pp. 1522-1529.
- Liete da Silva, A. M. and V. L. Arienti 1990. "Probabilistic load flow by a multilinear simulation algorithm," *IEE Proc. C: Generation, Transmission, Distribution*, Vol. 137, No. 4, July, pp. 276-282.

- Meliopoulos, A. P. Sakis, G. J. Cokkinides, and X. Y. Chao 1990. "A new probabilistic power flow analysis method," *IEEE Trans. on Power Systems*, Vol. 5, No. 1, February, pp. 182-190.
- Milligan, M., M. Schwartz and Y.H. Wan 2003. "Statistical Wind Power Forecasting for U.S. Wind Farms," National Renewable Energy Laboratory conference paper: NREL/CP-500035087, November.
- Min, L., S.T. Lee, P. Zhang, V. Rose, and J. Cole 2008. "Short-Term Probabilistic Transmission Congestion Forecasting", *IEEE Conference on Deregulation, Restructuring and Power Technologies*, Nanjing, China, April 6-9.
- Nogales, J., J. Contreras, A. J. Conejo, and R. Espínola 2002. "Forecasting next-day electricity prices by time series models," *IEEE Trans. Power Systems*, vol. 17, May, pp. 342-348.
- Szkuta, B. R., L. A. Sanabria, and T. S. Dillon 1999. "Electricity price short-term forecasting using artificial neural networks," *IEEE Trans. Power Systems*, vol. 14, August, pp. 851-857.
- Zhang, P. and S. T. Lee 2004. "Probabilistic load flow computation using the method of combined Cumulants and Gram-Charlier expansion," *IEEE Trans. on Power Systems*, Vol. 19, No. 1, February, pp. 676-682.

Appendix A: Literature Survey

Literature Review of Probabilistic Transmission Congestion Forecasting Approaches

Abstract

This paper reviews techniques for probabilistic congestion forecasting. The paper focuses on the long-term generation and transmission planning techniques. It also discusses some related topics, including the energy price forecasting, probabilistic production costing, probabilistic load flow, and security-constrained optimal power flow. In addition, discussion about the available commercial programs for congestion identification is included.

Index Terms --- Probabilistic congestion forecasting, generation and transmission planning, energy price forecasting, probabilistic production costing, probabilistic load flow, security-constrained optimal power flow.

I. Introduction

The U.S. Department of Energy issued the national congestion study in August 2006 (DOE 2006). The simulations of Eastern and Western interconnection congestion used 2008 as the base year to estimate congestion for a later year. This study is in the context of midterm operation planning, which is exposed to financial risks associated with forward electricity and fuel prices. For a long-term time frame, there are many uncertainties associated with the loads, generation, network topology and energy prices. Therefore, Probabilistic Congestion Forecasting (PCF) is essential for long-term power system planning. Long-term PCF, generally, is to apply a mathematical approach using probabilities to more correctly describe load and generation's time dependence and to include uncertainties in order to predict or forecast the ranges of transmission congestion. Results of PCF will help to find an optimal solution to the expansion planning problem of a power system with multiple transmission-constrained zones wherein existing and new power plants would be located. Review of the presented approaches and discussion of their advantages and drawbacks help to identify existing or potential new approaches and criteria for PCF. Unfortunately, there are no papers or reports directly addressing this topic. However, many papers and reports about new methods have been published in the technical literature due mostly to the improvement of the long-term generation and transmission planning introduced by the electric power sector deregulation. And also, this paper will review some interesting and related topics including:

- Energy price forecasting,
- Probabilistic production costing in generation planning,
- Probabilistic load flow (PLF) in transmission planning,

- Security-constrained optimal power flow (SCOPF).

This paper is organized as follows. Section II deals with publications that propose different synthesis algorithms to solve the generation resource planning problem. Section III reviews the methods related to transmission system planning. Section IV presents some interesting topics related to PCF. Some of the features of the available software for congestion identification are discussed in section V. Final remarks about the literature review close the paper.

II. Long-term Generation Resource Planning

Electric utility restructuring provides challenging options for generation resource planning in competitive electricity markets. Generation companies (GENCOs) can invest independently in the generation expansion. In the meantime, transmission system congestion is aggravated as GENCOs and other market participants make their investment decisions based on profit maximization. Indeed, the conflicts between economics and reliability are inevitable in the restructured electricity planning.

In electric utility monopolies, the least-cost operation within a pre-specified level of system security was the principle for applying conventional integrated resource planning algorithms (Bjorkvoll et al. 2001). However, in electricity markets, GENCOs' objective for generation resource planning is to maximize expected payoffs over planning horizons, while a secure operation of competitive power systems is sought by the Independent System Operator (ISO) through the coordination among market participants. Accordingly, conventional optimization techniques including linear programming (Khokhar 1997), dynamic programming (Neelakanta et al. 1999; Jia et al. 2001), decomposition method (Bloom 1982; McCusker et al. 2002) were successfully applied to multi-objective planning problems.

The emerging techniques applied to Generation Expansion Planning (GEP) were reviewed in (Zhu and Chow 1997). The Genetic Algorithm (GA) (Park et al. 2000; Park et al. 2002) and fuzzy set theory (Hu 2002) were used to solve the GEP problem. The Evolutionary Programming (EP) technique with Gaussian mutation and quadratic approximation technique was applied to solve the GEP problem (Park et al. 1999) as well. A simple, population based, Differential Evolution (DE) technique is used to solve the combinatorial optimization problems (Storn and Price 1997). Game-theoretic models (Chuang et al. 2001) were applied to generation resource planning with perceived difficulties for obtaining optimal solutions in competitive electricity markets.

Due to the volatile nature of electricity markets, uncertainties in operating conditions representing forecasted market price of electricity, load growth, financial risks, equipment availability, and the like are taken into consideration in generation resource planning algorithms. Facing uncertainties, power system planners apply certain techniques such as stochastic optimization (Gorenstin et al 1993; Melo et al 1991) to manage the probabilistic nature of generation planning. These proposed solutions are viewed as a tradeoff between economics and secure planning alternatives, which could provide planners with a list of alternatives based on expected payoffs and system security in electricity markets.

III. Long-term Transmission System Planning

Traditionally, the integrated planning of generation and transmission systems has been the responsibility of vertically-integrated utilities under state regulatory oversights (EPRI 1987; EPRI 1988; EPRI 1989). The common objective of integrated planning models is to find least-cost options for generation and transmission investment, system operation, and curtailment penalty charges.

However, in a competitive electricity market, it is hard to reach an integrated planning. GENCOs, as independent and for-profit market entities, are freely and actively making their own plans for generation expansion, which could dramatically impact existing transmission flows and congestions. Customers can also select their own electric energy suppliers based on economics, power quality, and security. As a result, transmission system planning is facing challenges for managing its operation economics and security (David and Wen 2001; Wong et al. 1999; Tabors 2000; Baldick and Kahn 1993; Leeprechanon et al. 2001; Gil et al. 2002). The objective in market-based transmission planning is to maintain system security within an acceptable level while maximizing the social welfare (or minimizing the investment and operation costs). To realize the objective and determine when, where, and what type of new transmission facilities are to be installed, various models and algorithms are introduced (Binato et al. 2001; Romero and Monticelli 1994; Alguacil et al. 2003; Fang and Hill 2003; Bahiense et al. 2001) in which the system security remains the most important operation issue.

In response to the novel requirements that competitive power markets place upon transmission planning, the Transmission Expansion Assessment Methodology (TEAM) (CAISO 2004) for assessing the economic benefits of transmission reinforcements is proposed for California. The methodology has five key principles: consideration of multiple perspectives (consumers, generators, transmission operators, and society at large); full network representation using a linearized DC load flow; market-based pricing, accounting for strategic behavior by generators; modeling of uncertainty, including the value of transmission as insurance against extreme events; and recognition of how supply, demand-side, and transmission resources can substitute for each other.

Either under state regulatory or competitive electricity markets, transmission planning is posed as an optimization problem with an objective function, subject to a set of constraints (Latorre et al. 2003). Several methods have been proposed to obtain the optimum solution for the transmission expansion problem, mostly using classical optimization techniques like linear programming (Chanda and Bhattacharjee 1994; Kim et al. 1988; Villasana et al. 1985), dynamic programming (Dusonchet and El-Abiad 1973), nonlinear programming (Youssef and Hackam 1989), and mixed integer programming (Bahiense et al. 2001; Lee et al. 1974). Optimization techniques like Benders (Binato et al. 2001; Tsamasphyrpu et al. 1999) and hierarchical decomposition (Romero and Monticelli 1994) have been also used, as well as the combination of decomposition techniques with other approaches, solving the problem with a “branch and bound” algorithm (Pinto and Nunes 1990).

IV. Related Topics

A. Energy price forecasting

Electricity prices are significantly influenced by the structure of the electricity market as it evolves over time. In a truly competitive environment, marketers will offer new products; new participants will find it attractive to participate in the market, and thus new financial instruments will become available for risk management. In the presence of liquidity and price discovery, the arbitrage between various energy products and their derivatives will be eliminated over time and equilibrium prices will be established. Determination of the evolving market structure, as inefficient plants are placed on stand-by or shut down and new players enter the market, is essential for forecasting the long-term prices of various energy products.

The electricity price forecasting methods include price simulation methods (Hamoud and Bradley 2001; Bastian et al. 1999; Hong and Hsiao 2002) and stochastic methods such as artificial neural network (ANN) methods (Hippert et al. 2001; Wang and Ramsay 1998; Szkuta et al. 1999; Nicolaisen et al. 2000; Valenzuela and Mazumdar 2001; Nogales et al. 2002), time series methods (Park et al. 1999), etc.

Energy price forecasting by simulation methods mimics the actual dispatch with system operating requirements and constraints. However, these methods require detailed system operation data and power grid models. For market participants, it is difficult to gather all required information, because most market participants prefer to reserve critical information from their competitors. Without sufficient information, obtaining accurate energy price prediction becomes hard. Furthermore, simulation methods are complicated to implement and consume lots of computation time.

Stochastic estimation methods possess the advantage of being easy to be implemented and requiring less information intensity. Artificial Neural Network (ANN) is a type of commonly used estimation method because of its strong learning capability. ANN techniques that have been widely used for load forecasting are now used for price prediction (Hippert et al. 2001; Wang and Ramsay 1998; Szkuta et al. 1999; Nicolaisen et al. 2000). In particular, Wang and Ramsay (Wang and Ramsay 1998) have proposed a hybrid approach based on neural networks and fuzzy logic, with examples from the England-Wales market and daily mean errors around 10 percent. Also, Szkuta et al. (Szkuta et al. 1999) have proposed a three-layer ANN with back propagation, showing results from the Victorian electricity market, with daily mean errors around 15 percent. Fourier and Hartley Transform as “filter” to the price data inputs of an ANN are reported in (Nicolaisen et al. 2000). Stochastic models of prices, as in (Valenzuela and Mazumdar 2001), are also competing with traditional time series models in order to predict daily or average weekly prices (Nogales et al. 2002; Contreras et al. 2003).

B. Probabilistic production cost (PPC) model

Electric power system planners use production costing models to forecast the operation cost of their generating systems to serve customer loads over a period of time, ranging from one month to several years.

Due to uncertainties in load forecasting, and to random outages of generation equipment, the production cost has to be calculated on a probabilistic basis. PPC then is used to calculate the mean of the operation cost for each stage and the probability of supply shortfalls, taking into account load fluctuations and availability of generation equipment. In order to calculate these indices efficiently, many of the available methods use a load duration curve approach introduced by Baleriaux (Baleriaux et al. 1967) and Booth (Booth 1972), with later improvements by several authors (Stremel et al. 1980; Sidenblad and Lee 1981; Schenk et al. 1984; Wang 1988; Gross et al. 1988).

One major disadvantage of the load duration curve based models is the loss of chronological information, that is, it is not possible to represent those aspects of production costing which are time dependent by nature: unit commitment, unit ramp rates, operation of hydro plants and other energy-limited devices etc.

An alternative for overcoming these limitations is to use a chronological simulation algorithm (Breipohl et al. 1994; Chiang et al. 1999), which captures time-dependent aspects, together with a Monte Carlo sampling routine to capture the probabilistic aspects of uncertainty in unit availability, load level, inflow characteristics etc. Examples of chronological production costing programs include PWRSYM (Babb 1983), BENCHMARK (Manhire et al. 1982), PROSYM (PROSYM 1988), and DYNATRAN (EPRI 2001). The major advantages of the chronological simulation approach are its flexibility (it is easy to incorporate new or complex operating features) and the ease of explaining and interpreting the results (which can be seen as the simulation of the system operation over several scenarios). One possible limitation of this approach is related to the computational requirements: if many Monte Carlo repetitions of the chronological simulation algorithm are required in order to estimate the desired indices, the computational effort may be excessive. Variance reduction techniques (Oliveira et al. 1989; Breipohl et al. 1990), based on the smart sampling of system scenarios, can be used to alleviate this problem in PPC.

Given that the representation of generation aspects in probabilistic production costing is becoming fairly well established, the integration of transmission aspects in the PPC was the next needed step. Most previous multi-area production costing models have been based on the transportation network model (EPRI 1990; Ahsan et al. 1983; Hobbs and Ji 1995; Noyes 1983). But as Lee (EPRI 1990) points out, there is a need for both transportation network (capacity flow) and DC network (DC flow) models; a preference towards one or the other can be based on the particular application. Lee (EPRI 1990) presents a multi-area model with DC load flow constraints for a small system. In (PTI 1997), convolution methods are used to calculate the probability distribution of flows over selected transmission lines in large systems by taking advantage of the fact that flows in linearized models can be expressed as a linear function of the generation of individual units. However, that method cannot impose limits on flows. Such limits can be imposed in DC-based chronological models, which naturally lend themselves to probabilistic analysis using Monte Carlo methods (Pereira et al. 1992). The DC-based model then is expanded to full AC model in some commercial software, such as GE MAPS (Simons et al. 1993) and DYNATRAN (EPRI 2001).

C. Probabilistic Load Flow

There are two ways to adopting probabilistic approach to study load flow problems: Stochastic Load Flow (SLF) and Probabilistic Load Flow (PLF). In SLF study, the load and generation at an instant time are treated as random variables, and the impact of this uncertainty at each instant time is evaluated. Therefore, SLF method is to solve short-term uncertainties and is useful to system operation and short-term planning. In the evaluation of uncertainties over a long-term period, the PLF method is reviewed for long-term planning study purpose. Many PLF methods have been proposed to study load flow uncertainty problem (Borkowska 1974; Dopazo et al. 1975; Allan et al. 1981; Zhang and Lee 2004; Dimitrovski and Tomsovic 2004; Allan and Liete da Silva 1981; Wang and Alvarado 1992; Meliopoulos et al. 1990; Leite da Silva et al. 1984; Leite da Silva et al. 1985; Liete da Silva and Arienti 1990). These methods can be classified as simulation method, analytical method, or by a combination of both.

The simplest evaluation of the PLF problem is through Monte Carlo simulation (MCS). This method requires that the data involved to be assigned a probability distribution that characterizes the possible variation in the parameters. The random values from these distributions are selected and used to arrive at an estimate of load flow solution.

In order to reduce the computational effort, the analytical methods were used (Borkowska 1974; Dopazo et al. 1975; Allan et al. 1981; Zhang and Lee 2004; Dimitrovski and Tomsovic 2004; Allan and Liete da Silva 1981; Wang and Alvarado 1992; Meliopoulos et al. 1990). In (Borkowska et al. 1974), Borkowska first proposed a DC load flow model to take node data uncertainty into account and to find the distributions of branch flows. Reference (Dopazo et al. 1975) used a direct approach based on the principle of statistical least square estimation to compute the effects of uncertainties in input data on all output quantities and to obtain the expected value and variance of the solution of a load flow problem. A discrete frequency domain convolution technique by applying fast Fourier transformations and linearized power flow equation was proposed in (Allan et al. 1981) to increase the computation accuracy and to find the PDF of all output quantities. Reference (Zhang and Lee 2004) used a DC load flow model combined the concept of Cumulants and Gram-Charlier expansion theory to consider the bus injection uncertainties and to achieve enough accuracy with less computation effort to compute the approximate PDF and cumulative distribution function of network branch flows. A new boundary load flow method based on fuzzy/interval numbers was proposed in (Dimitrovski and Tomsovic 2004) to find the accurate boundary load flow solutions. Allan and Liete da Silva (Allan and Liete da Silva 1981) proposed a new PLF algorithm based on linearized models to account for the nonlinear network equations and to compute the distributions of output quantities. Wang and Alvarado (Wang and Alvarado 1992) used interval arithmetic to consider the uncertainty of nodal data and to find all possible solutions included within the bounds given by interval arithmetic. A new PLF method based on a description of bus power injections as random variables was proposed in (Meliopoulos et al. 1990) to consider the bus power injection uncertainties and operating practice of power systems. The main advantage of the analytical methods mentioned above is to avoid the computer simulations, but more assumptions and complex mathematical algorithms are required for these methods.

A MCS based on linear power flow equations combined with analytical convolution technique is used in (Wang and Alvarado 1992) to simplify computation process and maintain sufficient computation accuracy. Meliopoulos et al (Meliopoulos et al. 1990) proposed a new PLF algorithm combining MCS and multi-linearized load-flow equations to sufficiently and efficiently evaluate all load-flow result quantities. To consider network topology uncertainty, a new algorithm is proposed in (Leite da Silva et al. 1984). In this method, the distributions of the load-flow solution are conditionally evaluated based on each possible network configuration. The final solution then is obtained from weighted sum of density distributions by using the probability associated with each network configuration.

D. Security Constrained OPF (SCOPF)

Optimal Power Flow is a traditional power flow plus the introduction of an optimization cost function to the set of power flow equations. An Optimal Power Flow (OPF) program determines the settings of the selected control variables to achieve an optimal operation of a power system. Generally speaking an OPF is used to answer “How To” questions regarding power system operation. Typically the primary goal of a generic OPF is to minimize the costs of meeting the load demand for a power system while maintaining the security of the system. There is a need to include suitable security constraints into the whole market pricing mechanism, so that the correct market signals can be sent to all market participants while operating the system within reasonable security margins. Security constrained Optimal Power Flow (OPF) has been used to solve such kind of problems.

Thermal constrained OPF is discussed in (Bresesti et al. 2002). Recently, OPF-based models have been used for addressing voltage stability issues. Maximization of the loading parameter (Irisarri et al. 1997) or sensitivity index (Milano et al. 2005) in traditional voltage collapse studies, are as index to detect voltage instability, which is then used as a stability constraint to set up voltage stability constrained OPF. Another strategy is to use multi-objective OPF technique to maximize both the social benefit and the distance to a voltage instability point (Rosehart et al. 2003). Also, Canizares reviewed the voltage stability constrained OPF techniques before 2001 in (Canizares et al. 2001). On the other hand, some researches (Gan et al. 2000; Scala et al. 1998) have tried to incorporate the transient stability constraints directly into OPF, mainly by approximating the differential equations to algebraic equations. The chief advantage of this kind of method is that well-developed transient stability analyzing methodologies can be adopted. However, converting differential equations into algebraic equations by discrete scheme may not only suffer the inaccuracy of computation because of approximation but also cause convergence difficulties due to the introduction of a large number of variables and equations at each time step to the original OPF. In another research direction (Chen et al. 2001), the transient stability constrained OPF problem is equivalently converted into an optimization problem in the Euclidean space instead of tackling it directly.

An optimal power flow program has to solve an optimization problem where the objective function, equality and inequality constraints are non-linear. Many approaches have been presented in the literature over the years to solve the optimal power flow problem. Some of these approaches include: the Lambda iteration method which is also called the equal

incremental cost criterion method, the gradient method, the Newton method, the linear programming method, and the interior point method.

V. Existing Commercial Programs

The available commercial programs for energy storage evaluations and integrated planning were reviewed in (EPRI 1993) and (EPRI 1998), respectively. In 2003, the U.S. Department of Energy report (KEMA 2003) compared the analysis tools to assess national-interest transmission bottlenecks. This section will combine these reports with the EPRI programs for power system planning to introduce the programs that can be used to identify long-term transmission congestion. Providers and programs in this section are:

- EPRI (DYNATRAN, EGEAS, TRACE, PLF, CAR)
- ABB (TRACE, GridView)
- GE (MAPS)
- Henwood Energy Systems (PROSYM)
- LCG Consulting (UPLAN)
- PTI (MUST)
- Siemens (PROMOD)

DYNATRAN (EPRI 2001) simulates the operation of an electric utility combined generation and transmission system for the purpose of evaluating the production cost of running the system over typical generation planning horizons. The main strength of DYNATRAN's generation system modeling over other production program is its ability to assess some of the dynamic operating benefits of energy storage plants and its methodological emphasis on storage evaluation. The ability of DYNATRAN to model the effect of transmission network constraints on generation system operation opens up its applications to many new planning and operational problems.

Electric Generation Expansion Analysis System (EGEAS) is a modular generation expansion package to evaluate least-cost resource plans, independent power producers, avoided costs, and plant life extension programs. EGEAS integrates demand-and supply-side options into the most cost-beneficial plans which meet the specified reliability criteria according to a user-selected objective function.

Transfer Capability Evaluation (TRACE) (EPRI 1997) calculates both the simultaneous and non-simultaneous power transfer capabilities, which form the basis for the computation of Available Transfer Capability (ATC) postings mandated by FERC. TRACE performs fast calculations based on accurate AC power flow modeling and realistic system operation limits. An optional DC power flow model is also available for even faster solutions where voltage problems do not exist. The program consists of two major software modules--one module does efficient contingency selection and the second module is a full AC optimal power flow program. Together, the two modules calculate the multi-area, simultaneous power transfer capabilities subject to thermal, voltage, and interface limits.

Probabilistic Load Flow (PLF) (EPRI 2004) models the generation and loads in a probabilistic way, and computes the probabilistic distribution functions of line flow. PLF studies give system planners a better understanding of future system conditions and will provide more confidence in making judgments concerning alternative investments in transmission system. The latest version of PLF can handle the transmission network outage in probabilistic way and provides the generation dispatch features. It allows users to perform economic dispatch and user-defined dispatch.

Community Activity Room (CAR) (Lee 2003) is a newly developed analytical and visualization software technology mainly for modeling MW limits of power transfers among regions. CAR uses the metaphor of a many-sided room to show the ranges of operation within which power transactions and market activities can freely take place. CAR graphics define the limits of the power market, locating congested bottlenecks and suggesting the combinations of net import and net export from various control areas that will avoid congestion. The user interface capabilities of CAR are very unique. CAR as a product can improve the use-ability of congestion forecasting if it is integrated with TRACE and PLF as part of the whole package. By implementing a system like this the job of congestion forecasting will be easier. Rather than looking into tabular displays they can utilize this graphical tool. CAR has the potential to display market prices in the same way as congestion is displayed.

GridView has built-up databases for all NERC regions in the US. This product includes detailed modeling of generation, transmission, load and market structure modeling such as generation variable cost, forced outage rates, transmission interfaces and flow gates, hourly load profiles for different load serving entities etc. GridView's methodology combines generation, transmission, loads, fuels, and market economics modeling in one integrated framework to deliver location dependent market indicators (price, etc.), transmission system utilization measures and market performance indices. GridView uses a security constrained unit commitment and economic dispatch. This product also includes an optimal power flow. The package includes a market scenario module that can be used for simulation of different situations in the system.

Multi-Area Production Simulation (MAPS) is a model that calculates hour-by-hour production costs while recognizing the constraints on the dispatch of generation imposed by the transmission system. Using simulation techniques like Monte Carlo based on trends from production costs the Location Marginal Price (LMP) are forecasted. MAPS uses a detailed electrical model of the entire transmission network, along with generation shift factors determined from a solved ac load flow, to calculate the real power flows for each generation dispatch. This enables the user to capture the economic penalties of re-dispatching the generation to satisfy transmission line flow limits and security constraints. The chronological nature of the hourly loads is modeled for all hours in the year. In the electrical representation, the loads are modeled by individual bus. In addition to the traditional production costing results, MAPS can provide information on the hourly spot prices at individual buses and on the flows on selected transmission lines for all hours in the year, as well as identifying the companies responsible for the flows on a given line.

PROSYM is the chronological electric power production costing simulation computer software package. It is designed for performing planning and operational studies, and as a result of its chronological nature, accommodates detailed hour-hour investigation of the operations of electric utilities. Because of its ability to handle detailed information in a chronological fashion, planning studies performed with PROSYM may reflect actual electric utility operations.

UPLAN-E provides various levels of detail and capabilities for operational evaluation, production cost analysis, and electricity pricing, depending on the planning requirements. For the short-term, the program provides chronological production cost algorithms for analyzing the hourly costs and operations of an individual utility, an interconnected utility, a multi-area composed of several utilities, and a regional power market made up of sellers and buyers interconnected through a complex transmission network grid. For the longer term, UPLAN-E provides monthly and annual production accurate results for up to 30 years.

MUST (Managing and Utilizing System Transmission) efficiently calculates the impact of transactions on key network elements, identifies the most limiting contingencies and constraints, calculates FCITC transfer capability, and the sensitivity of FCITC to transactions and generation dispatch changes. MUST calculates transmission transfer capabilities in grids the size of the U.S. Eastern Interconnection in very short times (low seconds), and displays the impact of transactions and generation dispatch variations on transfer capabilities.

PROMOD VI is an integrated production costing and transmission planning software used to evaluate operating cost impacts of changes of load and generation. The PROMOD VI simulation incorporates full transmission modeling and a security-constrained unit commitment and dispatch, with Monte Carlo modeling of generator unit random outages.

VI. Conclusion

This paper is the result of an extensive literature survey of publications that may have some contribution to the understanding of probabilistic congestion forecasting. As a summary, it has reviewed some existing methodologies for generation and transmission planning. Some uncertainty related topics, such as energy price forecasting, probabilistic production costing, and probabilistic load flow, have also been discussed in the paper. It is concluded that an efficient and holistic method, which uses probabilities to more accurately describe load and generation time dependence, is essential to characterize and forecast long-term transmission congestion.

References

- Ahsan, Q., K.F. Schenk., and R.B. Misra 1983. "Expected Energy Production Cost of Two Interconnected Systems with Correlated Demands," *IEEE Trans. Power Apparatus and Systems*, 99(5), pp. 2155-2164.
- Alguacil, N., A. L. Motto, and A. J. Conejo 2003. "Transmission expansion planning: a mixed-integer LP approach," *IEEE Trans. Power Systems*, vol. 18, no. 3, August, pp. 1070–1077.
- Allan, R. N., A. M. Liete da Silva, and R. C. Burchett 1981. "Evaluation methods and accuracy in probabilistic load flow solutions," *IEEE Trans. on Power Apparatus and Systems*, Vol. PAS-100, No. 5, May, pp. 2539-2546.
- Allan, R. N. and A. M. Liete da Silva 1981. "Probabilistic load flow using multilinearizations," *IEE Proc. C: Generation, Transmission, Distribution*, Vol. 128, No. 5, pp. 280-287.
- Babb 1983. "POWRSYM – Production costing program. Version 52G", Tennessee Valley, April.
- Bahiense, L., G. C. Oliverira, M. Pereira, and S. Granville 2001. "A mixed integer disjunctive model for transmission network expansion," *IEEE Trans. Power Systems*, vol. 16, no. 3, August, pp. 560–565.
- Baldick, R. and E. Kahn 1993. "Transmission planning issues in a competitive economic environment," *IEEE Trans. Power Systems*, vol. 8, no. 4, November, pp. 1497–1503.
- Baleriaux, H., E. Jamouille, and L. de Guerchetin 1967. "Simulation de L'Exploitation d'Electricite Couple a des Stations de Pompage," *Revue E.*, vol.5, no. 7, pp.3-24.
- Bastian, J. , J. Zhu, V. Banunarayanan and R. Mukerji 1999. "Forecasting energy prices in a competitive market," *IEEE Computer Applications in Power*, vol. 12, July, pp. 40-45.
- Binato, S., M. V. F. Pereira, and S. Granville 2001. "A new benders decomposition approach to solve power transmission network design problems," *IEEE Trans. Power Systems*, vol. 16, no. 2, pp. 235–240, May.
- Bjorkvoll, T., S. E. Fleten, M. P. Nowak, A. Tomasgard, and S. W. Wallace 2001, "Power generation planning and risk management in a liberalized market," *Proceeding of IEEE Porto Power Tech 2001*, Porto, Portugal.
- Bloom, J. A. 1982. "Long-range generation planning using decomposition and probabilistic simulation," *IEEE Trans. Power App. Systems*, vol. 101, no. 4, April, pp. 797–80.
- Booth, R.R.1972. "Power system simulation model based on probability analysis," *IEEE Tran. on PAS*, vol. PAS-91, no. 1, pp. 62-69.
- Borkowska, B. 1974. "Probabilistic load flow," *IEEE Trans. on Power Apparatus and Systems*, Vol. PAS-93, No. 3, May/June, pp. 752- 759.
- Breipohl, A.M., F. H. Lee, J. Huang, Q. Feng 1990. "Sampling size reduction in stochastic production simulation," *1990 Winter Power Meeting*, paper 90-WM 122-2 PWRS.
- Breipohl, A.M., F.N. Lee, Min-Ywan Lin, Hui-Dong Wang, and Jia-Yo Chiang 1994. "An Adaptive Production Simulation Structure for Operational Planning with. Annual

- Emission/Fuel constraints," *IEEE Trans. on Power Systems*, Vol. 9, No. 3, August, pp. 1634-1642.
- Bresesti, P., D. Lucarella, P. Marannino, R. Vailati, F. Zanellini 2002. "An OPF-based procedure for fast TTC analyses," *2002 PES Summer Meeting*, vol. 3, pp.1504-1509.
- CAISO 2004. Transmission Economic Assessment Methodology (TEAM), CAISO. [online]. Available:
<http://www.caiso.com/docs/2004/06/03/2004060313241622985.pdf>
- Canizares, C. A., W. Rosehart, A. Berizzi, and C. Bovo 2001. "Comparison of Voltage Stability Security Constrained Optimal Power Flow Techniques," *Proc. Of IEEE-PES Summer Meeting 2001*, Vancouver, BC, Canada, July.
- Chanda, R. S. and P. K. Bhattacharjee 1994. "Application of computer software in transmission expansion planning using variable load structure," *Electric Power Systems Research*, no. 31, pp. 13-20.
- Chen, L., Y. Tada, H. Okamoto, R. Tanabe, and A. Ono 2001. "Optimal operation solutions of power systems with transient stability constraints," *IEEE Trans. Circuits Sys. I*, vol. 48, March, pp. 327-339.
- Chiang, J. Y., A.M. Breipohl, F.N. Lee, and R. Adapa 1999. "Probabilistic load variation modeling for estimating the variance of annual production cost," *IEEE Trans. Power Systems*, Vol. 14(1), pp. 362-368.
- Chuang, A.S., F. Wi, and P. Varaiya 2001. "A game-theoretic model for generation expansion planning: problem formulation and numerical comparisons," *IEEE Trans. Power Sys.*, vol. 17, 4, November, pp. 885-891.
- Contreras, J., R. Espínola, F. J. Nogales, and A. J. Conejo 2003. "ARIMA models to predict next-day electricity prices," *IEEE Trans. Power Systems*, vol. 18, August, pp. 1014-1020.
- David, A. K. and F. Wen 2001. "Transmission planning and investment under competitive electricity market environment," *2001 Proceeding IEEE Power Engineering Soc. Summer Meeting*, vol. 3, pp. 1725-1730.
- Dimitrovski, A. and K. Tomsovic 2004. "Boundary load flow solutions," *IEEE Trans. on Power Systems*, Vol. 19, No. 1, February, pp. 348-355.
- DOE 2003, *Analysis and Selection of Analytical Tools to Assess National-Interest Transmission Bottlenecks*, U.S. Department of Energy, March.
- DOE 2006. *National Electric Transmission Congestion Study*, Annual Report, U.S. Department of Energy, August. [Online]. Available:
http://www.oe.energy.gov/DocumentsandMedia/Congestion_Study_2006-9MB.pdf.
- Dopazo, J. F., O. A. Klitin, and A. M. Sasson 1975. "Stochastic load flows," *IEEE Trans. on Power Apparatus and Systems*, Vol. PAS-94, No. 2, March/April , pp. 299-309.

- Dusonchet, Y. P. and A. H. El-Abiad 1973. "Transmission planning using discrete dynamic optimization," *IEEE Trans. Power Apparatus and Systems*, vol. PAS-92, July, pp. 1358–1371.
- EPRI 1983. *Composite generation transmission expansion planning*, EPRI EL-5179, Project 2473-9, EPRI.
- EPRI 1988. *Mathematical decomposition techniques for power system expansion planning*, EPRI EL-5299, Project 2473-6, EPRI.
- EPRI 1989. *Decomposition techniques for multi-area generation and transmission planning under uncertainty*, EPRI EL-6484, Project 2940-1, EPRI.
- EPRI 1990. *Multi-Area Power System Reliability and Production Costing*. Technical Report EL-6912, EPRI.
- EPRI 1993. *Comparison of Planning Code for Energy Storage Evaluations*, Technical Report -101648, EPRI.
- EPRI 1997. *Trace User's Manual – A Beginner's Guide to the Use of the Trace*. Technical Report -107846, EPRI.
- EPRI 1998. *Inventory of Available Methods and Processes for Assessing the Benefits, Costs, and Impacts of Demand-Side Options: V3-Description and Review of Computer Tools for Integrated Planning*, Technical Report -108506, EPRI.
- EPRI 2001. *DYNATRAN: Dynamic analysis of interconnected systems with energy storage*. Technical Report-1006531, EPRI.
- EPRI 2004. *Probabilistic Transmission Planning: Summary of Tools, Status, and Future Plans*, Technical Report-1008612, EPRI.
- Fang, R. and D. J. Hill 2003. "A new strategy for transmission expansion in competitive electricity markets," *IEEE Trans. Power Systems*, vol. 18, no. 1, February, pp. 374–380.
- Gan, D., R.J. Thomas, R.D. Zimmerman 2000. "Stability-constrained optimal power flow," *IEEE Trans. Power Systems*, vol. 15, no. 2, May, pp.535-540.
- Gil, H. A., E. L. Silva, and F. D. Galiana 2002. "Modeling competition in transmission expansion," *IEEE Trans. Power Systems*, vol. 17, no. 4, November, pp. 1043–1049.
- Gorenstin, B. G., N. M. Campodonico, J. P. Costa, and M. V. F. Pereira 1993. "Power system expansion planning under uncertainty," *IEEE Trans. Power Systems*, vol. 8, no. 1, February, pp. 129–136.
- Gross, G., N.V. Garapic and B. McNutt 1988. "The mixture of normal approximations technique for equivalent load duration curves," *IEEE Trans. Power Systems*, vol. 3, no. 2, pp. 368–374.
- Hamoud, G. and L. Bradley 2001. "Assessment of transmission congestion cost and locational marginal pricing in a competitive electricity market," *2001 IEEE, Power Engineering Society Summer Meeting*, pp. 1468 -1472.

- Hippert, H. S., C. E. Pedreira, and R. C. Souza 2001. "Neural networks for short-term load forecasting: A review and evaluation," *IEEE Trans. Power Systems*, vol.16, February, pp. 44–55.
- Hobbs, B.F. and Y. Ji 1995. "A Bounding Approach to Multi area Production Costing," *IEEE Trans. Power Sys.*, 10(2), pp. 853-859.
- Hong, Y. and C. Hsiao 2002. "Locational marginal price forecasting in deregulated electricity markets using artificial intelligence," *Generation, Transmission and Distribution, IEE Proceedings*, vol. 149 Issue 5, September, pp. 621 -626.
- Hu, Z. G. 2002, "Study on methodology of generation expansion planning for power restructuring," *2002 Proceeding of PowerCon*, vol. 1, pp. 388–392.
- Irisarri, G. D., X.Wang, J. Tong, and S. Mokhtari 1997. "Maximum loadability of power systems using interior point nonlinear optimization method," *IEEE Trans. Power Systems*, vol. 12, no. 1, February, pp. 162–172.
- Jia, N. X., R. Yokoyama, and Y. C. Zhou 2001. "Advanced DP based method for power expansion planning in the restructured market," *2001 Proceeding of IEEE Porto Power Tech 2001*, Porto, Portugal.
- Khokhar, J. S.1997, *Programming Models for the Electricity Industry*, New Delhi, Delhi: Commonwealth Publishers, pp. 21–84.
- Kim, K. J., Y. M. Park, and K. Y. Lee 1988. "Optimal long term transmission expansion planning based on maximum principle," *IEEE Trans. Power Systems*, vol. 3, November, pp. 1494–1501.
- Latorre, G., R.D. Cruz, and J. M. Areiza 2003. "Classification of Publications and Model on Transmission Expansion Planning," *IEEE Trans. Power Systems*, vol. 18, May, pp. 938–946.
- Lee, S. T., K. L. Hocks, and E. Hnyilicza 1974. "Transmission expansion using branch-and-bound integer programming with optimal cost-capacity curves," *IEEE Trans. Power Apparatus and Systems*, vol. PAS-93, pp. 1390–1400.
- Lee, S.T. 2003. "Community Activity Room as a New Tool for Transmission Operation and Planning Under a Competitive Power Market," *2003 IEEE Power Tech Conference*, Bologna, Italy, June 22-26.
- Leeprechanon, N., S. S. Moorthy, R. D. Brooks, A. K. David, and F. Wen 2001. "Transmission planning in deregulated systems: a model for developing countries," *2001 IEEE Porto Power Tech Conf.*, Porto, Portugal.
- Leite da Silva, A. M., R. N. Allan, and V. L. Arienti 1985. "Probabilistic load flow considering network outages," *IEE Proc. C: Generation, Transmission, Distribution*, Vol. 132, No. 3, May, pp. 139-145.

- Leite da Silva, A. M., V. L. Arienti, and R. N. Allan 1984, "Probabilistic load flow considering dependence between input nodal powers," *IEEE Trans. on Power Apparatus and Systems*, Vol. PAS-103, No. 6, June, pp. 1524-1530.
- Li, W. and R. Billinton 1993. "A minimum cost assessment method for composite generation and transmission system expansion planning," *IEEE Trans. Power Systems*, vol. 8, no. 2, May, pp. 628-635.
- Liete da Silva, A. M. and V. L. Arienti 1990. "Probabilistic load flow by a multilinear simulation algorithm," *IEE Proc. C: Generation, Transmission, Distribution*, Vol. 137, No. 4, July, pp. 276-282.
- Manhire, B. and R.T. Jenkins 1982. "BECHMARK: A Monte Carlo hourly chronological simulation model which includes effects of ramp-rates and reservoir constraints," 1982 *Proceedings of the conference on generation planning: Modeling and decision making*, Chattanooga, Tennessee, August 10-12.
- McCusker, S. A., B. F. Hobbs, and Y. Ji 2002, "Distributed utility planning using probabilistic production costing and generalized benders decomposition," *IEEE Trans. Power Systems*, vol. 17, no. 2, May, pp. 497-505.
- Meliopoulos, A. P. Sakis, G. J. Cokkinides, and X. Y. Chao 1990. "A new probabilistic power flow analysis method," *IEEE Trans. on Power Systems*, Vol. 5, No. 1, February, pp. 182-190.
- Melo, A. C. G., G. C. Oliveira, M. Morozowski, and M. V. F. Pereira 1991. "A hybrid algorithm for Monte Carlo/enumeration based composite reliability evaluation," 1991 *Probabilistic Methods Applied to Electric Power Systems Conf.*, pp. 70-74.
- Milano, F., C.A. Canizares, and A.J. Conejo 2005. "Sensitivity-based security-constrained OPF market clearing model," *IEEE Trans. Power Systems*, vol. 20, no. 4, November, pp.2051-2060,
- Neelakanta, P. S. and M. H. Arsali 1999. "Integrated resource planning using segmentation method based dynamic programming," *IEEE Trans. Power Systems*, vol. 14, no. 1, February, pp. 375-385.
- Nicolaisen, J. D., C. W. Richter Jr., and G. B. Sheblé 2000. "Price signal analysis for competitive electric generation companies," *Proc. Conf. Elect. 2000 Utility Deregulation and Restructuring and Power Technologies*, pp. 66-71.London, U.K., Apr. 4-7.
- Nogales, F. J., J. Contreras, A. J. Conejo, and R. Espínola 2002. "Forecasting next-day electricity prices by time series models," *IEEE Trans. Power Systems*, vol. 17, no. 2, May, pp. 342-348.
- Noyes, L.R. 1983. "Two-Area Probabilistic Production Costing by the Method of Bivariate Cumulants," *IEEE Trans. Power Apparatus and Systems*, PAS-102(2), pp. 433-443.
- Oliveira, G.C., M.V.F. Pereira, and S.H.F. Cunha 1989. "A technique for reducing computational effort in Monte Carlo based composite reliability evaluation," *IEEE Trans. Power Systems*, vol. 4, no. 4, November, pp. 1309-1315.

- Park, J. B., J. H. Kim, and K. Y. Lee 2002. "Generation expansion planning in a competitive environment using a genetic algorithm," *2002 Proceeding of IEEE Power Engineering Soc. Summer Meeting*, vol. 3, pp. 1169–1172.
- Park, J. B., Y. M. Park, J.-R. Won, and K. Y. Lee 2000. "An improved genetic algorithm for generation expansion planning," *IEEE Trans. Power Systems*, vol. 15, no. 3, August, pp. 916–922.
- Park, Y. M., J. R. Won, J. B. Park, and D.-G. Kim 1999. "Generation expansion planning based on an advanced evolutionary programming," *IEEE Trans. Power Systems*, vol. 14, no. 1, February, pp. 299–305.
- Pereira, M.V.F. et al. 1992. "Chronological Probabilistic Production Costing and Wheeling Calculations with Transmission Network Modeling," *IEEE Trans. Power Systems*, Vol. 7(2), pp. 885-889.
- Pinto, L. M. V. G. and A. Nunes 1990. "A model for the optimal transmission expansion planning," *10th Power Systems Computation Conf., 1990*, pp. 16–23.
- PROSYM 1988. "PROSYM – Utility planning software," The simulation group, Sacramento, CA.
- PTI 1997. "TOPS: Transmission-Oriented Production Simulation Program," Software Program-238, Power Technology, Inc. Schenectady, NY.
- Romero, R. and A. Monticelli 1994. "A hierarchical decomposition approach for transmission network expansion planning," *IEEE Trans. Power Systems*, vol. 9, no. 1, February, pp. 373–380.
- Rosehart, W. D., C. A. Cañizares, and V. Quintana 2003. "Multi-objective optimal power flows to evaluate voltage security costs in power networks," *IEEE Trans. Power Systems*, vol. 18, no. 2, May, pp. 578–587.
- Scala, M. La, M. Trovato, and C. Antonelli 1998. "On-line dynamic preventive control: An algorithm for transient security dispatch," *IEEE Trans. Power Systems*, vol. 13, May, pp. 601-610.
- Schenk, K.F., R.B. Misra, S. Vassos and W. Wen 1984. "A New Method for the Evaluating of Expected Energy Generation and Loss of Load Probability," *IEEE Tran. on PAS*, vol. PAS-103, no. 2, pp. 294-303.
- Sidenblad, K.M. and S.T.Y. Lee 1981. "A probabilistic production costing methodology for system with storage," *IEEE Tran. on PAS*, vol. PAS-100, pp. 3116-3124.
- Simons, N.W., L.L. Garver, and R. Mukerji 1993. "Transmission Constrained Production Costing: A Key to Pricing Transmission Access," *Public Util. Fortnightly*, 131(3).
- Storn, R. and K. Price 1997. "Differential evolution—A simple and efficient heuristic for global optimization over continuous spaces," *J. Global Optimiz.*, vol. 11, pp. 341–359.
- Stremel, J. P., R. T. Jenkins, R.A. Babb, and W.D. Bayless 1980. "Production costing using the cumulant method of representing the equivalent load curve," *IEEE Tran. on PAS*, vol. PAS-99, no. 5, pp. 1947-1956.

- Szkuta, B. R., L. A. Sanabria, and T. S. Dillon 1999. "Electricity price short-term forecasting using artificial neural networks," *IEEE Trans. Power Systems*, vol. 14, August, pp. 851–857.
- Tabors, R. D. 2000. "Unbundling of transmission: the operation and economics of a for-profit transmission company," *33rd Hawaii Int. Conf. System Sciences*, Maui, HI.
- Tsamaspfyrpu, P., A. Renaud, and P. Carpentier 1999. "Transmission network planning: An efficient Benders decomposition scheme," *Proceeding of 13th PSCC in Trondheim, 1999*, pp. 487–494.
- Valenzuela, J. and M. Mazumdar 2001. "On the computation of the probability distribution of the spot market price in a deregulated electricity market," *Proc. 22nd Power Ind. Computer Application Int. Conf.*, pp. 268–271, Sydney, Australia, May.
- Villasana, R., L. L. Garver, and S. L. Salon 1985. "Transmission network planning using linear programming," *IEEE Trans. Power Apparatus and Systems*, vol. PAS-104, February, pp. 349–356.
- Wang, A. J. and B. Ramsay 1998. "A neural network based estimator for electricity spot-pricing with particular reference to weekend and public holidays," *Neurocomputing*, vol. 23, pp. 47–57.
- Wang, X. 1988. "Equivalent energy function approach to power system probabilistic modeling," *IEEE Trans. Power Systems*, vol. 3, no. 3, pp. 823–829.
- Wang, Z. and F. L. Alvarado 1992. "Interval arithmetic power flow analysis," *IEEE Trans. on Power Systems*, Vol. 7, No. 3, August, pp. 1341–1349.
- Wong, W., H. Chao, D. Julian, P. Lindberg, and S. Kolluri 1999. "Transmission planning in a deregulated environment," *1999IEEE Transmission and Distribution Conf.*, vol. 1, pp. 350–355.
- Youssef, H. K. and R. Hackam 1989. "New transmission planning model," *IEEE Trans. Power Systems*, vol. 4, February, pp. 9–18.
- Zhang, P. and S. T. Lee 2004. "Probabilistic load flow computation using the method of combined Comulants and Gram-Charlier expansion," *IEEE Trans. on Power Systems*, Vol. 19, No. 1, February, pp. 676–682.
- Zhu, J. and M.-Y. Chow 1997. "A review of emerging techniques on generation expansion planning," *IEEE Trans. Power Systems*, vol. 12, no. 4, November, pp. 1722–1728.

Appendix B: Summary of Probabilistic Models

Introduction

In this research project, probabilistic modeling was used extensively to address the uncertainty of the many variables affecting transmission congestion. In this Appendix, the various probabilistic models are summarized and comments are made by the researchers to provide the readers some sense of confidence about the assumptions made in this research project, especially with respect to the nature of the probabilistic models being used. Where insufficient data analysis or research has been done in this project to validate the assumptions, they will be noted.

Model of the Annual Peak Load Increase

The area projection of load demand is modeled as a function of its last year's peak load demand and the load demand growth rate. It is assumed that the annual peak load of one area is increased with a time series exponential expression as the following equation:

$$L_i(t) = L_i(t-1)e^{\lambda_i(t)} \quad (\text{B-1})$$

where $L_i(t-1)$ is the value of the last year's peak load demand of the i^{th} area; $\lambda_i(t)$ is the increase exponent of this year which is decided by the load demand growth rate.

Load growth is a complex process which basically is the summation of a very large number of variables involving electricity customers' addition of things that use electricity and their electricity usage patterns which may respond to uncertain factors such as weather, etc. Thus it is simplistic to assume a constant growth rate. In Equation (B-1), a growth rate that may change each year is assumed instead. It should be recognized that the growth rate may be a negative value as well as a positive value. The challenge is in estimating these growth rates. Econometric models are typically used to make such forecasts.

Because the load demand growth rate is uncertain and cannot be precisely predicted, for the purpose of this research project, the exponent λ_i is treated as a probabilistic variable which is assumed to follow the normal distribution. (Note that there is a reasonable argument for the normal distribution because of the Central Limit Theorem¹⁹, which states that the sum of a large number of independent random variables, no matter what their own probability distributions are like, will approach the normal distribution.)

See Section 2.1.1 for a more detailed analysis on the reasonableness of using the normal distribution to approximate the annual peak demand growth rate. It can be seen that even with

¹⁹ "Central Limit Theorem." From MathWorld--A Wolfram Web Resource.
<http://mathworld.wolfram.com/CentralLimitTheorem.html>

the limited number of samples, the general shape of the histogram is approximately symmetrical and centered on a mean value, which fits a bell-shape curve reasonably well.

Probabilistic Model of Generation Capacity

The probability density function of the in-service date for an installed new generator or the retirement date of an old generator has the following form:

$$f(x) = \begin{cases} \lambda e^{-\lambda(x-t_0)}, & x \geq t_0 \\ 0, & x < t_0 \end{cases} \quad (\text{B-2})$$

where t_0 is the planned in-service date or the planned retirement date; λ is the delay rate.

Since this is a discrete annual assessment and the exponential distribution is a continuous analog of the [geometric distribution](#), the [geometric distribution](#) is used in this report. The probability density function of the in-service date for an installed new generator or the retirement date of an old generator has the following form:

$$f(x) = \begin{cases} \lambda(1-\lambda)^{(x-t_0)}, & x \geq t_0 \\ 0, & x < t_0 \end{cases} \quad (\text{B-3})$$

By sampling the in-service date and retirement date in Monte Carlo simulation, the generation capacity for a specific year can be calculated in a probabilistic way.

In order to build this model, the following variables need to be decided:

- t_0 : The planned in-service date or the planned retirement date.
- λ : The delay rate parameter of the in-service date or the retirement date of one specific generator

Delay rate λ is the key parameter of this exponential distribution model for in-service or retirement date. The value of delay rate should be decided by the user based on the historical data or reasonable assumptions. For example, if there is a 12 percent probability of delay in this project, the value of λ will be set as 0.88, The PDF of the in-service date for a new generator or the retirement date for an old generator is as shown in Figure B-1.

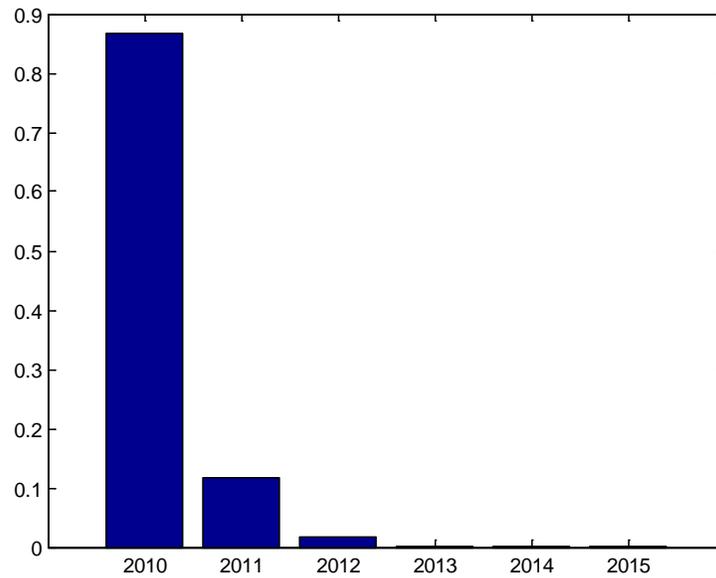


Figure B-1: PDF of in-service date for a new generator or retirement date for an old generator

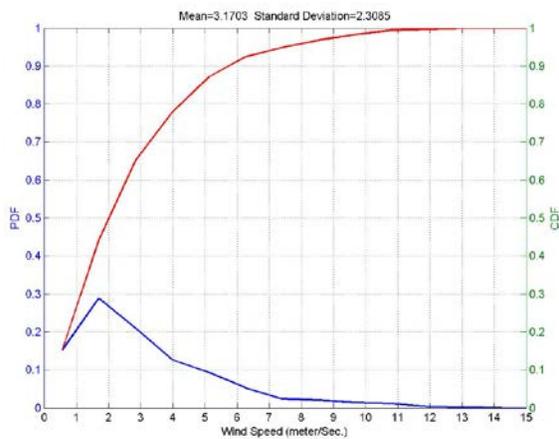
In this project, no analysis was done on the actual shape of the probability distribution of the inservice dates of power plant units. The discrete probability mass function as shown in Figure B-1 is a reasonable assumption if one assumes that the study is being done using the most recent scheduled inservice dates that are available. Generally, there are more factors that can delay the inservice date than those that can advance the inservice date. In some situations, e.g., when there is a financial incentive to complete construction ahead of schedule, there is some probability that the inservice date probability may be nonzero for $x < t_0$. Also, because projects tend to be completed within reasonable time after the scheduled date, the shape of the probability distribution should decline over time. An exception would be if a project eventually is canceled. In that case, there would be a finite probability that the inservice date is infinity.

The modeling of retirement date is likely to be somewhat different from the modeling of the inservice date. Here, the financial incentive is to postpone retirement as much as possible, as long as the generating unit continues to have economic values derived from its operation. This is because the original investment is a sunk cost, and by the scheduled retirement date or even sooner, the full capital cost has probably been fully recovered. The cost of maintaining, repairing and replacing old or worn out parts would be much lower than the cost of building another plant. Therefore, the probability distribution of the retirement date will likely spread out for quite a few years beyond the scheduled year. This of course can be modeled by a longer delay rate of the probability distribution. The general shape of the probability distribution as shown in Figure B-1 would still be reasonable.

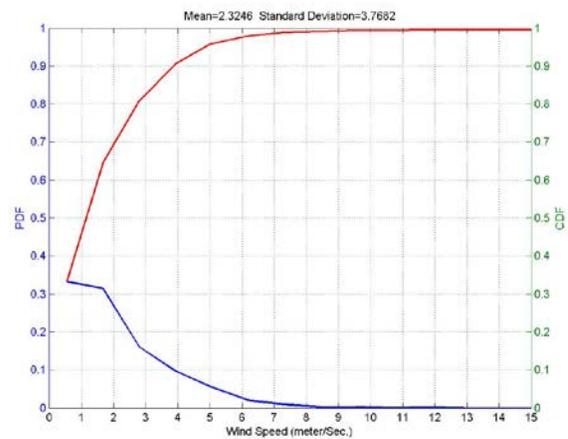
Model of Wind Generation Peak Condition Uncertainty

For this study, wind speed data (from year 2001 to year 2006) from Forecast Systems Laboratory (FSL)²⁰ were used as historical data. For each station, the data were collected at 0:00 and 12:00 UTC of each day. The data collected at 12:00 UTC of each day, which is 4:00 a.m. standard Pacific Time, were used to investigate the PDF of peak wind speed. *This limited approach is to test the idea of using wind speed statistics to investigate the uncertainty of wind speed over a long term of several years, i.e., seasonal uncertainty. It is also intended to provide some data to look into the spatial diversity over the long distances between Northern and Southern California and other adjacent regions to California. This limited analysis is not intended to be a definitive study.*

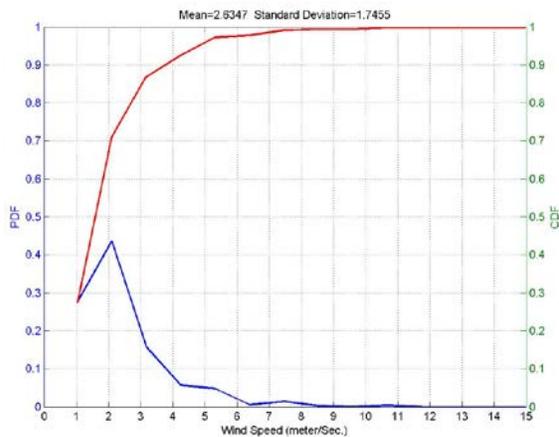
Historical data of wind speed were obtained for four different locations, as plotted below.



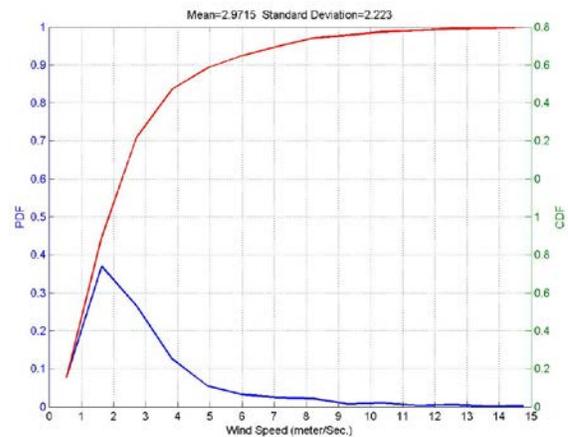
a. Oakland (CA) Wind Speed PDF



b. Vandenberg (CA) Wind Speed PDF



c. Denver (CO) Wind Speed PDF



d. Albuquerque (NM) Wind Speed PDF

Figure B-2: PDF of Wind Speed in Four Stations: Oakland (CA), Vandenberg (CA), Denver (CO), and Albuquerque (NM).

²⁰ The data are from the FSL, Radiosonde Data Archive, [Online]. Available: <http://raob.fsl.noaa.gov/>.

Figure B-2.a shows the probabilistic density function of wind speed in Oakland, CA. Figures B-2 b, c and d show the probabilistic density functions of wind speed over the same time period and the same measuring hour for Vandenberg, S. California, Denver and Albuquerque. The mean wind speed in Oakland (CA), Vandenberg (CA), Denver (CO) and Albuquerque (NM) is 3.17m/s, 2.32m/s, 2.63m/s and 2.97m/s respectively²¹.

The distribution functions of wind speed at these four locations are close to the Rayleigh Distribution, which is generally applied to model the annual or monthly wind speed distribution. For a typical Rayleigh distribution, the probability density function of wind speed has the following form:

$$f(x, \sigma) = \frac{x}{\sigma^2} e^{\left(\frac{-x^2}{2\sigma^2}\right)} \quad (\text{B-4})$$

where σ is the parameter; x is the wind speed.

Figure B-3 shows a typical wind power output curve of a 1000 MW wind turbine from a manufacturer. We project the statistical characteristic of wind speed in Oakland (CA) onto the wind power output. Figure B-4 shows the PDF of the wind generation output, which is close to exponential distribution. To consider and model wind power output uncertainty in the long term transmission congestion forecast, it could be assumed that the wind generation outputs in long term planning model follow exponential distribution.

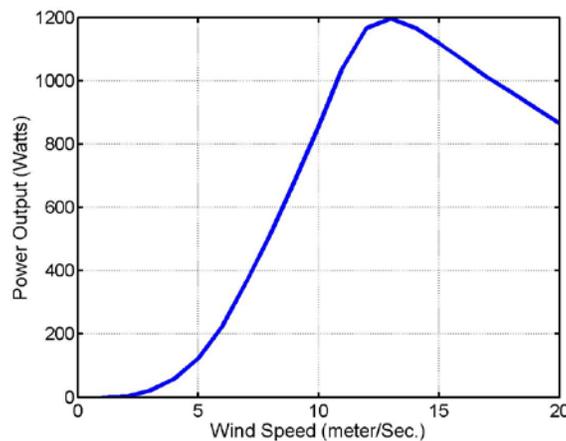


Figure B-3: Sample Wind Power Curve vs. Wind Speed²²

²¹ 1m/s = 2.2369mph.

²² <http://www.bergey.com/Technical/XL.1.R.xls>

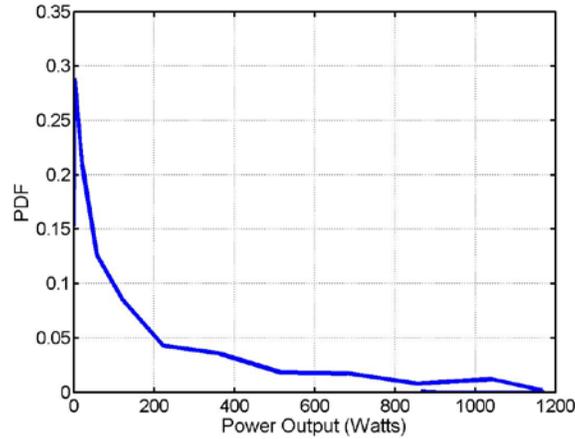


Figure B-4: PDF of Wind Generation Outputs

Model of the Day-Ahead Load Forecasting Uncertainty

The actual load demand is uncertain and might be different from the Day Ahead forecasted value, but it still can be predicted. Based on our observation, load deviation to the forecasted value is not an unchanged value at different time. So, it is assumed that the forecasted load follows the normal distribution at a certain time. Mean value and standard deviation of the hourly load are functions of time:

$$L(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu(t))^2}{2\sigma(t)^2}} \quad (\text{B-5})$$

where

$\mu(t)$ is the mean value of the day-ahead (DA) hourly forecasted load at time t ;

$\sigma(t)^2$ is the variance of the hourly forecasted load at time t that is introduced to consider the load forecasting errors.

In this application of the normal distribution assumption, there is strong support. Short term load measured at the system level is the resulting aggregation of a large number of random variables, i.e., the electricity usage of all the individual electricity customers. Though there are certain factors that make them not completely independent of one another, for example, time of the day, common weather conditions, etc., there are sufficient variations in the other factors affecting electricity usage that the overall effect is an illustration of the Central Limit Theorem, previously cited in this Appendix.

Generation Forced Outage Rate

To model generation uncertainties, a forced outage rate (FOR) is used to express the outage rate of a generation unit. This is the same as using a Binomial probability distribution function. Forced Outage Rate controls [generator](#) forced outages, a partial or complete loss of generating

capability for a certain period of time. The method of modeling is well known and used in generation reliability models, e.g., for Loss Of Load Probability (LOLP) models, which uses the binomial distribution for individual generator outages and combine them into a discrete probability mass function of the total available generation capacity.

Model of the Day-Ahead Wind Forecasting Uncertainty

Figure B-5 illustrates the fluctuations of one-minute wind power delivered to the California grid during a single 24-hour period on a summer day in California (CA ISO, 2005). The California daily wind generation profile in Figure B-5 is typical of a summer day in the areas affected by the coastal marine-layer. Wind speed and generation build up during the afternoon as the marine-layer spreads from the high-pressure region over the cool Pacific Ocean to the low pressure regions over the hot interior valleys. It then reaches a peak in early evening, and begins to fall off during the early morning hours reaching a minimum between about 10:00 a.m. and 2:00 p.m.

Unfortunately, the daily patterns of wind power are opposite to the typical daily system load curve in Figure 3-1. During the evening, wind generation is reaching its peak when system load is reaching its trough. During the day time, the reverse situation exists. When load is increasing rapidly in the morning, wind generation is diminishing to its minimum.

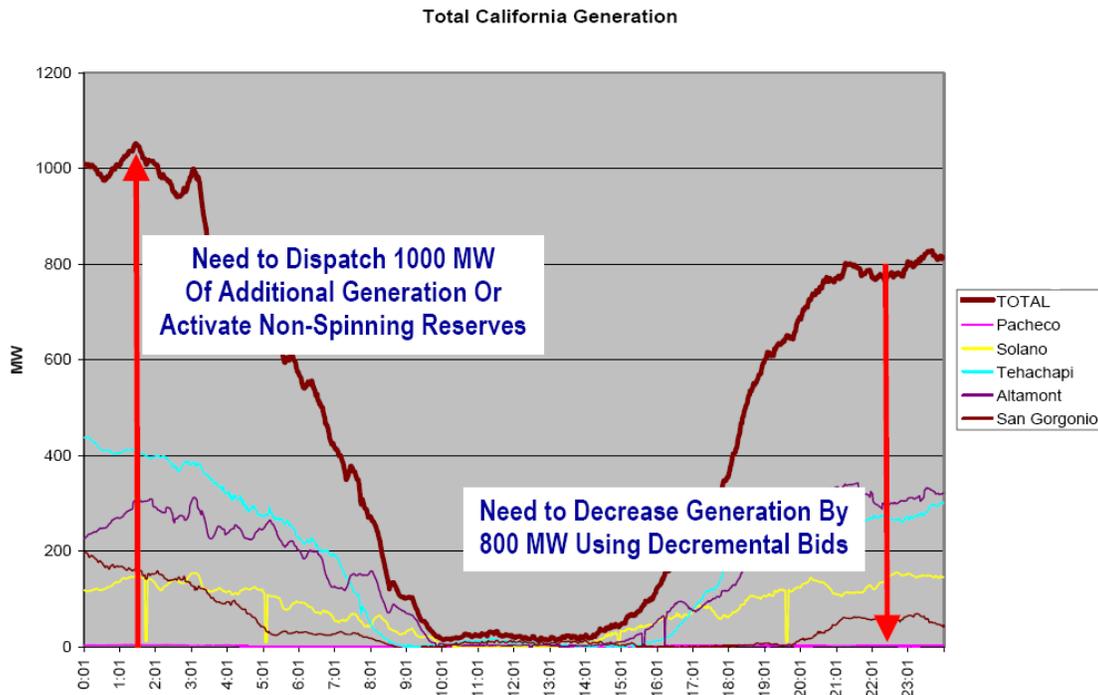


Figure B-5: Typical Variation of Total and Regional One-Minute Wind Generation in California on a Summer Day (CA ISO, March 2005)²³

²³ This figure is from the CEC report "California Regional Wind Energy Forecasting System Development".

For the congestion forecasting study, we will also consider the wind forecasting error in this study²⁴. Day-Ahead (DA) wind forecasting error could be 15 percent of the total installed wind generation capacity or 50 percent of the forecasted wind power production.

Another factor to consider is that the wind power outputs from different locations cannot be assumed to be independent of one another. It will be affected by common weather conditions. Some wind power output prediction methods also use statistical time-series models. For example, Autoregressive integrated moving average (ARMA) model in (Milligan et al. 2003) is shown as:

$$X_t = \sum_{j=1}^p a_j X_{t-j} + \sum_{k=0}^q b_k e_{t-k} \quad (\text{B-6})$$

where the current time t 's observation X_t depends on a linear combination of past observations of X_{t-j} plus a moving average of series e_{t-k} , which is a white-noise process characterized by zero mean and variance as function of time. p is the order of the autoregressive process of X on itself and q is the order of the moving-average error term.

It should be noted by the readers that the subject of modeling the uncertainty of short term wind power output is early in its research. The behavior of the wind power output is highly dependent on the geographical and weather effects of specific sites. Different parts of the world, even with the best wind regimes for wind power plants, exhibit very different intraday wind output profiles, and each site also may exhibit large variations in the seasonal profiles. Therefore, it is essential that the best data be collected for actual wind sites. Instead of using simplified probabilistic models for wind power output, it may be preferable to use historical samples of wind output and use them directly in Monte Carlo simulations.

Probabilistic Model of Transmission Capacity

Similar to the probabilistic model of generation with the consideration of FORs, FOR is also used to express the outage rate of a transmission line to model transmission uncertainties.

As an example, the FOR of a line is estimated using the following equation:

$$FOR = \text{Outage Freq} \times \text{Repair Time} / 8760 \quad (\text{B-7})$$

In the above equation, the outage frequency is estimated using the following equation:

$$\text{Outage Freq} = a + b \times (Z / Z_{puPerMile}) \quad (\text{B-8})$$

where a (1/year) is the constant parameter of the forced outage frequency, b (1/year/mile) is the proportional parameter of the forced outage frequency, $Z_{puPerMile}$ (pu/mile) is the average

²⁴ This is not development of wind forecasting methods, but rather considering the forecasting error in the study and specifying the uncertain impact on system operation and planning.

impedance (p.u.) per mile used to estimate the line length, RepairTime (hour) is the average repair time (hour) after a forced outage.

Note that the simple probabilistic model using the Binomial distribution is a reasonable approach when there is insufficient data to identify the best probability distribution function for fitting the data.

The rationale is that, in the case of a complete lack of information about a probabilistic process, the best model is to establish the upper and lower maximum bounds of the random variable. This modeling approach is called Unknown But Bounded (UBB), and its original author was Fred C. Schweppe.²⁵ This may be viewed also as a uniform probability distribution function over the range between the lower and the upper bounds. In the case where the random variable is not a continuous variable, but a discrete variable which can take on either the value of on or off, then the best model is a discrete Binomial distribution of the two states. The probability of outage is then based on the statistical data on the average time over a long duration that one would find the variable in the outage state.

Conclusions

This Appendix has summarized the probabilistic models assumed in this research project. Comments have been made by the researchers to provide the readers some sense of confidence about these modeling assumptions. These notes could be useful for researchers or users of the results of this research project.

²⁵ Uncertain Dynamic Systems, by Fred C. Schweppe, Prentice-Hall, Inc., 1973.