

VARIABILITY IN OUTPUT AND RELIABILITY OF BROADLY DISTRIBUTED
WIND FARMS AND SOLAR ARRAYS AS A FUNCTION OF THE SYSTEM
SCALE

BY

HARIKRISHNA RAO

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Electrical and Computer Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2012

Urbana, Illinois

Adviser:

Professor Peter W. Sauer

ABSTRACT

This study was motivated by the current difficulties faced by the electric power industry in maintaining acceptable levels of electric grid reliability in broadly distributed wind and solar resources. This thesis presents a study on the geographic diversification of wind turbines and solar arrays to minimize the variance in the expected renewable output according to the mean variance portfolio theory. The uncorrelatedness in the wind and solar patterns amongst geographically diverse sites could compensate shortfalls in electric generation in a region. By studying the effects of deployment of renewable capacity in progressively larger deployment zones within the U.S. we hope to identify an optimal scale for a broadly distributed system as dictated by its reliability. The study monitors measurements of the loss of load probability (LOLP), expected unserved energy (EUE) and the ramping of the controllable load to study the electric reliability of hybrid wind and solar resources as function of the system scale.

To my family and dear friend Yannick Degeilh

ACKNOWLEDGEMENTS

I would like to sincerely thank my adviser Professor Peter W. Sauer for his tremendous support and the opportunity to pursue my academic interests in electric power studies at the University of Illinois. I thank Professor Angus F. Rockett for proposing the research topic for my thesis. I cannot thank Yannick Degeilh enough for being my constant companion in the hours it took to accomplish this study. His guidance in the optimization studies presented in this thesis helped me greatly. Special thanks to my parents and grandparents for the unwavering support and motivation they provided me throughout this journey called life!

TABLE OF CONTENTS

LIST OF TABLES.....	vi
LIST OF FIGURES.....	vii
CHAPTER 1: INTRODUCTION.....	1
1.1 Motivation.....	1
CHAPTER 2: WIND AND SOLAR DATA.....	6
2.1 Literature review.....	6
2.2 Identification of a viable region and system for analysis.....	8
2.3 Wind speed and solar insolation measurements.....	11
2.4 Weibull distribution and conversion of data into system inputs.....	14
2.5 Assumption of correlation between turbines in a site.....	18
CHAPTER 3: THE MODELED ERCOT SYSTEM.....	20
3.1 Geographic diversification and portfolio theory.....	20
3.2 Optimization problem and geographic configurations.....	23
3.3 ERCOT load and conventional generation sources.....	29
CHAPTER 4: ESTIMATION OF ELECTRIC RELIABILITY MEASUREMENTS...	32
4.1 Generation of the simulated optimal geographic configuration.....	32
4.2 Reliability measures: Calculation of LOLP, EUE and ramping.....	34
4.3 Comparison of reliability measures for various geographic configurations.....	41
CHAPTER 5: CONCLUSION.....	43
5.1 Results and inferences.....	43
5.2 Future work.....	55
AFTERWORD.....	57
APPENDIX: MATLAB CODE.....	59
REFERENCES.....	76

LIST OF TABLES

Table 4.1 Installed wind capacity per deployment zone.....	33
Table 5.1.LOLP and EUE as a function of the system scale.....	43
Table 5.2 LOLP and EUE as a function of reserves and renewables.....	46
Table 5.3 Effects of load correlation on LOLP as a function of system scale.....	49
Table 5.4 Ramping as a function of load correlation.....	50
Table 5.5 Ramping as a function of the expected renewable output.....	51

LIST OF FIGURES

1.1 Geographic range selection	3
1.2 ERCOT Zones	3
2.1 Hourly correlation between sites vs. distance.....`	10
2.2 Manufacturer’s power curve for GE 1.5 MW turbines.....	15
2.3 Derived power curve for GE 1.5 MW turbines.....	15
4.1 Distribution of wind resources by zone	34
4.2. LOLP estimate over an increasing number of samples.....	36
4.3 Coefficient of variation of LOLP over a number of samples.....	37
5.1 EUE as a function of system scale.....	44
5.2 Zero LOLP – Reliable system scenario.....	52
5.3 LOLP Failure.....	53
5.4 EUE improvement with increase in expected renewable energy.....	55

CHAPTER 1

INTRODUCTION

1.1 Motivation

Renewable sources of energy such as wind and solar resources are ‘free,’ perpetual and clean, but have yet to be deemed reliable as compared to conventional sources of energy such as coal, gas and nuclear power. The reliability of these renewable sources is a major concern with the rapid growth and increased impetus placed on renewable energy sources and with the electric power industry seeking to find ways to reduce the variability in the output of these resources. Considerable research into the application of the portfolio diversification theory to renewable energy such as wind and solar farm siting has been performed. Research by Drake and Hubacek, illustrate the application of the ‘efficient frontier’ method whereas Roques, Saguan and Hiroux create portfolios of different wind sites and state that these methods hold promise in reducing output variability [1,2]. Rombauts, Delarue and D’haeseleer also use the optimal portfolio based allocation of wind power but also consider cross border constraints [3]. This thesis circumvents this issue placed by politics by selecting a region that is relatively autonomous and is not subject to man-made political restrictions. Geographic diversification does manage in reducing the variability, but it is still unclear how distributed these resources need to be in terms of their geographic location to optimally reduce the uncertainty of power output. The reduction in the variability depending on the geographic distribution needs to be quantified in order to identify this optimal scale. This study aims to measure reliability parameters to quantify the reduction in power output variability of distributed solar and wind hybrid resources as a function of the geographic distance between them. It also will investigate the ‘ramping’ of conventional generation in instances when there is a shortfall in the power supplied to loads from renewable resources.

Multiple research studies have documented that definite benefits exist in site diversification pertaining to the reduction in the variability in the power output of renewable energy sources such as wind farms. However, similar claims pertaining to solar arrays and a combination of hybrid wind and solar resources are scant. The bulk of the research conducted in the E.U., U.K. and even the U.S. thus far documents the potential benefits of geographic wind diversification at large. No specific attempts have been made to combine various other sources of renewable energy and quantify the resulting reduction in variability while accounting for the conventional generation and the load served in the area. Since this thesis relies heavily on the load and conventional generation, it was necessary to use a region for which substantial data was available and easy to access. The Electric Reliability Council of Texas (ERCOT) region in addition to being very large geographically, also has volumes of freely available public data which made it an excellent choice for this studying the effect of wind and solar farm diversification.

This study uses 77 meter diameter 1.5 MW GE turbines and 1.5 MW Solar arrays comprising of BP 3200SX solar panels as the smallest individual units of wind and solar resources in its wind and solar sites, respectively. The ERCOT region is studied as the interactions between the four regions: Houston, North (previously divided into the North and Northeast zones), South and West, each of which are over a 100 miles in width and length. The study uses 12 sites in each of the ERCOT regions and in addition also studies 9 sites in the area called U.S. Global which refers to regions (in New York, Washington, California, Colorado, Illinois, etc.) more than 1000 miles away from Texas and 12 sites in the U.S. Medium Range zone, referring to areas between 500 and 1000 miles away from Texas (i.e., neighboring states such as Oklahoma, Arizona, New Mexico, etc.) to examine the effect of geographic diversification in reducing the correlation between wind and solar sites within known ranges of distance. These potential wind and solar farm deployment regions are depicted on the following maps in Figures 1.1 and 1.2:

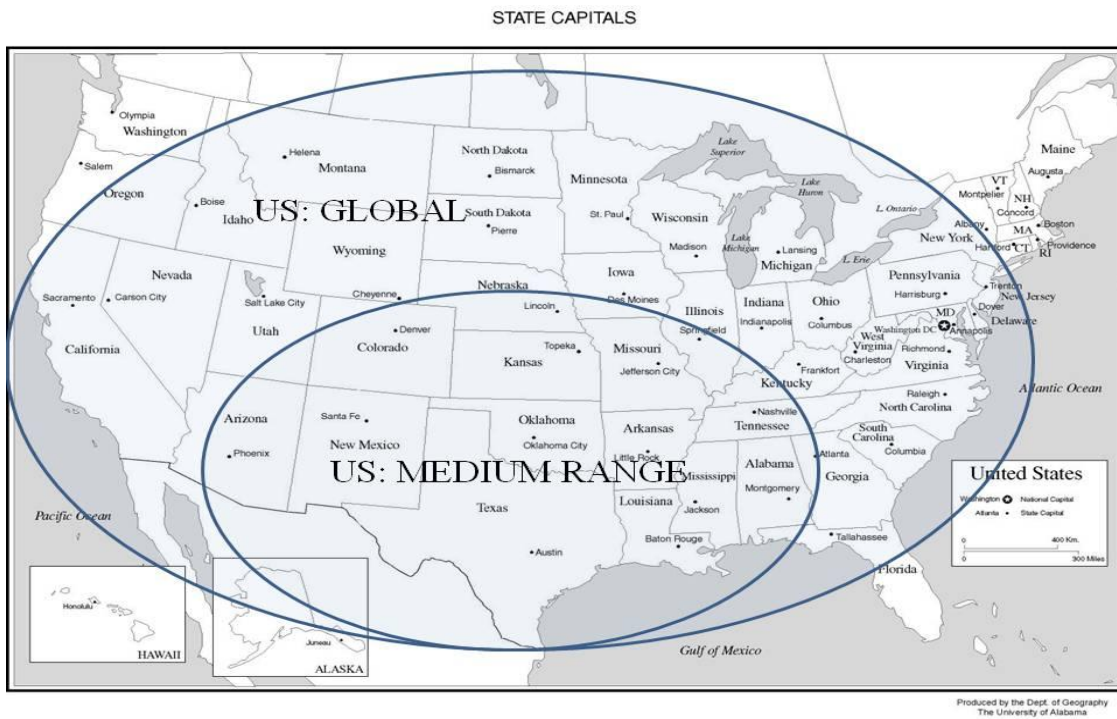


Figure 1.1 Geographic range selection



Figure 1.2 ERCOT zones

With the work conducted in this study it is hoped that further research in electric resource planning might be facilitated that uses a more holistic approach towards incorporating multiple sources of renewable energy within a region. It would be beneficial to exploit the readily available benefits in naturally occurring phenomena such as the correlation between wind and solar patterns as a function of physical distance and the load-following nature of certain resources such as solar arrays during a diurnal cycle. It is hoped that the groundwork completed here would allow more detailed studies to be conducted that encompass greater geographic areas and more forms of renewable resources.

This thesis is organized into five chapters, each detailing the research philosophy and methodology followed. Chapter 1 introduces the topic and the motivation for this study, following which in Chapter 2, a literature review discusses prior research that serves as the foundation to this thesis. The second chapter also discusses the considerations of selecting a region for this thesis and the importance placed on the ERCOT region for this thesis. A description of the wind and solar data collected and the methods used to process this data into system inputs for the optimization is also discussed. Certain assumptions regarding the correlation between wind and solar farms critical to validating the optimization method followed in this study are outlined as well. Then Chapter 3 expounds the principles of portfolio diversification and its application to wind farms. The validity of the modification of the portfolio theory for wind and solar farms siting in terms of geographic diversification is explained and the approach is discussed. The mathematics for adapting the idea of portfolio theory to wind farms is explained to provide the basics necessary for a rigorous discussion of the optimal distribution algorithm. The optimization process is then described along with the load and generation models for the ERCOT grid which are also inputs to the system. Chapter 4 explains the implementation of the optimization algorithm and the calculation of the reliability measures of system reliability such the loss of load probability (LOLP), expected unserved energy (EUE), ramping, etc., using

Monte Carlo simulations. Next the methodology employed to compare various reliability metrics obtained from various geographic configurations to explain the reliability as function of the system scale is explained. Just before the conclusion, Chapter 5 presents the obtained results for the optimization algorithm and the simulated reliability metrics, offers inferences and concludes with a short discussion of future work that could be based on this thesis. This is followed by an afterword and an appendix of the MATLAB code used in this study.

CHAPTER 2

WIND AND SOLAR DATA

2.1 Literature review

The work presented in this thesis develops on considerable literature that has been published so far on hybrid (wind and solar energy) system reliability and on the idea of applying portfolio diversification to wind farm siting. A lot of research has been conducted in the European Union (E.U.) and the United Kingdom (U.K.) to examine the effects of ‘spatial distribution’ or geographic diversification on wind system reliability. For example, Drake and Hubacek used a very literal application of the portfolio diversification theory in the U.K. to construct an efficiency frontier and graphically resolved the problem to support the claim that geographic diversification reduces the variability in the power output [1]. This thesis, in contrast, examines the claim using a robust approach using an automated algorithm for identifying the optimal geographic configuration of sites for minimizing the system variability while also measuring reliability with specialized measurements such as LOLP, EUE, etc. Another study by Roques, Hiroux and Saguan in the E.U. on cross-country portfolios inspects the effects of distributing turbines across Spain, France, Germany, Austria and Denmark to reduce variability [2]. Studies have also been performed to study broadly distributed renewable resources such as those by Atwa et al. [4]. However, studies combining various ideas, and then examining stability as a function of the geographic system scale are relatively scarce. Degeilh and Singh adapt the portfolio diversification theory to wind farm optimization as a function of the correlation between wind farm sites [5]. This thesis is an extension of their work by considering solar resources along with wind turbines as well as increasing the scale of the optimization by selecting a much larger geographic selection of deployment sites. In addition this study inspects the changes in system reliability as the scale of the system is altered and also examines the effects of the

correlation of the load with the renewable resource. Though Arkadani, Riahy and Abedi offer a method to optimize the design of a hybrid energy system, their optimization seeks to reduce the cost associated with the power production and includes storage capacity in the system [6]. This thesis shies away from the commonly used method of minimizing the cost function. We seek to isolate the effects of wind and solar farm correlation on system reliability so as to examine the independent benefits of geographic diversification and the related uncorrelatedness between distant sites. With this approach we prevent the generation costs from overshadowing the correlation-based optimization. Large scale hybrid systems have often been studied using probabilistic models successfully. This thesis follows a similar scheme, using probabilistic models as proposed by Testa, De Caro and Scimone and investigates their claim that these hybrid systems seem to be more sensitive to solar irradiation than wind speed [7]. However, the system scale we use is much larger than the low voltage, low power systems considered by them.

With respect to reliability studies, Dahman et al. solely investigate the effects of wind farm correlation on reliability on a realistic wind system in California. They propose transmission upgrades and the addition of resources to the grid [8]; this thesis in contrast seeks to investigate whether the reliability could be improved by geographic diversification and also considers solar resources. Gerard O'Connor's work is a detailed attempt at observing the effects of geographic diversification on reducing the variability in power output for an interconnected set of wind turbines [9]. He attempts to determine the optimal configuration by minimizing the cost of operating a hybrid system of various fuel types. In contrast, this thesis considers a larger selection of sites and optimizes their distribution solely on the correlations between sites, which is more representative of the effect of geographic diversification. His Texas case study is extended in this thesis by increasing the number of sites chosen and examining the effects on reliability by measuring the LOLP, EUE and ramping as opposed to observing the smoothing effects on power curves. Mabel, Raj and Fernandez conducted

a similar study in the Indian state of Tamil Nadu and used the LOLP and unserved energy as reliability metrics [10]. However, the effects on ramping were not investigated by them. This thesis also attempts to answer the question posed on the effect of very large systems on geographic diversification while also accounting for the effects of renewables on system events such as controllable load ramping.

The portfolio diversification theory can also be applied to other renewable sources such as solar energy and tidal energy which also have variations in their time series. Tina and Gagliano state that wind and solar resources are excellent components of each other and can be used in combination to improve the capacity factor of hybrid systems [11]. This thesis considers a much larger data set and also considers the correlation of the renewable resources with the load. The applicability of the portfolio theory to a larger combination of resource types remains to be validated. With the work done before by many others in this field, we are fortunate and thankful to have a strong foundation laid out for future research efforts.

“Nanos gigantum humeris insidentes”

2.2 Identifying a viable region and system for analysis

According to literature already published, the idea of portfolio diversification has been considered to reduce the variability in wind power output due to the reduction in the correlation between wind patterns. Geographic locations that are spread far apart are able to diversify the ‘risk’ of the entire wind power resource because while some sites encounter poor winds, there exists a possibility of another site encountering a state of high winds, thereby compensating them. Thus wind farm geographic diversification can be employed to smooth out the fluctuation in wind power output from uncorrelated sites. This is currently achieved through wind power variations in one part of the country canceling out variations in wind power in another part of the country as claimed

by Drake and Hubacek [1], depending on the distance between them. The reduction in variability is related to the distance between sites, which stems from the reduction in correlation as distance between sites increases. For instance, Sinden found that the hourly correlation coefficient between U.K. wind farm sites decreases to approximately 0.1 over distances in excess of 100 km [12].

Since the United States is much larger in geographic size than most countries in the E.U. it is possible to examine the true feasibility of such a diversification scheme as the region is unencumbered by political boundaries, thereby allowing us to monitor the interaction between wind resources, conventional generation and loads all within the same electricity provider's region of operation. ERCOT's relative electric power autonomy within the United States helps avoid the issues of state boundaries and load distribution boundaries between various electric providers. The absence of boundaries between ERCOT's zones along with the ease of obtaining detailed load and generation data facilitated the development of this thesis tremendously.

To provide an estimate of its great size: ERCOT manages the flow of electric power to 23 million customers in Texas - representing 85% of the state's electric load, covering almost 75% of the entire physical area of the state of Texas. The portion of the electric grid in the State of Texas that is under the administration of ERCOT is essentially unconnected to electrical grids in other states in the U.S. It is the independent system operator for the region and connects 40,500 miles of transmission lines and more than 550 generation units. By virtue of in essence being a single power system instead of as a conglomeration of independent cooperating utility companies, the collection of data and statistics is quite straightforward. In comparison, contacting a multitude of independent electric providers as is the case in other regions such as the Eastern interconnect or the Midwest would have been quite a task!

Texas is fortunate to have strong winds and solar insolation patterns for the greater part of the year and is therefore an ideal region for the deployment of a hybrid wind-solar energy resource. The state itself being the largest in the United States covers an area of about 268,820 square miles and is 773 miles wide and 790 miles long. According to Brendan Kirby in his assessment of the ERCOT CREZ study, the large geography of Texas definitely helps add diversity and reduce wind volatility as the correlation between zones will likely be near zero in the minute-to-minute regulation time frame [13]. To illustrate his point, Figure 2.1 shows that the correlation between wind plants drops dramatically with distance. The coastal area, for example, is 350 miles from Abilene and 500 miles from Floyd County and the hourly correlation between the coast and the rest of Texas is essentially zero.

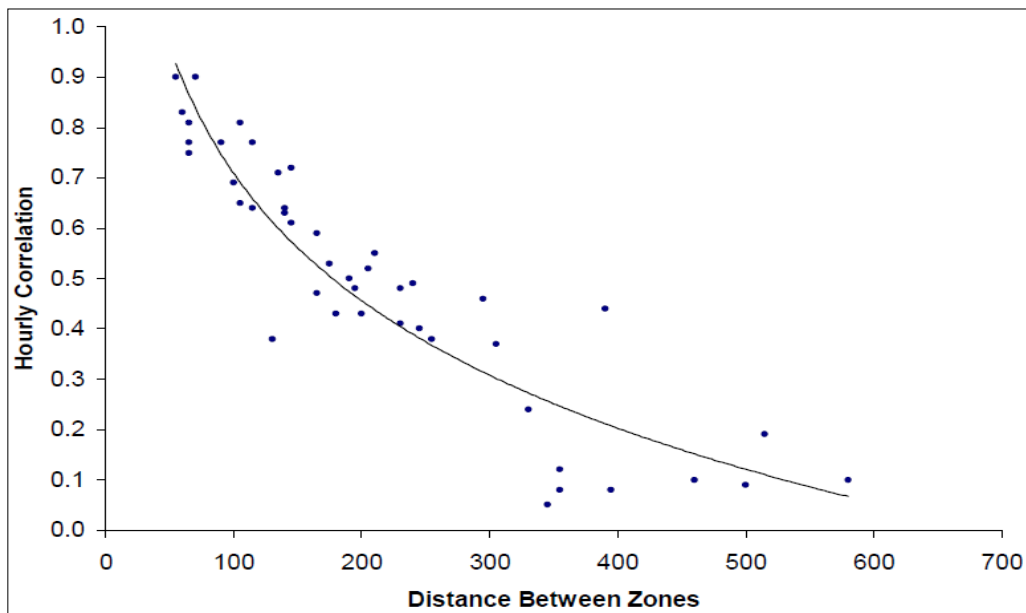


Figure 2.1 Hourly correlation between sites and distance

Figure 2.1 from Kirby's research shows that there is significant correlation between the wind farms clustered in various zones within Texas, i.e., within a range of 300-400 miles; however, the correlation reduces significantly with the increase in

distance [13]. The non-correlation between wind patterns, especially at great distances would be even greater. This reduction in correlation will thereby allow the evaluation of the effects of geographic site diversification in reducing the fluctuations in wind power output by siting wind farms in areas far from each other.

2.3 Wind speed and solar insolation measurements

For the optimization of the hybrid wind and solar resource, actual wind speed and solar insolation data was used in order to obtain a system as close to real physical conditions in the region. However this posed a challenge as the data required to be gathered for the optimization problem was gathered from multiple sources. Initially the wind and solar data was rather difficult to obtain as some of the resources were not available for public use and a large number of areas were not monitored in detail. Initially the attempt to gather time series data of wind speeds and solar insolation was also thwarted by the inability to find both measures for the same geographic location. Attempts were made to gather the data from government run as well as non-governmental weather monitoring websites. Often these internet resources for the data had incomplete time series with data measurements missing for several months in a year due to maintenance or breakdown or absence of the measuring stations. A compromise on the absolute accuracy of the wind and solar measurements was imminent; fortunately the National Solar Radiation Database (NSRDB) was found. The NSRDB is a serial collection of hourly values of actual wind speed and solar radiation (global horizontal (GHI), direct normal and diffuse horizontal) for a year. Though the NSRDB provides wind speed and solar radiation measurements over the course of a year, it is important to note that it is a careful construction of the average measurements in a typical meteorological year (TMY). The TMY data set is “composed of 12 typical meteorological months (January through December) that are concatenated essentially without modification to form a single year with a serially complete data record for primary measurements. These monthly data sets contain actual time-series meteorological measurements and modeled solar values, although some hourly records

may contain filled or interpolated data for periods when original observations are missing from the data archive” [14]. This database was the source for the wind speed and solar radiation for sites in states other than Texas.

It is critical to note that the algorithm used to produce the TMY data set assigns priority to the solar radiation elements, thus the selected months may or may not be typical for other elements such as the wind speed. Additionally, even though wind speed was used in the selection of the typical months, its relatively low weighting with respect to the other elements prevents it from being sufficiently typical for simulating wind energy conversion systems [14]. This however does not impact us adversely as we are considered about the correlation between wind patterns rather than the windspeeds.

Since it was necessary to maintain the natural correlation between wind speeds for a geographic region, special care was taken to obtain wind speed measurements for the sites selected within Texas. For each of the four zones in the ERCOT region, 12 counties were selected that had a large urban population or city in them such that the county had a significant load and thereby could be investigated for the potential of deploying additional resources locally to bolster the load demand. For the sites selected within Texas, a website maintained by the state government through the Texas Commission on Environmental Quality (TCEQ) was used to obtain wind speed measurements for various counties as measurements for a continuous string of 8760 hourly measurements in the year 2006. The TCEQ uses an algorithm called the AERMET preprocessor to record and pre-process meteorological data collected at monitoring stations in all 254 Texas counties. The AERMET is designed to be run as a three-stage process and operate on three types of data: National Weather Service (NWS) hourly surface observations, NWS twice-daily upper air soundings, and data collected from an on-site measurement program such as from an instrumented tower. The first stage extracts data and assesses data quality. The second stage merges the available data

for 24-hour periods and writes these data to an intermediate file. The third and final stage reads the merged data file and develops the necessary boundary layer parameters for dispersion calculations by AERMOD. Two files are written for AERMOD: a file of hourly boundary layer parameter estimates and a file of multiple-level observations of wind speed and direction, temperature, and standard deviation of the fluctuating components of the wind. The AERMET algorithm accounts for variables such as the Albedo, Bowen ratio and the roughness characteristics of the surface. The detailed 30-year weather records from the National Weather Service and measuring stations used to create this comprehensive database of wind measurements helped preserve trends in naturally occurring wind patterns in Texas [15]. This precaution ensured that the naturally occurring correlations between various geographic regions within Texas were retained, thereby enabling the examination of the benefits offered by employing the idea of wind farm geographic site diversification. However, the solar insolation data set for these counties was considered to be the time series for big cities and towns in these counties from the NSRDB. This is an acceptable assumption as the TMY time series is, in fact, a representation of the typical average annual solar insolation for the region in question.

Once the NSRDB and TCEQ websites were verified to be reliable sources of wind and solar data, the data collection process began with identifying regions and cities within the ERCOT regions for which both wind speed and solar insolation information was available. Sites close to large urban settlements were chosen such that load and conventional generation data would also be readily available for these regions from the ERCOT transmission database. The wind speed in m/sec was converted to miles/hr and then converted into the associated power output as described in the next section. The GHI solar insolation was used for calculating the solar array output power as described in the following section. The resulting time series of 8760 hourly observations of wind and solar power in kilowatts for the length of a entire year for each

of the four ERCOT regions as well as the U.S. Global and U.S. Medium range areas were then fed into the optimization algorithm. The optimization to attain the optimal wind and solar farm configurations and calculate the associated reliability parameters for combinations of various regions and energy sources is explained in Chapter 4.

2.4 Weibull distribution and conversion of data into system inputs

The accurate modeling of the wind speed variation of a particular geographic location is very important in estimating the power production of a wind turbine. As the Weibull distribution closely reflects the actual distribution of hourly wind speeds at many locations it is widely used for the modeling of wind speeds. The Rayleigh distribution is a special case of the Weibull distribution with a scale factor alpha of 2. The strength of wind varies from one given geographic location to another; and thus to model the amount of power a wind turbine could produce, a probability distribution function is often fit to the observed data. Thapar, Agnihotri and Sethi mention a technique where the manufacturer of the wind turbines provides a graph of the observed wind output power of a wind turbine as a function of the wind speed and this power curve is extrapolated as a linear function for the entire spectrum of required wind speeds [16]. We adapted this technique, and the wind speed data is first converted to miles/hour from m/sec and then the given manufacturer's power curve of power generated in watts for a specific windspeed (in m/sec) is interpolated linearly at increments of 0.1 miles/hour to obtain the indirect power curve of power vs. windspeed in 0.1 miles/hr increments as shown in Figures 2.2 and 2.3.

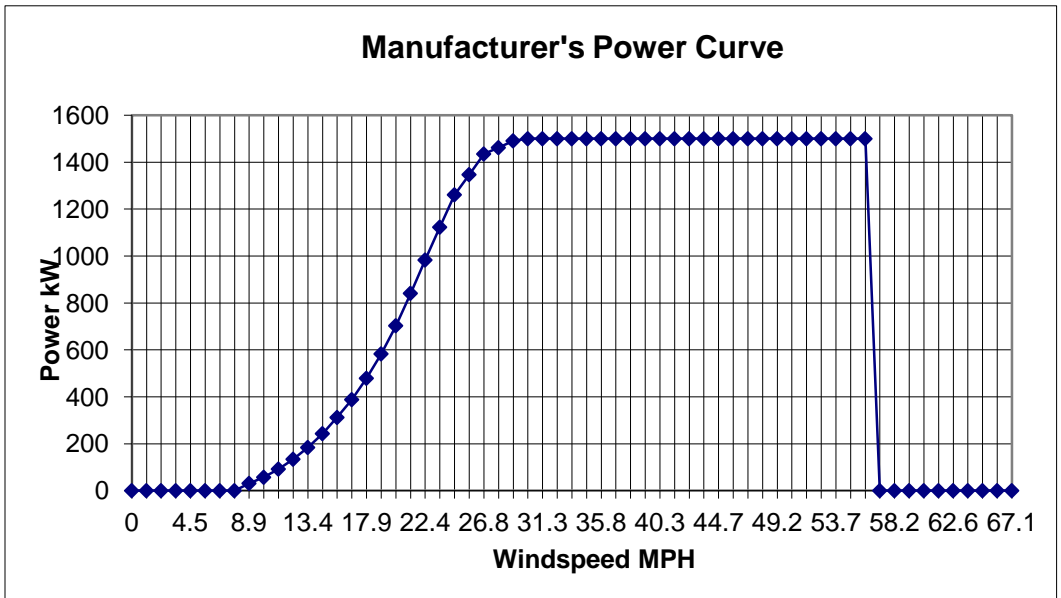


Figure 2.2 Manufacturer's power curve for GE 1.5 MW turbines

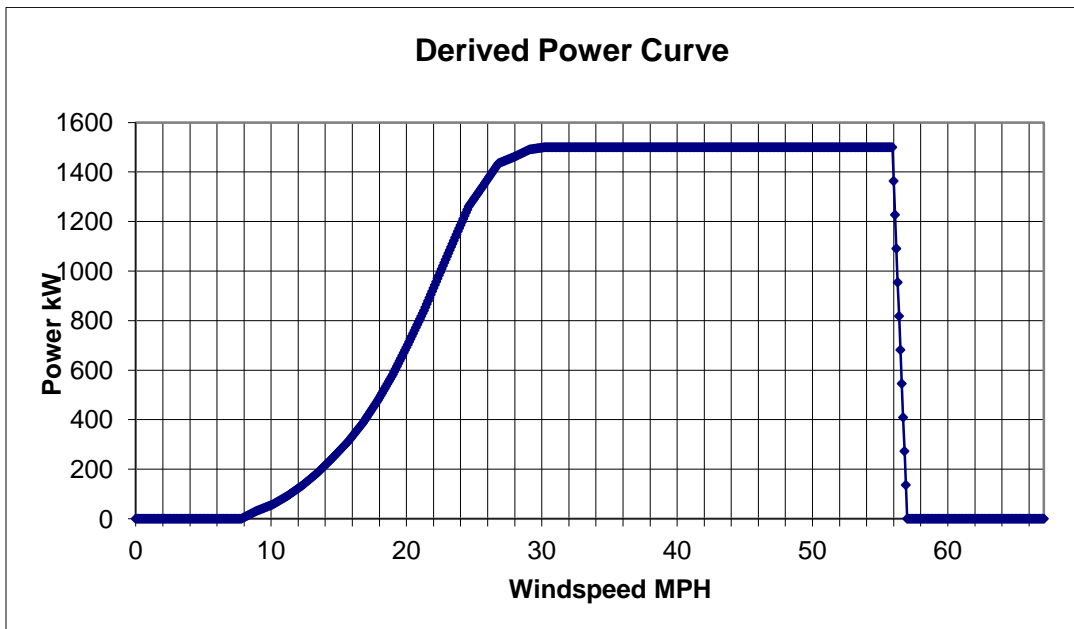


Figure 2.3 Derived power curve for GE 1.5 MW turbines

The Weibull distribution is given by the following equation which is widely considered by many to closely model the wind-speed measurements obtained in various locations.

$$f(x) = \left(\frac{b}{a}\right) \left(\frac{x}{a}\right)^{b-1} e^{-\left(\frac{x}{a}\right)^b}$$

This idea was implemented using Microsoft Excel and the indirect power curve is obtained by extrapolating the power curve obtained from the manufacturer's data for every 0.1 mph increment linearly using the slope of the values from the manufacturer's curve. Then using the 'look up' function, the power for every 0.1 mph can be obtained from the indirect power curve. Then this indirect power curve is used to create the Weibull probability distribution function (pdf). The Weibull pdf can be generated by varying the values of x over the range of required wind speeds in 0.1 miles/hr increments using the equation:

$$F(x) = 1 - e^{-\left(\frac{x}{a}\right)^b}$$

Once the Weibull fit power curve is obtained, then the power for a turbine over a certain length of time for which the windspeeds are known can be calculated simply as the integral of the product of the Weibull fit pdf and the indirect power curve:

$$P = \int_0^{\infty} f(x) \cdot \Phi(x) \cdot dx$$

Thus by first extrapolating the manufacturer's power curve to obtain the derived power curve and then by fitting it using the Weibull distribution, we can identify the power output of the turbine for an entire year. The aggregate power output for the turbine over a time series of varying windspeeds of any length is then obtained as a sum of the individual hourly power outputs for the corresponding average hourly wind speed recorded for that hour.

For the solar power calculations, the GHI solar incidence was considered as it was a more complete measure of the solar insolation intercepted by the panel at a location. The GHI values report the total amount of direct and diffuse solar radiation received on the surface during a 60 minute period [14]. The NSRDB resource had detailed GHI solar insolation measurements with respect to the angle of inclination for the geographic location. The solar panels considered for this study are the BP3200SX panels which are rated to produce 200 watts each. The values were scaled up for values of a 1.5 MW solar array by using 5000 panels per array providing a surface area of 7056 m² for each 1.5 MW solar array. The BP3200SX solar panel has been reported to have an efficiency of 13% according to Schwartz and Puffer, which is very competitively placed in the usual attainable range of 12% to 16% efficiency for commercially available solar panels [17]. According to Tina and Gagliano and Supriya and Siddharthan, the relationship between the maximum power per unit area of PV array surface available from the array and the irradiance is linear and depends on the efficiency of the solar panel being used [11, 18]. The NSRDB time series for the GHI were directly converted into the power output per solar panel according to the equation:

$$P_{pv} = A_C \cdot \eta \cdot I_\beta$$

where A_C is the array surface area in square metres, η is the efficiency of the PV array as stated by the manufacturer and I_β is the irradiance (in kW/m²) on a surface with inclination β to the surface the panels are mounted on.

2.5 Assumption of correlation between turbines in a site

For the purposes of this thesis, it is assumed that the power output from individual wind turbines in a particular site is perfectly correlated. According to Degeilh, it is necessary to make such an assumption for the sake of simplicity in obtaining wind and solar data and using it for the optimization algorithm [19]. It is nearly impossible to obtain detailed second-scale or minute-scale wind and solar time series within a particular site of interest without having monitoring equipment set up at each individual turbine. Thus, with the severity in the lack of information it is assumed that the wind and solar data obtained from the NSRDB and TCEQ for a particular site will suffice and would be assumed to be the wind and solar profile encountered by each individual resource in that site. This essentially means that we assume perfect correlation between all the individual resources within a particular site.

It is known that on a second-scale and minute-scale individual turbines have differences in their power output as there are subtle differences in the wind speed profile encountered by individual turbines within a site. It is understood that every turbine never encounters wind of the exact same windspeed or direction at a particular instant even within a small site. Literature published seems to support the claim that on longer time scales, such as that of a per-hour basis the turbines in a site seem to have a smoothing effect on the overall power output of a particular site. According to Prasad, Bansal and Sauturaga, the rapid short-term, small-scale fluctuations in the windspeed to some degree are smoothed out at the power output from a single wind turbine unit, both by the extent of the rotor and by the power control of the wind turbine [20]. Further still, according to studies conducted by Richardson and McInerney, “Relationships have been derived to estimate the spectrum of the summed power from a number of wind turbines from the power of a single wind turbine.” Because of the random effect of the wind, the resulting spectrum represents a smoother time series. Measurements were made to verify the theory, with excellent agreement. The smoothing property of multiple wind turbines is a desirable property of wind turbines” [21]. According to

them, two nearby wind turbines will produce power traces which are not statistically independent of each other. In this case, the cross terms of the cross covariance do not vanish; thus the assumptions that on a longer time scale the turbines in a site can all be expected to have the same wind power output and attributing it to their perfect correlation is justifiable.

This assumption is crucial to the optimization process used, which assumes that the aggregate power out from a site is simply the sum of the individual power outputs from the individual turbines in the site, which is approximated to be the average power from an individual turbine scaled to the number of turbines installed at the site in question as stated by:

$$P_i = p_i \times n_i, \text{ for } i \in [1 \dots n]$$

This assumption is central to the optimization process followed in this thesis and will be employed later on in this study as well.

CHAPTER 3

THE MODELED ERCOT SYSTEM

3.1 Geographic diversification and portfolio theory

The modern portfolio theory approach as used in finance to diversify the risk in investment portfolios has found its way into electric power engineering and is the basis for the idea of wind farm site diversification. According to the theory of diversification, the unsystematic risk events in a portfolio can be smoothed out such that the positive performance of some investments will neutralize the negative performance of others, although this holds true only if the securities in the portfolio are not perfectly correlated. Applying an extension of Dunlop's investment philosophy to wind farm siting [22], if the wind power outputs of different turbine sites are not correlated, then the aggregate power output of these sites can be smoothed by using them to complement each other. Therefore in instances when one farm faces a deficit in power output, it is compensated by the excess generation from another site that owing to its uncorrelatedness in wind speed profiles with the other site produces excess power at the very same instant.

The overall goal is to somehow obtain a portfolio with the minimum variance so as to have a portfolio that bears the least amount of investment risk and has a more predictable pattern of income to the investor. According to the mean variance portfolio theory (MVP), an efficient portfolio is the one which gives the highest return for a certain level of risk appropriated. The expected return of a portfolio P , $E(r_p)$ consisting of N investment assets, with each asset having: expected return r_i and standard deviation σ_i . Then consider in proportion X_i to be the weighted average of the N assets expected returns.

Then the expected return of the whole portfolio can be written as:

$$E(r_p) = \sum_{i=1}^N X_i E(r_i)$$

Then the portfolio standard deviation σ_p is described by the formula:

$$\sigma_P = \sqrt{\sum_{i=1}^N X_i^2 \sigma_i^2 + \sum_{i=1}^N \sum_{i=1, i \neq j}^N X_i X_j \rho_{ij} \sigma_i \sigma_j}$$

where ρ_{ij} represents the correlation between the returns r_i and r_j of two assets. Thus it can be observed that for the minimum value of the variance of the portfolio represented by

$$Var(P) = \sigma_P^2 = \sum_{i=1}^N X_i^2 \sigma_i^2 + \sum_{i=1}^N \sum_{j=1, i \neq j}^N X_i X_j \rho_{ij} \sigma_i \sigma_j$$

The minimum variance is obtained when the second term is minimized, which occurs when the correlation between individual assets in the portfolio is minimized. Thus it can be explained that by having individual assets that are uncorrelated, the variance, thereby the investment risk of the entire portfolio, can be reduced.

The correlation ρ_{ij} can also be represented as: $\rho_{ij} = \frac{cov(i,j)}{\sigma_i \sigma_j}$, where $Cov(X_i, X_j)$

is the covariance of random variables X_i and X_j , which leads to the relation:

$$Var(P) = Var\left(\sum_{i=1}^N X_i\right) = \sum_{i=1}^N Var(X_i) + 2 \sum_{i < j}^N Cov(X_i, X_j)$$

This idea can be extended directly to apply to wind farms with individual wind farms representing investment assets and the total ‘global’ wind resource representing the portfolio. For the purpose of this thesis we consider the total aggregate power output of the wind farm to be the sum of the individual power outputs of the wind turbines in the farm. Assuming this to be the basis of the power calculations, then the global power

output variance can be calculated as the variance of the sum of various wind farm outputs. We then have the relation for the global power output variance for N wind farms as:

$$Var(G) = Var (\sum_{i=1}^N P_i)$$

where P_i denotes the power output of a particular wind farm and G refers to the ‘global’ wind resource. Then we can obtain the relation:

$$Var(G) = Var (\sum_{i=1}^N P_i) = \sum_{i=1}^N Var(P_i) + 2 \sum_{i<j}^N Cov(P_i, P_j)$$

where again, the correlation coefficient ρ_{ij} between the two wind farms is related to their covariance $Cov(P_i, P_j)$ as:

$$\rho_{ij} = \frac{Cov(P_i, P_j)}{\sigma_i \sigma_j} = \frac{Cov(P_i, P_j)}{\sqrt{Var(P_i)Var(P_j)}}$$

By then extending this idea to n_i individual turbines within a wind farm i , we have the equation:

$$Var(G) = \sum_{i=1}^n n_i^2 Var(p_i) + 2 \sum_{i<j}^n n_i n_j Cov(p_i, p_j)$$

It is important to note that this equation is valid only when a complete correlation of wind turbines within a wind farm is assumed.

Without this, the relation between the wind farm power output and the power outputs of individual turbines, would not hold as given by:

$$P_i = p_i \times n_i, \text{ for } i \in [1 \dots n]$$

Research by Roques, Hiroux and Saguan, following this idea of reducing the global portfolio variance, has been performed in Europe to identify optimal wind power portfolios across Austria, Denmark, France, Germany and Spain [2]. In their approach they successively use two objective functions to define optimal cross-countries wind power portfolios: (i) “Optimizing wind power output” which consists in maximizing wind power production and minimizing hourly variability at all times; and (ii) “Maximizing wind power contribution to system reliability” which consists in maximizing wind power production and minimizing variability during peaking-hours. It is also their opinion that the basis of wind portfolio optimization through geographic diversification is the combination of wind production from sites with low or negative correlations. In this thesis a similar approach is employed to reduce the variability in power output by discovering a combination of sites and energy sources which have low or negative correlations amongst them.

With this understanding we will now proceed to produce an optimal configuration of turbines that minimizes the global variance of the power output of both the wind and solar resources.

3.2 Optimization problem and geographic configurations

In this study, the objective is to distribute wind and solar resources within a stipulated geographic region so as to minimize the output variability of all the renewable resources deployed within that region. The load served in that region and the conventional generation resources already available need to be accounted for. We have six possible regions, with four of them corresponding to the four ERCOT zones, all

within the geographic boundaries of Texas. The other two regions, U.S. Medium range and U.S. Global, refer to sites selected in states neighboring Texas within a range of 500-1000 miles and more than 1000 miles away from Texas state boundaries, respectively. Each of these six regions has about 9 to 12 sites.

For the optimization study the user has the option of varying the regions of deployment, load and conventional generation. The load considered for this study is the load supplied in each ERCOT zone and the generation is the conventional generation already deployed in each ERCOT zone as a combination of nuclear, coal and gas turbine power resources. The algorithm from this thesis allows the user to examine individually the effects on the load of a particular ERCOT region or in any combination of the four ERCOT regions, the same for the conventional generation zones. This allows the user to investigate the effects of loads on one region by considering the generation from elsewhere or from multiple zones including itself, allowing a comprehensive study permitting multiple combinations of conventional generation and load zones. As for the deployment zones, any combination of the six zones may be used to investigate scenarios such as if a region far away from Texas was to be the source for wind and solar power.

With the following convention being used: $\left\{ \begin{array}{l} \textit{Houston ERCOT} - 1 \\ \textit{North ERCOT} - 2 \\ \textit{South ERCOT} - 3 \\ \textit{West ERCOT} - 4 \\ \textit{US Global} - 5 \\ \textit{US Medium range} - 6 \end{array} \right.$

the user could select loads from regions 1 and 3 and conventional generation from regions 2 and 4 and deployment from 1, 2, 5 and 6. This would command the algorithm to find the optimal configuration of wind and solar resources in Houston, North, U.S. Global and U.S. Medium range zones to supply the load zones of Houston and South ERCOT in addition to the conventional generation from North and West ERCOT. Any

such combinations can be studied. For our purposes of studying the effect of geographic distance, we can see the difference in reliability by altering the deployment zones. To study the effect of distance on a short scale, we could use combinations of load, conventional generation and deployment within the ERCOT region inside Texas state boundaries. By investigating the change in the reliability metrics of the system by considering different combinations, we can seek to understand how differences in the geographic scale of the system alter the variability in the power output.

The optimization then essentially reduces to an attempt to deploy m hybrid renewable resources (a combination of wind turbines and solar arrays) over n sites. The number of turbines and solar arrays are constrained by the maximum output power per site and the total renewable power output for the system. These upper and lower bounds seeks to limit m , which is the number of resources for each site, and thereby prevents the system from overzealously deploying turbines and solar arrays in unreasonable amounts wherever it finds uncorrelated sites thereby limiting $\sum_{i=1}^n m_i * p_i$, the total renewable power output of the system. The lower bound ensures that the optimization, while configuring the distribution of solar arrays and wind turbines, remembers to fulfill the minimum required renewable power requirement to the grid. This combination of bounds ensures the judicious deployment of these resources in the sites that best lower the overall variability of the controllable load (the system load remaining after subtracting the renewable power output; this load would be supplied by conventional generation) by selecting sites that are relatively most uncorrelated.

This thesis follows a similar optimization algorithm as implemented by Yannick Degeilh where the variance of the controllable load, $Var(L-G)$, is the objective function to be minimized with respect to the wind turbine and solar array dispatch [19]. The constraints consist of the total number of wind turbines and solar arrays m to be installed and the expected renewable power output mean value EXP (expected power)

the system configuration is expected to supply. As a matter of fact, this optimization needs to be carried out many times for a range of values of EXP so as to make clear what can be expected from the optimal distribution of m wind turbines and solar arrays.

The optimization problem of minimizing the variance of the controllable load (difference of the load and the total renewable power output), such that the following conditions are met, can be represented as:

$$EXP = \text{Selected value of renewable output,}$$

such that:

$$\begin{aligned} \text{Installed wind power capacity per site} &\leq \text{Max. installed wind power per site} \\ \text{Installed solar power capacity per site} &\leq \text{Max. installed solar power per site} \end{aligned}$$

This is mathematically represented as

$$\text{Min Var} (L - G) = \text{Min Var} \left(L - \sum P_i^W + \sum P_i^S \right),$$

where $G = \sum P_i^W + \sum P_i^S$ which upon expansion is the same as

$$\begin{aligned} \text{Min. Var}(L - G) &= \sum_{i=1}^n n_{iw}^2 \text{Var}(p_{iw}) + 2 \sum_{i<j}^n n_{iw} n_{jw} \text{Cov}(p_{iw}, p_{jw}) + \\ &\quad \sum_{i=1}^n n_{is}^2 \text{Var}(p_{is}) + 2 \sum_{i<j}^n n_{is} n_{js} \text{Cov}(p_{is}, p_{js}) + \\ &\quad 2 \sum_{i<j}^n n_{iw} n_{js} \text{Cov}(p_{iw}, p_{js}) + \text{Var}(L) - \\ &\quad 2 \sum_i^n n_{iw} \text{Cov}(L, p_{iw}) - 2 \sum_i^n n_{is} \text{Cov}(L, p_{is}) \end{aligned}$$

subject to the constraints,

$$\sum_{i=1}^n (n_{iw}) * 1.5 \text{ MW} \leq \text{Max. installed wind output per site}$$

$$\sum_{i=1}^n (n_{is}) * 1.5 \text{ MW} \leq \text{Max. installed solar output per site}$$

$$\sum_{i=1}^n n_i E(p_i^W) + n_i E(p_i^S) = EXP$$

$$i \in [1 \dots n], n_i \geq 0$$

By linearity of the “expected value” operator $E ()$, we have $\sum_{i=1}^n n_i E(p_i) = E(G)$.

The optimization problem therefore deals with quadratic function subject to the constraints mentioned above. As indicated by Degeilh, the solution of this problem is convex and the local minimizer of the problem is also a global minimum [19]. The solution to this optimization was obtained using the Quadprog() function available in MATLAB, which uses an active set method which searches for a stationary point along a set of active constraints which are redefined until an optimum is obtained [23]. The algorithm seeks to minimize the following,

$$\min_x \frac{1}{2} x^T H x + c^T x$$

with the constraints

$$\begin{cases} Ax \leq B \\ A_m x = B_{eq} \\ x_m \leq x \leq x_M \end{cases}$$

which Quadprog () solves as

$$[x, f_{opt}, flag, c] = \text{quadprog}(H, f, A, B, A_{eq}, B_{eq}, x_m, x_M, x_0)$$

with H as the Hessian matrix of the objective function, which in our case is:

$$x = [n_1^W, \dots, n_n^W, n_1^S, \dots, n_m^S]$$

$$H = \begin{bmatrix} \text{Var}(p_1^W) & \text{Cov}(p_1^W, p_2^W) & \text{Cov}(p_1^W, p_3^W) & \dots & \text{Cov}(p_1^W, p_n^W) & \text{Cov}(p_1^W, p_1^S) & \dots & \text{Cov}(p_1^W, p_m^S) \\ \text{Cov}(p_1^W, p_2^W) & \text{Var}(p_2^W) & \text{Cov}(p_2^W, p_3^W) & \dots & \text{Cov}(p_2^W, p_n^W) & \text{Cov}(p_2^W, p_1^S) & \dots & \text{Cov}(p_2^W, p_m^S) \\ \vdots & \vdots & \vdots & \text{Var}(p_n^W) & \ddots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \text{Var}(p_1^S) & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \text{Var}(p_m^S) \end{bmatrix}$$

where, H the Hessian matrix is composed of the variances of individual wind and solar sites and the covariances between these wind and solar sites amongst themselves as depicted above. Mathworks, the creators of MATLAB, states that the solution procedure involves two phases. The first phase involves the calculation of a feasible point (if one exists). The second phase involves the generation of an iterative sequence of feasible points that converge to the solution [23]. The solution of the optimization is then returned in x , the optimal objective function is returned in f_{opt} and if the optimization is successful, then $flag$ is zero.

The number of samples affects the convergence of the solution of the algorithm. We found that for 150 samples or iterations of the algorithm it converges to a steady value; however, for a 100 samples, the value obtained is not far off from the steady value. For samples numbering less than 50, the convergence of the values calculated by the algorithm is not absolute and large fluctuations are observed in the values.

Thus for this study we always ran the algorithm for 100 samples to guarantee an accurate level of convergence for the solution. For numbers of samples above 100, the algorithm faced hardware problems such as running out of memory. A similar problem was encountered when we included more than one region of conventional generation because the algorithm by using the common random numbers method generates very large matrices that caused the script to abort. In the best interest of processing time and with the hardware constraints for this study, 100 samples were always considered and

only the conventional generation from ERCOT Houston (region 1) was used. However, the total system was considered as the aggregate demand of all four regions of ERCOT.

3.3 ERCOT load and conventional generation sources

For this study, data on the load and conventional generation presently deployed in the ERCOT region in Texas was gathered and used along with the wind and solar data time series to develop the optimal wind and solar farm siting configuration. The ERCOT website offers free public information about the generation schedule and load from the zones it covers. From the hourly load and generation schedules, we were able to identify peak load patterns and model the controllable load, i.e., these loads are those that are served by the conventional generation and can be regulated as per the signals from the grid control units based on a real-time set point received from an outside source such as the automatic generation control (AGC) as an immediate response to frequency deviations on the ERCOT grid. The optimization seeks to identify the optimal configuration through a combination of wind and solar resources to minimize the variability of the controllable load.

From the ERCOT website hourly load data for all four ERCOT zones were readily available for multiple years. Then to identify the conventional generation already deployed in the ERCOT region, a generation schedule was obtained which listed the conventional generation units and their type, i.e., gas, hydro, nuclear, biomass, etc., for every region of the ERCOT grid.

Once the list of individual generation units for each ERCOT region and their fuel type were known, the generation capacity was modeled in MATLAB. Olsina and Larisson conducted studies on hybrid system models consisting of diesel generators and wind turbines, modeling these resources as a multi-state Markov model with either three or

four states [24]. Similarly in this thesis, all generation resources are considered to have either fully functional (operating), failure (shut down) or forced-out states (both when the unit is required and when the unit is not required, resulting in two forced-out states). Then using data available on the mean time to failure (MTTF) which is the average amount of time the unit is ‘up’ and active and the mean time to repair (MTTR) which is the average time required for the unit’s repair, the typical forced outage rates (FOR) for the type of generator by fuel type, the failure rate (λ) and the repair rate (μ) were calculated. Po Hu describes the calculation of the failure rate and repair rate as follows [25]:

$$\lambda = \frac{1}{\text{MTTF}} = \frac{\text{number of failures of a component in the given period of time}}{\text{total period of time the component was operating}}$$

$$\mu = \frac{1}{\text{MTTR}} = \frac{\text{number of repairs of a component in the given period of time}}{\text{total time the component was being repaired}}$$

However the probability of finding a unit on forced outage (when the unit is taken offline on purpose, for purposes of maintenance, etc.) is calculated as:

$$\text{FOR} = \frac{\lambda}{\lambda + \mu}$$

With these values of FOR, we obtain the generation schedules for every hour of the year by sampling their steady state distribution for the generation units of the ERCOT regions depending on their fuel type. The FOR values used for wind, solar and diesel generation units were obtained from Taljan, Pantoš and Gubina’s research on system reliability by changing the penetrations of various fuel types [26].

We combine the results of the optimization, i.e., the number of wind turbines and solar arrays with the statistical power outputs of a single wind and solar resource (p_i^W, p_i^S) extracted from the NSRDB and TCEQ data to obtain the artificial history of

one year of renewable power output. The algorithm also allows the user to specify a desired level of reserve levels for these conventional generation resources. If desired, by using the “lhsdesign” function which generates a latin hypercube of 8760 independent samples for each conventional generating unit (depending on the FOR), the statistical accuracy of the simulation can be improved.

CHAPTER 4

ESTIMATION OF ELECTRIC RELIABILITY MEASUREMENTS

4.1 Generation of the simulated optimal geographic configuration

The MATLAB program used in this study is organized into two modules. The first module, “texaswindsolar,” is the optimization algorithm and is used to derive the optimal configuration for the combination of deployment regions, load and generation zones selected. First the lower and upper bounds of acceptable expected power are defined for the system. This is done to ensure that the system always produces power within an acceptable range for normal operation, thereby preventing shortages or overloads. Then the maximum power output limits are set for the wind turbines and solar arrays. The biggest wind farms today range in the 800 MW to 1 GW range and according to a report published by Foster-Wheeler, on average, large solar farms are around 300 MW [27]. The algorithm allows the user to choose whether or not to consider correlation of the load patterns with the wind and solar profiles. By default this option allows us to investigate how the system performs when the resource is correlated with the load. It is also possible to choose only the peak load if desired for a quick assessment of the system’s robustness at full capacity. The expected value, covariance and correlation matrices are calculated for the all the wind and solar time series data. The objective function is built and the Hessian matrix is constructed with the calculated values of the variances and covariances from the covariance matrix. This provides the optimal configuration of renewable resources which are the least correlated with each other to minimize the variance for the aggregate renewable system output. It is interesting to note that if the correlation of the renewable resources with the load is

considered, the system then minimizes the variance of the controllable load which tends to favor renewable resources that are positively correlated with the load. Then for various values of aggregate system power output within the range of normal operation, as defined earlier, the systems expected value constraints are specified for the configured resources.

In these experiments, the most reliable configuration of wind farms and solar arrays was deployed in all six deployment zones. On careful inspection of the deployment of renewable resources, it was noticed that very little solar capacity was utilized with the system deploying just three arrays in the U.S. Medium range region. This is understandable as explained by Cartwright and Apt; the solar arrays even over a large geographic distribution (within 280 kilometres) are still highly correlated [28], thus the system would try to avoid the deployment of these highly correlated resources as far as possible. The wind farm deployment for a total of 23,656.5 MW was as seen in Table 4.1.

Table 4.1 Installed wind capacity per deployment zone

Deployment Zone	Number of 1.5 MW turbines deployed	Capacity deployed (MW)
1	0	0
2	1221	1831.5
3	1529	2239.5
4	720	1080
5	5306	7959
6	6995	10492.5

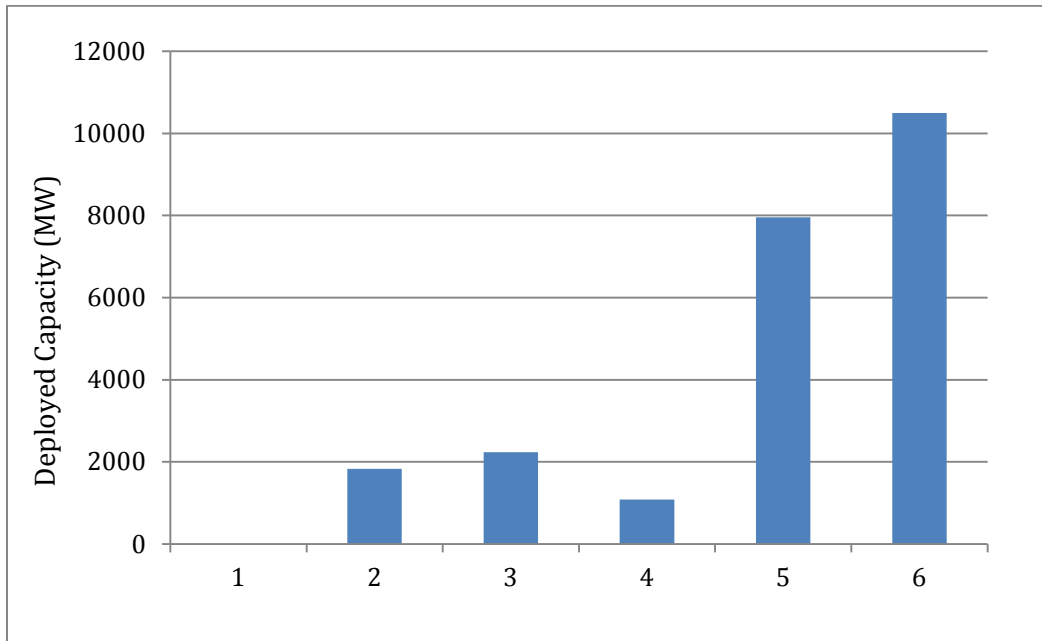


Figure 4.1 Distribution of wind resources by zone

It is clearly visible from Figure 4.1 that the system prefers to deploy the most capacity in the U.S. Medium range and U.S. Global zones to exploit geographic diversification. The system has a natural tendency to spread the wind farms as far apart as possible to minimize the variance of the entire system and more importantly does not even deploy resources in the same region as the conventional generation.

4.2 Reliability measures: Calculation of LOLP, EUE and ramping

Once the optimal configuration of the wind and solar resources is determined for the combination of deployment, load and generation areas, the second module of the MATLAB code, 'simulation' is used to determine the reliability indices and study the effect on system ramping events. As described in the previous section, in the simulation the load and generation schedules (of both renewable and conventional generations sources) are built for the selected regions using the latin hypercube function. As described earlier, the Monte Carlo simulation creates an artificial operational history of

8760 hours during which the conventional generation, depending on the FOR assumed for the particular fuel type, is declared ‘out of order’. Then the total aggregated system output (conventional generation + expected renewable power output) minus the aggregated load is calculated. This allows us to inspect if a loss of load event occurs when the system is unable to supply the load demand at any instant during the simulation for different levels of expected renewable output within the range of normal operation. With this calculation being performed for every hour of operation, the reliability indices, i.e., the loss of load probability (LOLP), the expected unserved energy (EUE) and the ramping events are calculated over the entire year, i.e., for 8760 hours. The Monte Carlo simulation is run for a selected sample size large enough for the results to converge. We shall compare the coefficient of the variation of the LOLP with a tolerance of 5% as an indicator of the convergence of the simulation as described by Degeilh. The number of simulations that need to be run are determined by the tolerance as described below:

$$\frac{\sigma_{LOLP}}{\widehat{LOLP}} \leq tol(LOLP)$$

where σ_{LOLP} is the standard deviation of the LOLP estimate, \widehat{LOLP} is the LOLP estimate and $tol(LOLP)$ is the tolerance limit of the LOLP, considered as at least 5%. In order to be as reliable as possible the simulation attempts to get the standard deviation of the *LOLP* estimate (the true value being unavailable) to be *tol* % inferior to the estimate of its expected value, which basically means that once this criterion is satisfied, the value of the reliability index *LOLP* is very unlikely to vary significantly with further rounds of simulation. A test was run to observe the change in the criterion with increasing rounds of simulation. The results were plotted in Figure 4.2.

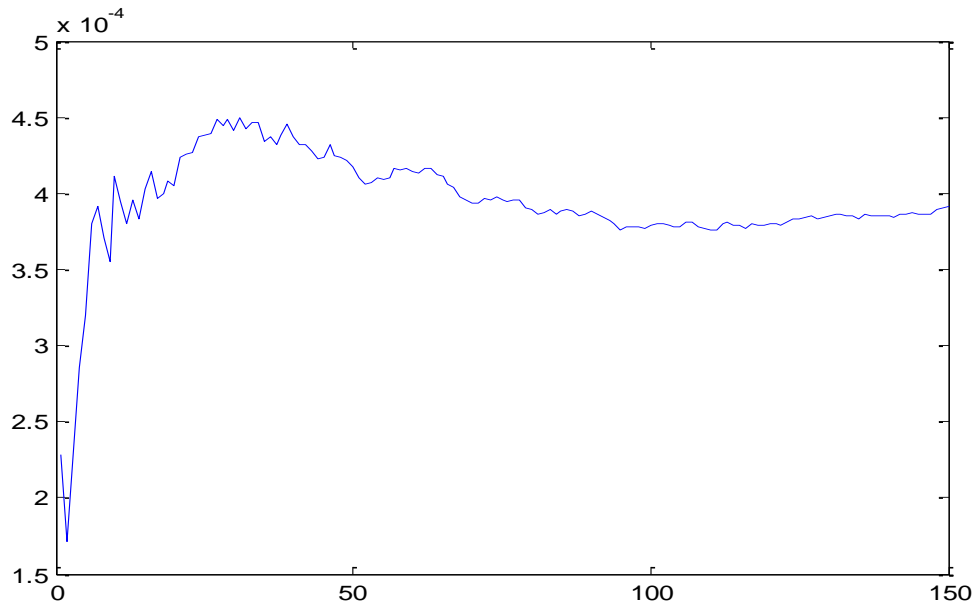


Figure 4.2. LOLP estimate over an increasing number of samples

For example, Figure 4.2 shows a trial of measuring the LOLP for an optimal configuration within the ERCOT region over 150 samples. We see that the value of the LOLP converges to a value with time. Initially there is a large fluctuation which settles down to a more stable estimate after about 90 samples; thus by considering about 100 to 150 samples, the LOLP estimate is closer to the actual value.

From Figure 4.3 it can be inferred that for more than 100 samples of simulations, the improvement in the criterion is not markedly increased. For around a 100 samples the criterion converges to a value of about 5% and is deemed sufficient for this study in the interest of time and computational resources.

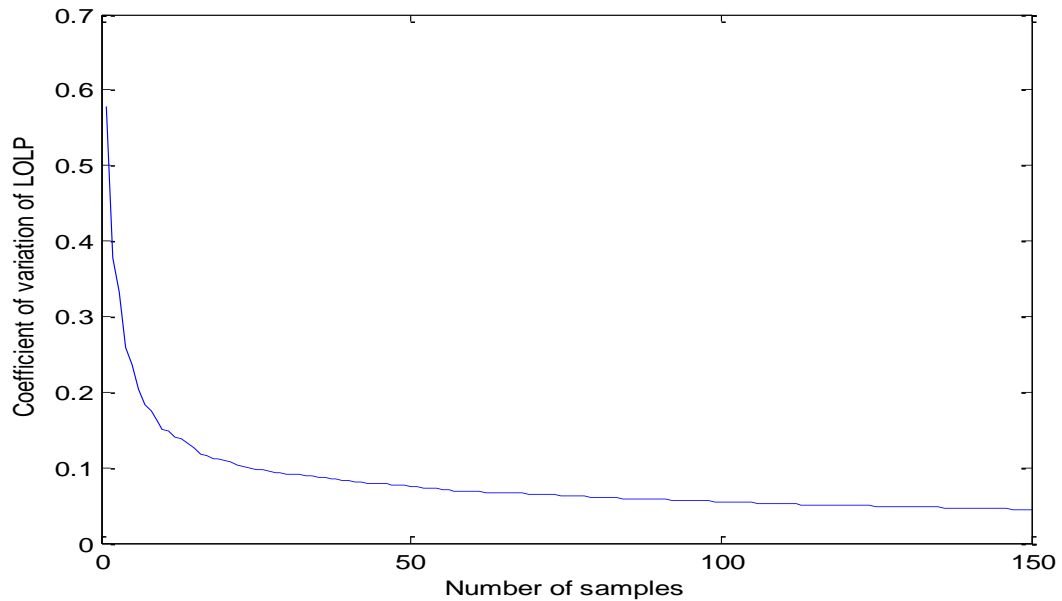


Figure 4.3 Coefficient of variation of LOLP over a number of samples

The simulation also enables the user to study the system reliability, independent of the conventional generation. It might be argued that the reliability of the renewable resources system is tied to the reliability of the conventional generation by virtue of different conventional generation units having different forced outage rates. However, if we could study the reliability of the renewable resources in a scenario of randomly organized conventional generation, it offers a more even level of comparison. We avoid comparing combinations of regions with unequal distribution of forced outage for conventional generation, i.e., one region might have more coal and nuclear plants than another which has a lot of hydro power. This is achieved by using the principle of ‘common random numbers’ to schedule the forced outage of the generation resources. Wright and Ramsay state that the technique of using common random numbers in reducing the variance is very common and well recommended [29]. The idea is illustrated by Martin Haugh, as follows:

Let $(X_1^m \dots X_r^m \text{ and } X_1^n \dots X_r^n)$ be the set of r samples that we use to estimate $E(X^m)$ and $E(X^n)$, respectively. Then $Z_i := X_i^m - X_i^n, i = 1, \dots, r$. Then if all Z_i 's are independently and identically distributed random (IID) variables, then for, $\varphi = \frac{\sum_{i=1}^r Z_i}{r}$, the variance is given by [30]:

$$\text{Var}(\varphi) = \frac{\text{Var}(X_i^m) + \text{Var}(X_i^n) - 2 \text{Cov}(X_i^m, X_i^n)}{r}$$

So to reduce the term $\text{Var}(\varphi)$, we would like to make $\text{Cov}(X_i^m, X_i^n)$ as large as possible. We can generally achieve this by using the same set of random numbers to generate X_i^m, X_i^n , which are highly positively correlated. This is essentially what is done in generating the FOR schedule of the conventional generation when the common random number function is invoked in the MATLAB program. Milligan and Porter offer insight to the process in great detail about the current techniques used to assess the capacity value of wind and monitor the reliability of the grid with the introduction of wind resources [31]. The nation-wide trend appears to be the use of probabilistic metrics, which typically use Monte Carlo simulation techniques to simulate future system performance under various uncertain variables (such as load variations and thermal resource availability). It is standard procedure for utilities to study LOLP and EUE in great detail, often performing it with an hourly level of granularity over an entire year. The number of simulations typically run is in the neighborhood of hundreds of simulations for each of the 8760 hours in the year. We follow a similar method in this thesis and follow the convention introduced in this paper, though not on such large values of iterations. In this thesis the convention is as follows.

If $(L - \text{Conventional generation} - (\sum P_i^W + \sum P_i^S)) > 0$, where $G = \sum P_i^W + \sum P_i^S$ represents the total expected renewable system capacity, a 'loss of load event' occurs and the EUE is the value of that difference.

Loss of load probability (LOLP) measures the probability that at least one shortfall event will occur over the time period being evaluated. By definition, since a probability must be greater than or equal to zero and less than or equal to one, LOLP is calculated as the number of simulations in which a shortfall occurs divided by the total number of simulations. Thus the LOLP is calculated as the number of instances where the system is unable to meet the system load. In this thesis we calculate the mean LOLP as the average of the LOLP values obtained for every instance of the Monte Carlo simulation as described by:

$$LOLP = \frac{\sum_{i=1}^N \sum_{j=1}^{8760} S_e}{N * 8760}$$

where S_e = every hour in a year, total conventional generation = $\sum P_i^G$ and total wind and solar (renewable) generation = $\sum P_i^G + \sum P_i^S$. $S_e = 1$ if $(L - \sum P_i^G - (\sum P_i^W + \sum P_i^S)) > 0$, otherwise $S_e = 0$ and N = the number of iterations of Monte Carlo simulations run for that year.

As the LOLP also provides no information regarding duration or magnitude of resource shortfalls, we consider it in conjunction with the expected unserved energy (EUE) in units of megawatt-hours (MWh,) which measures the expected amount of energy not being served per hour. The EUE is calculated by adding up all of the unserved energy over every simulation and dividing by the total number of hours simulated, obtaining the mean

$$EUE = \frac{\sum_{i=1}^N \sum_{j=1}^{8760} E_h}{N * 8760}$$

where, EUE = expected unserved energy (in MWh), E_h = the amount of unserved energy for this hour (in MWh), and N = the total number of hours simulated in the Monte Carlo study.

EUE is also typically calculated over one year using hourly level of granularity and provides some indication of the magnitude of shortfalls but only in aggregate. It does not reflect the frequency, duration or magnitude of individual shortfall events. Thus it only complements the LOLP but cannot be considered a comprehensive measure in itself. Thus by considering both the LOLP and the EUE as complementary measures of reliability, we can achieve a more comprehensive measure of the system reliability. With this assumption, all the studies conducted in this research carefully monitor the LOLP and EUE to derive inferences.

The ramping is measured as the amount of change in generation (increase or decrease) in MWh from one hour to the next during system operation for the 8760 hours in a year. In this thesis we investigate the ramping patterns of both the renewable resources as well as the controllable load. The expected ramping and its standard deviation are calculated for the period of one year of system operation.

$$Ramping = \frac{\sum_{i=1}^N \sum_{j=1}^{8760} (T_j - T_{j-1})}{N * 8760}$$

where *ramping* = change in generation (MWh), T_j = the difference between the amount of generation for an hour and the preceding hour (in MWh), and N = the total number of hours simulated in the Monte Carlo study.

4.3 Comparison of reliability measures for various geographic configurations

We have now looked at the various methods followed to build models to represent the load, conventional generation and renewable resources for the system used in this thesis. We will now look at the methodology used to assess the system reliability as a function of the geographic system scale by varying the deployment regions for renewable resources. We will examine the changes in composition of the optimal configurations and reliability metrics as a function of the change in the scale of the system.

First, we shall study the effect of the system scale. We shall obtain the optimal configuration for full load and full conventional generation conditions by incrementally deploying resources in ERCOT regions independently, then we shall consider the U.S. Medium range zone and finally include the U.S. Global zone as well. This will help us understand the preferences of the system in siting renewable resources and examine whether it prefers more distributed resources or a clustered siting of wind and solar resources. We shall also study the differences in correlations between sites that are close together vs. those that are far away and observe how these parameters influence the system's choice in site selection of renewable resources.

It was interesting to note that as long as the contribution from the renewable resources was significant, the LOLP and EUE were improved and the system variability was better than that of the system without renewable resources. However, when the renewable resources were low enough to be overshadowed by the conventional generation, the LOLP and EUE of the system were completely dependent on the conventional generation. This was expected and validates that the optimization performs as designed.

Following this, the second study aimed at establishing the effect on system reliability as a contribution of the renewable resources. The system reliability was first monitored for a range of reserves from 0% to 20% without any renewable resources, then the same study was run with the implantation of renewable resources totaling 20% of the available conventional generation to the system. We decided to include renewable resource capacity equaling 20% of the total conventional generation, as it was a likely scenario along the lines of the 20% national RPS goal [32] to ensure that by 2020, at least 20% of the load served in the U.S. would be from renewable energy.

The third study focused on the system response to correlations between renewable resources and the load, examining the system preference for renewable resources that follow load patterns. We implemented this by allowing the optimization to consider the correlation of the resources with the load. It can be assumed that solar resources follow a diurnal pattern quite similar to the daily load. This would help us infer the system preference of including either wind or solar resources or a combination in the optimal configurations. This premise is based primarily on the fact that humans for the most part also follow a diurnal sleep cycle and consume more energy during the day and reduce the electric demand by not using as many appliances, lights, etc., at night. This might be skewed in areas such as industrial parks, etc., where the load is relatively invariant over the course of 24 hours.

Finally, we shall observe the changes in LOLP, EUE and ramping for various combinations of deployment regions by first considering deployment zones that are close together and then spreading them apart.

CHAPTER 5

CONCLUSION

5.1 Results and inferences

In the first section of the study, the impact of the system scale was investigated by examining optimal configurations produced by the algorithm for different combinations of deployment zones. We aimed to observe the patterns in renewable resource siting that the optimization followed in order to minimize the variance of the entire system. The system was configured by progressively reducing the geographic areas of deployment that were available for placing the wind turbines and solar arrays. We began this study by first permitting the system to site resources in all six deployment zones and then running the optimizations five more times while progressively removing one deployment zone at a time from the possible siting options. Therefore after five iterations the system would only be able to site the renewable in one deployment zone. For the study the entire load of the system and conventional generation from deployment zone 1 was included. The results are tabulated in Table 5.1.

Table 5.1. LOLP and EUE as a function of the system scale

Deployment Zones	LOLP	EUE
[6 5 4 3 2 1]	0.00088	0.2856
[5 4 3 2 1]	0.000905	0.2931
[4 3 2 1]	0.001	0.3254
[3 2 1]	0.001	0.34
[2 1]	0.0011	0.3676
[1]	0.0016	0.5443

The LOLP and EUE were also recorded to observe the improvement or deterioration of reliability as the geographic area of deployment was progressively reduced. As the area was reduced in size, the system would be forced to cluster the sites closer, thereby reducing the correlation between sites only to the extent possible. Thus as the geographic area of deployment was progressively reduced, the benefits related to geographic diversification would also reduce. We expected a deterioration in reliability as the deployment area was reduced in size, and this was in fact observed. For true geographic diversification benefits to be experienced, the deployment region needed to be spread out to include site in the U.S. Medium range region and was improved when the U.S. Global region was also included. Thus from this study it was evident that for viable benefits of diversification, the distances between sites on average need to be greater than 500 miles.

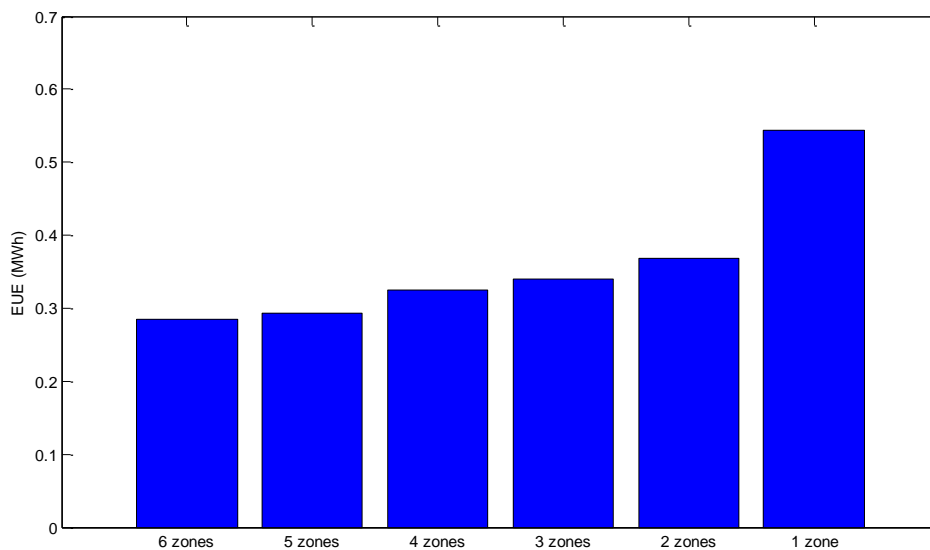


Figure 5.1 EUE as a function of system scale

From Figure 5.1 we see that the improvement in reliability does increase as the geographic scale of the system is increased; however, the returns from the system scale

show diminishing marginal returns after about 500 miles. It can be seen that the most improvement is seen when at least two zones in the ERCOT region are used for the deployment of the renewable resources. A slight improvement is gained by extending the scale further to include all four ERCOT zones. Then further improvement is seen when the system scale is extended to include the U.S. Medium range territory as well. However, when the system is dispersed over the U.S. Global region, the improvement in reliability is slight, almost negligible. The increased costs of transmission over such great distances render it uneconomical. The geographic diversification can thus be assumed to be optimized over a geographic scale of about 500 miles, with further increases in scale having insignificant improvements in system reliability.

Once the system scale was established, we were then engaged in assessing the effect of renewables on system reliability. For this test, the system's LOLP and EUE was calculated for varying levels of reserve capacity assigned as a percentage of the conventional generation. Thus as the reserve levels were decreased by 2% for each iteration of the study, the system became less reliable as the probability of the demand exceeding the conventional generation increased. We then performed the same procedure but with the inclusion of renewable reserves totaling an additional 20% of the conventional generation. We chose this amount as the maximum possible value of renewable energy in accordance with a national renewable portfolio standard (RPS) that by 2020, renewables would supply 20% of the total demand in the U.S. With this procedure we attempted to see if the inclusion of renewable energy sources into the system generation [32] while still limiting the aggregate generation capacity of the system improved the system reliability in terms of the EUE and the LOLP. The following observations were made, as tabulated in Table 5.2.

Table 5.2 LOLP and EUE as a function of reserves and renewables

Reserve Level (% Conv. gen.)	With no renewable capacity		With 20% exp. renewable capacity	
	LOLP	EUE	LOLP	EUE
20	0.75351×10^{-4}	0.0209	0	0
18	1.5527×10^{-4}	0.0413	0	0
16	2.5688×10^{-4}	0.0779	0	0
14	4.6293×10^{-4}	0.1472	0	0
12	8.3457×10^{-4}	0.2740	0.0034256×10^{-4}	0.00024187
10	15×10^{-4}	0.5097	0.10275×10^{-4}	0.0016
8	27×10^{-4}	0.9453	0.2055×10^{-4}	0.0045
6	46×10^{-4}	1.7236	0.35392×10^{-4}	0.0094
4	74×10^{-4}	3.0680	0.84885×10^{-4}	0.0213
2	115×10^{-4}	5.2640	1.6783×10^{-4}	0.0479
0	172×10^{-4}	8.7625	3.2538×10^{-4}	0.1011

The expected value of the conventional generation for the system was found to be (13,315 MW) a little above 13 GW, thus 2.6 GW or approximately 20% additional capacity in the form of expected renewable energy was introduced in the second part of the experiment.

Due to the low capacity factors of renewable resources, it takes about 24,290 MW or about 24 GW or about 10 times as much installed renewable capacity to supply an expected renewable capacity of 2.6 GW.

Table 5.2 shows that as the level of reserves is lowered, the LOLP and the EUE increase, thus; the probability of the system being unable to satisfy the load is raised. However, it is remarkable how 20% of added renewable capacity in addition is able to

offer as much as a two orders of magnitude of improvement in the LOLP, and four orders of magnitude of improvement with the system loads and conventional generation being unchanged. Thus if more units are introduced, the system would become more stable, as is expected. This tremendous improvement in the system reliability attests to the benefits of introducing additional renewable generation into the system while the grid continues to operate on its current resources.

It is interesting to note that in the second scenario, with 4% conventional reserves and 2600 MW of expected renewable capacity, grants about the same stability as having 20% conventional generation. Thus it could be estimated that to replace about 16% of conventional generation it would take about 20% or more of renewable capacity.

Further studies would be required to see the effects on reliability if the renewable resources were to instead replace conventional generation in these cases. Such a scenario is reached in the last trial of the experiment with no reserve capacity in conventional generation but 20% of additional expected renewable energy instead, which in essence emulates the scenario where the system reserves are entirely made up of renewables. In this case the system has an LOLP and EUE as the scenario where there is a reserve capacity equivalent to somewhere between 14% and 16% of conventional generation. This proves that it would take an extremely high amount of installed renewable capacity to slowly replace the conventional generation reserves. This would require an added capacity of renewable energy orders of magnitude higher than conventional energy related to the ratio of capacity factor of renewable sources to conventional sources.

We also wished to assess the effect of the correlation of the renewable resource with the load. It could be expected that if the load and renewable resource are well correlated, then both the generation and load follow similar patterns of peaks and troughs. This potentially presents an efficient, naturally occurring method of reducing the mismatch between loads and generation. When the algorithm considered load correlations with the renewable resource time series, it was noticed that there were almost no wind resources and the optimal system configuration consisted almost entirely of solar arrays. This is expected as the solar output follows a diurnal cycle with production during the day as does the load which in most cases peaks during the daytime. We performed the experiment by comparing the system configurations for when it was concentrated within Texas against when it was spread out over the U.S. Global and U.S. Medium range deployment zones. The expected renewable power output was maintained at 2600 MW, about 20% of the conventional generation. Since it was also found that the configuration with the largest deployment zone was the most stable, we selected it as our benchmark and compared it to various cases of a smaller system of the same renewable resource capacity but with an additional 5%, 10% and 20% of conventional generation reserve capacities. This permits us to compare the results of load correlation on system reliability for a large system as well as for a small system in terms of geographic diversification. The LOLP and EUE were measured and it is evident that the larger system is more stable than the smaller system for the base case with no reserve capacities. To examine the potential improvement in the smaller system, when the reserve capacity of conventional generation of the zone is increased, the reliability is also markedly increased as would be expected. Table 5.3 illustrates this result.

Table 5.3 Effects of load correlation on LOLP as a function of system scale

Deployment Zones	With load correlation		Without load correlation	
	LOLP	EUE	LOLP	EUE
[1 2 3 4 5 6]				
No reserves	3.2538×10^{-4}	0.1011	0.22834×10^{-4}	0.0060
[1 2 3 4]				
20% reserves	0	0	0	0
10% reserves	0.1712×10^{-4}	0.0031	0.01417×10^{-4}	0.0001767
5% reserves	0.92476×10^{-4}	0.0225	0.1417×10^{-4}	0.0029
No reserves	5.1832×10^{-4}	0.1583	0.82201×10^{-4}	0.0217

As for the effects of load correlation, it is seen that for both systems the optimal configuration with the load correlation does poorly compared to the system with the uncorrelated load. This is contrary to the assumption made earlier and can be explained that the load in the ERCOT region does not exactly follow a diurnal pattern as the solar resource and thus cannot be adequately served when peak loads occur after sunset. During the course of a year the ERCOT loads had two peaks on average during the day, in the morning and in the evening, whereas solar energy has peak output during mid-day. This is a likely case for the ERCOT region as the peak loads during the summer are high throughout the day due to the use of air conditioning; thereby a diurnal variation in the solar resource might cause a higher number of potential mismatches at night. This could be remedied by the addition of wind resources into the renewable energy as the wind is more erratic in its diurnal variation. As can be seen from the LOLP and EUE measurements, the addition of wind resources to the solar resources in the system improves the reliability. This could be the result of a decrease in the possibility of a mismatch after sunset, thereby complementing the solar resources.

Features such as the AGC to safeguard the smooth functioning of the grid often require the sudden ramping of generators set aside to supplement the base generation in the event of sudden peaks and troughs in the load. Since it is an often cumbersome and costly process to quickly ramp up or reduce the output of conventional generation, utilities often try to minimize the potential of the variation in the mismatch between loads and the generation. The ramping is studied to see how the wind solar hybrid resource presented in this thesis could potentially alleviate this concern of grid control. The controllable load and the renewable resource ramping were compared using the same configuration as in the load correlation experiment with the deployment in all six zones and the expected renewable output capped at 2600 MW. We wished to establish which of the renewable resources, wind or solar, would be a better choice to obtain a lower value of ramping of the controllable load. The mean values of the ramping and the standard deviation of the ramping were studied for the renewable resources as well as the controllable load (the load supplied by conventional generation). The system configuration with load correlation (primarily composed of solar resources) was compared to the system having load correlation (a considerable amount of wind resources.) This was facilitated using the load correlation by setting the ‘enableload’ variable to 1 in the MATLAB code. The observations are summarized in Table 5.4.

Table 5.4 Ramps as a function of load correlation

Ramping measurements	With load correlation	Without load correlation
Expected renewable ramps	717.893	569.7
Standard deviation of renewable ramps	828.605	580.603
Expected controllable load ramps	800.628	607.871
Standard deviation of controllable load ramps	658.454	548.489

It is seen from Table 5.4 that the system composed of primarily solar resources has both a higher expected (mean) amount of ramping and also has a higher standard deviation. Thus for a system composed mainly of solar resources, the conventional generation would need to undergo many more ramping events of significant magnitudes. Instead, if the system had more wind power renewable resources, it could be expected to experience much fewer ramping events of smaller magnitudes, thus reducing the stress on conventional generation resources placed under AGC control.

Furthermore, a study was also conducted for various levels of expected renewable energy output to see the effect of an increased amount of demand from the system with no reserves. The power range was varied from 1000 MW to 2600 MW and the controllable load ramping was monitored. The results are recorded in Table 5.5.

Table 5.5 Ramping as a function of the expected renewable output

Expected renewable output (MW)	LOLP	EUE	Standard deviation controllable load ramping	Expected controllable load ramping (MW)
1000	0.0016	0.5443	245.9235	325.8876
1200	0.0009	0.3044	265.0372	347.2161
1400	0.0006	0.1679	287.5936	370.9905
1600	0.0003	0.0928	311.8417	396.3393
1800	0.0002	0.0523	337.9634	422.9202
2000	0.0001	0.0304	366.3002	451.2474
2200	0.0001	0.0188	394.2848	478.9259
2400	0	0.0111	438.2277	523.5879
2600	0	0.006	446.3687	530.6626

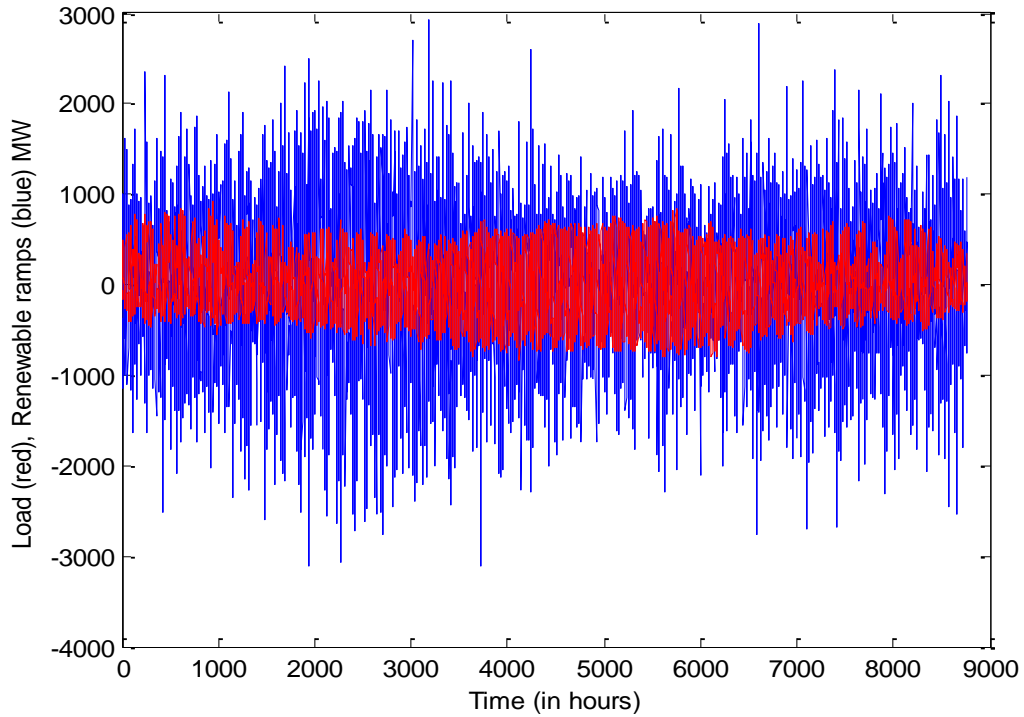


Figure 5.2 Zero LOLP – reliable system scenario

Figure 5.2 shows the stable case with zero LOLP for a system with 2600 MW of expected renewable output with no reserves. The load ramps are smaller in magnitude than the renewable ramps; thus, the system is able to keep up with the load demand.

Another instance, as shown in Figure 5.3, is for an expected renewable output of only 1000 MW and no system reserves. The LOLP is not zero as the controllable load ramps exceed the expected renewable capacity and does not have conventional generation reserves to cover the shortfall.

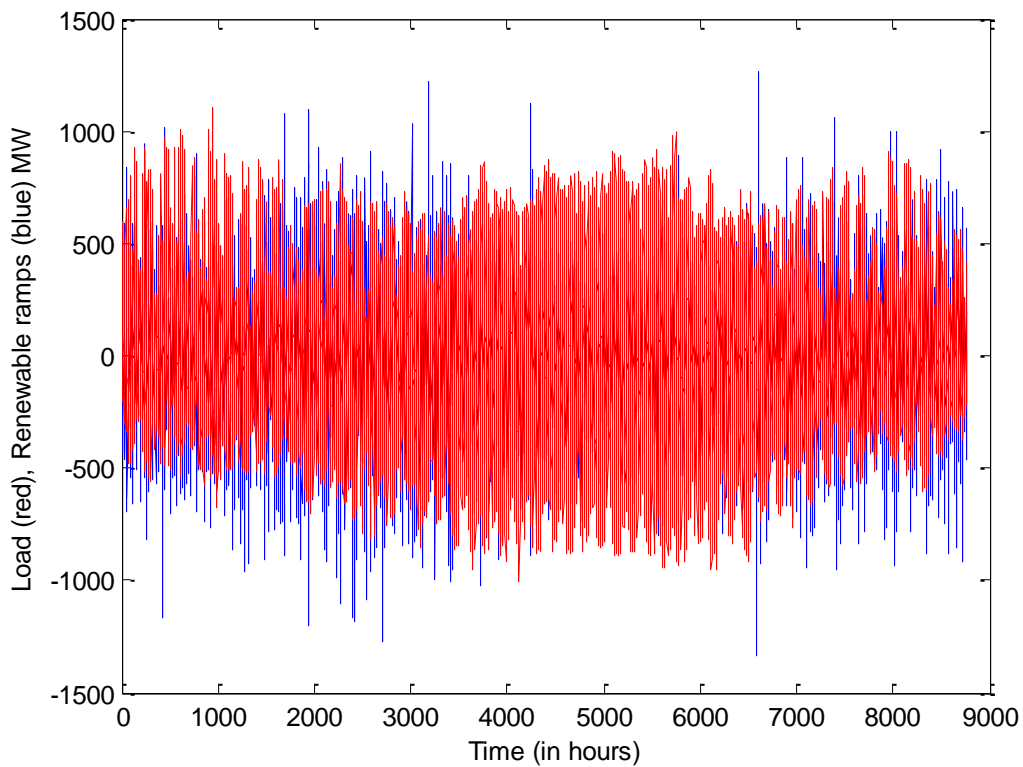


Figure 5.3 LOLP failure

During the course of the year, there are many instances when the controllable load exceeds the renewable ramps, occurring mostly in the early and middle part of the year. This could be attributed to the fact that loads are high during the winter due to heating requirements and during the summer due to cooling requirements. Figure 5.3 clearly depicts that renewable resources thus cannot be expected to alleviate the ramping of the controllable load with an acceptable level of reliability, especially in the absence of conventional generation reserves to cover mismatches during peak load events. From Table 5.5 it can be seen that for an increased expected renewable energy capacity of the system, there is a somewhat exponential increase in reliability in terms of LOLP and the EUE, which slows down after about 1900 MW of renewables are

introduced. However, as the expected capacity of the renewables increases, the expected ramping and the standard deviation also increase.

This result is attributed to the variance of the renewable output becoming the dominant term in the expression of the variance of the controllable load. On the other hand, the correlation of the ramps with the load supplied by renewable is very low at merely 0.1903, indicating that the renewables, despite the optimization efforts, are not a viable solution to alleviate ramping events. This can be explained mathematically as:

$$Var(L - R) = Var(L) + Var(R) - 2 * Cov(L, R)$$

It can be explained that as the covariance between the load and renewable energy shows little change, it is overshadowed by the variance of the renewable ramps. As the scale of the ramps increases, the dominance of the variance of the renewable ramps is scaled as a square of the ramps.

$$Var(a.R) = a^2 * Var(R)$$

As the correlation between the load and the renewable ramps is low, the covariance between the load and the renewable ramps only changes linearly with respect to the standard deviation of the renewable ramps and is overshadowed by the variance of the renewables.

$$\rho_{RL} = \frac{Cov(R, L)}{\sigma_R \cdot \sigma_L}$$

The covariance term does not increase as fast as the variance term; thus, the variance of the system on the whole is not diminished, indicating that using renewables is not the best choice in reducing the controllable load ramps of the system.

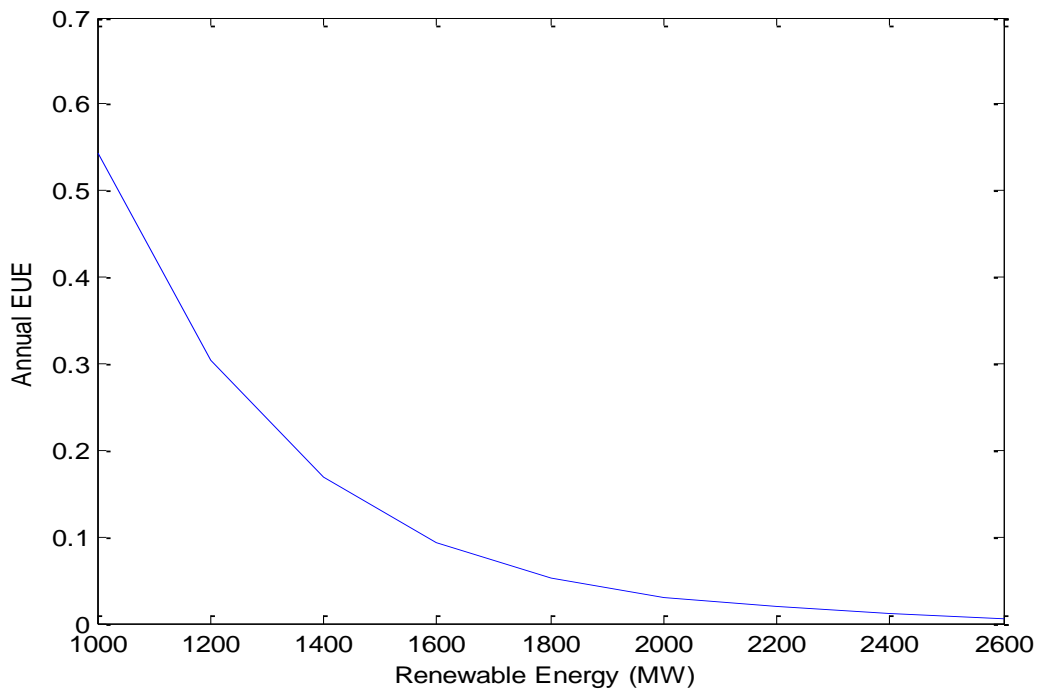


Figure 5.4 EUE improvement with increase in expected renewable energy

As can be seen from Figure 5.4, by adding more renewables after a certain level, there are diminishing marginal returns on the increase in system reliability. In this case it is seen that for an expected renewable capacity beyond 1900 MW, the improvement in reliability is limited.

5.2 Future work

With the research conducted in this thesis, it is hoped that future efforts would be facilitated to study other aspects not covered here. It would be important to involve a cost function to examine the economic viability of wind and solar site diversification schemes. Also the validity of the scheme should be examined, considering the currently available transmission resources, thermal limits, etc. This would help determine the

potential for adapting this idea in utility applications as wind and solar resources gain prominence in the near future as the world attempts to move towards a more environmentally responsible energy scenario.

The optimization algorithm only considers the measures of correlation between wind and solar sites; it could be enhanced by also including elements such as transmission capability, ramping limits, etc., into the constraints for optimization. This thesis only examines a few of the possible siting configurations for its analysis; a more comprehensive study including various permutations and combinations of the load, conventional generation and deployment regions would improve the accuracy of findings. A larger sample size of results on the benefits of geographic diversification could also lead to newer insights as more variables are accounted for in the algorithm. The cost savings from lowering the need for contingency resources due to the reduction in LOLP is of much interest and could be compared to the cost of geographic diversification, i.e., from longer transmission lines, more nodes, etc. Another study that would be beneficial would be to consider hybrid systems with more than two fuel types so as to further exploit the possible uncorrelatedness between resources and thereby also reduce the system variability through fuel type diversification. The addition of storage capacities to the hybrid system alters the reliability metrics and the LOLP, EUE improvements could be quantified. Finally, as this study uses an hourly granularity for a one-year time scale, the study could be enhanced by using data on shorter time scales such as 10 minute or even minute time scales, though the collection of wind and solar data might be cumbersome for shorter time periods. It might be argued that most of the sites selected in this study are usually considered to have good wind or solar resources, thereby precluding a bias of having better than average wind and solar resources. Thus, it would also be of interest to isolate the effects of geographic diversification even further by carefully selecting sites that on average over the course of a year have similar aggregate wind and solar resource outputs

AFTERWORD

We set out with the objective to examine the effects on system reliability as a function of the geographic diversification of wind and solar resources to take advantage of their uncorrelatedness to smooth out power output. The system evaluated was associated mainly with the ERCOT region in Texas and some distant sites in neighboring states. This thesis examines the results on system reliability through the placement of wind turbines and solar arrays constrained by the power production limits placed per site as well as one the entire system. The data collected for the load, conventional generation, wind and solar patterns from ERCOT, NSRDB and TCEQ then helped us create a system that simulated load and generation on an hour scale. The Monte Carlo simulations of the load and generation, in conjunction with the simulated energy production from renewable resources, helped examine the effects on system reliability. The optimization for identifying the optimal configuration of renewable resources was based on the correlations between the wind and solar sites. The active set algorithm is employed by MATLAB to find the global minimum for the selected load, generation and deployment areas. The variability was examined in terms of LOLP, EUE and ramping for various geographic configurations of deployment sites.

In conclusion, this thesis finds that geographic diversification has tangible benefits as the system scale is increased; however, care must be taken to limit the geographic diversification so as to avoid entering the region of diminishing marginal returns to system scale where the system becomes too big. In this thesis it was established that the geographic diversification benefits on improving system reliability become pronounced in the 500 mile range and beyond, but not improving much further as the scale is increased beyond 1000 miles. The LOLP and EUE as complements of each other grant a more comprehensive measure to ascertain the effects of geographic diversification on system reliability.

The reliability of the system is increased with the addition of expected renewable capacity; however, the reliability is heavily impacted by the availability of conventional generation reserves. The replacement of the conventional generation by renewable resources would require extremely high amounts of installed renewable capacity and would also lower the overall system reliability.

Along with conventional generation reserves, the addition of expected renewable resources can seek to alleviate the concerns with the ramping of the controllable load while maintaining system reliability. However, the renewable resources alone are not a viable solution to reduce the ramping of controllable loads as the LOLP and EUE of a system without conventional generation reserves are increased manifold. A hybrid resource with many more forms of renewables, such as biomass, hydro, etc., with an improved capacity factor over the basic wind and solar hybrid system used in this thesis, would improve the system reliability as the uncorrelatedness between the energy sources themselves is increased.

APPENDIX

MATLAB CODE

The code for the studies conducted in this thesis was written in MATLAB. It has two modules, the 'texaswindsolar' module gathers all the data from the user in the form of Microsoft Excel sheets of time series of the load, conventional generation, wind speed and solar insolation data. It then configures the optimal system by minimizing the variance of the various sites. The second module 'simulation' then uses the optimal configuration to assess the reliability of the system by measuring the LOLP, EUE and ramping.

```
clear
clc
close all

tic

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%Specify parameters

SaveFileName='Results'; %Name the save file in between the brackets;
date and time are appended automatically

EnableLoad=0; %1 if you want to consider the correlations with the
load in the optimization (irrelevant for the simulation)
EnablePeakLoad=0; %1 if you only want to keep the peak load data in
the optimization (the simulation part will still consider all the load
data)

%1 for Houston, 2 for North,3 for South, 4 for West, 5 for US wide 500
%iles, 6 US wide 1500 miles
EnableAreas=1; %if 0, then all renewable are considered in the
optimization
```

```

%If all the areas (i.e. the 4 areas) are selected in ChosenAreasGen,
then
%the existing wind generation will also be included; otherwise it is
not
ChosenAreasLoad=[1 2 3 4]; %Input the areas to be considered in terms
of load (numbers from 1 to 4)
ChosenAreasRenewables=[1 2 3 4 5 6]; %Input the areas to be considered
in terms of renewable (numbers from 1 to 6); 5 for US wide 500 miles
and 6 for US wide 1500 miles
ChosenAreasGen=[1]; %Input the areas to be considered for the
conventional generation and existing wind (numbers from 1 to 4)

%Specify range for the expected renewable output
ExpectedOutputRangeMin=1000; %0; %renewable output min
ExpectedOutputRangeMax=10000; %10000; %renewable output max
increment=200;

MaximumInstalledRenewableCapacity=1000000; %In total over all selected
sites in MW

WindFarmCap=1000; %in MW per farm
WindTurbineCap=1.5; %max capacity of a single wind turbine

SolarFarmCap=300; %300; %in MW per farm
SolarArrayCap=1.5; %max capacity of a single solar array

%Reserve level (conventional generator total capacity in % above the
peak load)
ReserveLevel=0;

%Specify Number of samples
NumberSamples=100;

EnableLHS=0; %better NOT use; too many gens, so it is likely to slow
down the code
EnableCommonRandomNumbers=1; %set 0 for the 1st run if the number of
gen or samples has increased

%THERE'S NOTHING ELSE TO SPECIFY BEYOND THIS POINT
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

if EnablePeakLoad==1

    EnableLoad=0;

```

```

end

if EnableAreas==1

    ChosenAreasRenewables=sort (ChosenAreasRenewables, 'ascend');

end

%Import data
%[CovMatrix,txt,row]=xlsread('CovarianceMatrix');
[TimeSeries,txt1,row1]=xlsread('AERMOD_Time_Series');
%xlsread('TimeSeriesExpectedValues');

%Specify areas
%Areas=[0 16 38 54 72 90]; %delimiters in terms of columns; Houston,
then North, then South, then West
Areas=[0 24 48 72 96 114 138];

%Process time series
TimeSeries=TimeSeries/1000; %Convert KW into MW

%Normalize solar data
for i=2:2:size (TimeSeries,2)

    TimeSeries (:,i)=TimeSeries (:,i)*1.5;

end

ExpectedValues=TimeSeries (8761,:);
NumberDV=numel (ExpectedValues);

%Compute the covariance matrix
TimeSeriesOnly=TimeSeries (1:size (TimeSeries,1)-2,:);
CovMatrix=cov (TimeSeriesOnly);
CorrMatrix=corr (TimeSeriesOnly);

if EnableAreas==1

    AreasTimeSeries=[];

    for i=1:numel (ChosenAreasRenewables)

        AreasTimeSeries=[AreasTimeSeries
TimeSeriesOnly (:,Areas (ChosenAreasRenewables (i)) +1:Areas (ChosenAreasRe
newables (i) +1)]];

    end

```

```

TimeSeriesOnly=[];
TimeSeriesOnly=AreasTimeSeries;

ExpectedValues=mean(TimeSeriesOnly,1);
NumberDV=numel(ExpectedValues);
CovMatrix=cov(TimeSeriesOnly);
CorrMatrix=corr(TimeSeriesOnly);

end

if EnableLoad==1

    %Import data
    [Load,txt2,row2]=xlsread('2011_ERCOT_Hourly_load','B2:E8761');

    temp=[];

    for i=1:numel(ChosenAreasLoad)

        temp=[temp Load(:,ChosenAreasLoad(i))];

    end

    temp=sum(temp,2);

    clear Load
    Load=temp;

    Load(end,:)=[]; %make Load the same size as TimeSeriesOnly

    %Compute expected value
    MeanLoad=mean(Load);

    %Compute covariances with the renewable resources
    ExtendedTimeSeries=[TimeSeriesOnly Load];
    ExtendedCovMatrix=cov(ExtendedTimeSeries);
    ExtendedCorrMatrix=corr(ExtendedTimeSeries);

end

if EnablePeakLoad==1

    PeakLoadTimeSeries=[];

    %Trim the time series to only keep the peak load data
    for i=1:24:size(ExtendedTimeSeries,1)-23

        temp=ExtendedTimeSeries(i:i+23,end); %pick up the daily load
    end
end

```

```

        ind=find(temp==max(temp)); %local index of the daily peak load
PeakLoadTimeSeries=[PeakLoadTimeSeries;ExtendedTimeSeries(i+ind-1,:)];

    end

    ExpectedValues=mean(PeakLoadTimeSeries(:,1:end-1),1);
    CovMatrix=cov(PeakLoadTimeSeries(:,1:end-1));
    ExtendedCovMatrix=cov(PeakLoadTimeSeries);
    ExtendedCorrMatrix=corr(PeakLoadTimeSeries);

end

% %Make the covariance matrix whole
% for i=1:size(CovMatrix,2)
%
%     for j=i+1:size(CovMatrix,1)
%
%         CovMatrix(i,j)=CovMatrix(j,i);
%
%     end
%
% end

%Build objective function obj=1/2*x'*H*x+f'*x
H=2*CovMatrix;
f=zeros(NumberDV,1);

if EnableLoad==1

    f=-2*ExtendedCovMatrix(end,1:end-1);

end

count=0;

for
ExpectedOutput=ExpectedOutputRangeMin:increment:ExpectedOutputRangeMax

    count=count+1;

    %Build constraint on the expected value of the total renewable
power output

    A1=[];
    b1=[];

    Aeq=ExpectedValues;
    beq=ExpectedOutput;

```

```

A2=ones(1,NumberDV)*1.5;
b2=MaximumInstalledRenewableCapacity;

A=[A1;A2];
b=[b1;b2];

%Fill in the upper bounds on the decision variables, i.e., wind
turbines
%and 1-MW solar array
ub=zeros(NumberDV,1);

for i=1:NumberDV

    if mod(i,2)==1 %then i is an odd number and the decision
variable is associated to a wind farm

        ub(i)=WindFarmCap/WindTurbineCap;

    elseif mod(i,2)==0 %then i is an even number and the decision
variable is associated to a solar farm

        ub(i)=SolarFarmCap/SolarArrayCap;

    end

end

MaxExpectedOutput=ExpectedValues*round(ub)

%lower bounds
lb=zeros(NumberDV,1);

%Invoke solver quadprog
[x{count},fval,exitflag,output,lambda] =
quadprog(H,f,A,b,Aeq,beq,lb,ub);

iter=1;

while exitflag==0 && iter<=100

    [x{count},fval,exitflag,output,lambda] =
quadprog(H,f,A,b,Aeq,beq,lb,ub,x{count});

    iter=iter+1;

end

```

```

if iter==100
    disp('Optimization did not converge; simulation aborted')
    break
end

ExpectedRenewablePowerOutput(count)=ExpectedValues*x{count}

if EnableLoad==1

ExpectedRenewablePowerOutputMinusLoad(count)=ExpectedRenewablePowerOut
put(count)-MeanLoad;

end

if EnableLoad==1

    StdRenewablePowerOutputMinusLoad(count)=sqrt(fval+var(Load));

end

WindTurbinesNumber=zeros(NumberDV/2,1);
SolarArraysNumber=zeros(NumberDV/2,1);

w=0;
s=0;

for i=1:NumberDV

    if mod(i,2)==1 %then i is an odd number and the decision
variable is associated to a wind farm

        w=w+1;
        WindTurbinesNumber(w)=x{count}(i);

    elseif mod(i,2)==0 %then i is an even number and the decision
variable is associated to a solar farm

        s=s+1;
        SolarArraysNumber(s)=x{count}(i);

    end

end

end

WindTurbinesNumber=round(WindTurbinesNumber)

```

```

SolarArraysNumber=round(SolarArraysNumber)

RenewableInstalledCapacity(count)=round(x{count}).'*ones(NumberDV,1)*1
.5

end

%
plot(ExpectedRenewablePowerOutput,StdRenewablePowerOutputMinusLoad./Ex
pectedRenewablePowerOutput), hold on;
% xlabel('Expected value (renewable power output - load) in MW');
% ylabel('Standard deviation renewable power output in MW');

if EnableLoad==1

plot(RenewableInstalledCapacity,StdRenewablePowerOutputMinusLoad./Rene
wableInstalledCapacity), hold on;
    xlabel('Installed renewable capacity in MW');
    ylabel('Standard deviation renewable power output in MW');

end

%Export data
% columnletter=char('A'-1+comparison);
%
xlswrite('file',transpose(LOLPoverWeek{comparison}),'LOLPoverWeek',[co
lumnletter,'1:',columnletter,'168']);
%
xlswrite('file',mean(transpose(LOLPoverWeek{comparison})), 'LOLPoverWee
k',[columnletter,'170:',columnletter,'170']);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%Simulation%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

for i=1:size(x,2)

    x{i}=round(x{i});

end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

```

```

ChosenAreasGen=sort (ChosenAreasGen, 'ascend');
ChosenAreasLoad=sort (ChosenAreasLoad, 'ascend');

%Specify generator area delimiters
GenAreas=[0 89 231 409 515]; %delimiters in terms of columns; Houston,
then North, then South, then West

%Import
data%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

clear Load TimeSeries AreasTimeSeries LoadData

%Extract capacities
[Capacities, CapacitiesTxt, CapacitiesRaw]=xlsread ('ERCOT_Generation', 'G
3:G517');

%Extract region
[Region, RegionTxt, RegionRaw]=xlsread ('ERCOT_Generation', 'C3:C517');

%Extract fuel type
[Fuel, FuelTxt, FuelRaw]=xlsread ('ERCOT_Generation', 'B3:B517');

%Extract load
[Load, LoadTxt, LoadRaw]=xlsread ('2011_ERCOT_Hourly_load', 'B2:E8761');

%Extract wind and solar outputs
[TimeSeries, txt1, raw1]=xlsread ('AERMOD_Time_Series');
%xlsread ('TimeSeriesExpectedValues');

%Extract existing wind power outputs for the entire ERCOT system
[WindPowerOutput, Windtxt1, Windraw1]=xlsread ('existing_wind', 'D5:D8764'
);

%Process time series
TimeSeries=TimeSeries/1000; %Convert KW into MW

%Normalize solar data
for i=2:2:size (TimeSeries, 2)

    TimeSeries (:, i)=TimeSeries (:, i)*1.5;

end

TimeSeries (end-1:end, :)=[]; %Eliminate last two rows (i.e. the
computation of the expected values)

AreasTimeSeries=[];

```

```

for i=1:numel(ChosenAreasRenewables)

    AreasTimeSeries=[AreasTimeSeries
TimeSeries(:,Areas(ChosenAreasRenewables(i))+1:Areas(ChosenAreasRenewables(i)+1))];

end

TimeSeries=[];
TimeSeries=AreasTimeSeries;

%Select load to be used based on ChosenAreasLoad
LoadData=[];

for i=1:numel(ChosenAreasLoad)

    LoadData=[LoadData Load(:,ChosenAreasLoad(i))];

end

LoadData=sum(LoadData,2);
LoadData(end,:)=[];

%Create useful data table based on ChosenAreasGen
CapData=[];
FuelData=[];

for i=1:numel(ChosenAreasGen)

CapData=[CapData;Capacities(GenAreas(ChosenAreasGen(i))+1:GenAreas(ChosenAreasGen(i)+1))];

FuelData=[FuelData;FuelTxt(GenAreas(ChosenAreasGen(i))+1:GenAreas(ChosenAreasGen(i)+1))];

end

TotalAvailableConventionalGenCap=sum(CapData);

%Turn fuel type into FOR

FOR=zeros(size(FuelData,1),1);
lambdaf=zeros(size(FuelData,1),1); %failure rate (h^-1)
mur=zeros(size(FuelData,1),1); %recovery rate (h^-1)
MTTF=zeros(size(FuelData,1),1); %mean time to failure (h)
MTTR=zeros(size(FuelData,1),1); %mean time to recovery (h)

```

```

for i=1:size(FuelData,1)

temp=FuelData(i);

if strcmp(temp, 'GAS')==1

FOR(i)=0.061;
%           MTTF(i)=450;
%           MTTR(i)=50;
%
%           lambda_f(i)=1/((MTTF(i)+MTTR(i))*(1-FOR(i)));
%           mur(i)=1/((MTTF(i)+MTTR(i))*FOR(i));

elseif strcmp(temp, 'BIOMASS')==1

FOR(i)=0.061;
%           MTTF(i)=1960;
%           MTTR(i)=40;
%
%           lambda_f(i)=1/((MTTF(i)+MTTR(i))*(1-FOR(i)));
%           mur(i)=1/((MTTF(i)+MTTR(i))*FOR(i));

elseif strcmp(temp, 'OTHER')==1

FOR(i)=0.05;
%           MTTF(i)=1200;
%           MTTR(i)=50;
%
%           lambda_f(i)=1/((MTTF(i)+MTTR(i))*(1-FOR(i)));
%           mur(i)=1/((MTTF(i)+MTTR(i))*FOR(i));

elseif strcmp(temp, 'COAL')==1 && CapData(i)>=250

FOR(i)=0.08;
%           MTTF(i)=950;
%           MTTR(i)=50;
%
%           lambda_f(i)=1/((MTTF(i)+MTTR(i))*(1-FOR(i)));
%           mur(i)=1/((MTTF(i)+MTTR(i))*FOR(i));

elseif strcmp(temp, 'COAL')==1 && CapData(i)<250

FOR(i)=0.04;
%           MTTF(i)=950;
%           MTTR(i)=50;
%

```

```

%             lambdaf(i)=1/((MTTF(i)+MTTR(i))*(1-FOR(i)));
%             mur(i)=1/((MTTF(i)+MTTR(i))*FOR(i));

elseif strcmp(temp,'NUC')==1

FOR(i)=0.12;
%             MTTF(i)=2940;
%             MTTR(i)=60;
%
%             lambdaf(i)=1/((MTTF(i)+MTTR(i))*(1-FOR(i)));
%             mur(i)=1/((MTTF(i)+MTTR(i))*FOR(i));

elseif strcmp(temp,'HYDRO')==1

FOR(i)=0.01;
%             MTTF(i)=960;
%             MTTR(i)=40;
%
%             lambdaf(i)=1/((MTTF(i)+MTTR(i))*(1-FOR(i)));
%             mur(i)=1/((MTTF(i)+MTTR(i))*FOR(i));

elseif strcmp(temp,'WIND')==1

FOR(i)=1; %0.04
%             MTTF(i)=1100;
%             MTTR(i)=150;
%
%             lambdaf(i)=1/((MTTF(i)+MTTR(i))*(1-FOR(i)));
%             mur(i)=1/((MTTF(i)+MTTR(i))*FOR(i));

elseif strcmp(temp,'SOLAR')==1

FOR(i)=1; %0.04
%             MTTF(i)=1980;
%             MTTR(i)=20;
%
%             lambdaf(i)=1/((MTTF(i)+MTTR(i))*(1-FOR(i)));
%             mur(i)=1/((MTTF(i)+MTTR(i))*FOR(i));

end

end

%Monte Carlo
Simulation%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%
```

```

count=0;

for
ExpectedOutput=ExpectedOutputRangeMin:increment:ExpectedOutputRangeMax

    ExpectedOutput
    count=count+1;

    %Weight TimeSeries data with the appropriate number of wind
    turbines /
    %solar arrays
    WeightedTimeSeries{count}=TimeSeries*x{count};

    %Scale the load if necessary based on ReserveLevel

    LoadScaleCoeff=(sum(CapData)/(1+ReserveLevel/100))/max(LoadData);
    LoadData=LoadData*LoadScaleCoeff;

    AvailableCapacity=zeros(NumberSamples,size(LoadData,1));
    LossLoad=zeros(NumberSamples,size(LoadData,1));
    UnservedEnergy=zeros(NumberSamples,size(LoadData,1));

    if EnableLHS==1

        %Sample generator states (use one Latin hypercube per hour)
        for h=1:size(LoadData,1)

            h
            GenSamples{h}=lhsdesign(NumberSamples,numel(FOR));
            GenState{h}=ones(NumberSamples,numel(FOR));

        end

    else for h=1:size(LoadData,1)

        h
        GenSamples{h}=zeros(NumberSamples,numel(FOR));
        GenState{h}=ones(NumberSamples,numel(FOR));

    end

end

if EnableCommonRandomNumbers==1

    load Samples

end

```

```

%Simulation
for sample=1:NumberSamples

    sample

    for h=1:size(LoadData,1)

        h
        sample

        if EnableCommonRandomNumbers==0

            for g=1:numel(FOR) %find out about generator states

                if EnableLHS==0

                    GenSamples{h}(sample,g)=rand;

                end

                if GenSamples{h}(sample,g)<=FOR(g) %then generator
g is outage

                    GenState{h}(sample,g)=0;

                end

            end

        end

    end

    AvailableCapacity(sample,h)=GenState{h}(sample,:)*CapData+WeightedTime
Series{count}(h); %Available Capacity contains both conventional
generation and renewables

    if numel(ChosenAreasGen)==4

        AvailableCapacity(sample,h)=AvailableCapacity(sample,h)+WindPowerOutpu
t(h);

    end

    if AvailableCapacity(sample,h)<LoadData(h) %then we have a
loss of load event

```

```

        LossLoad(sample,h)=1;
        UnservedEnergy(sample,h)=LoadData(h)-
AvailableCapacity(sample,h);

        end

    end

    SampleAnnualLOLPRecord=mean(LossLoad,2);
    SampleAnnualLOLP(sample)=SampleAnnualLOLPRecord(sample);

EstimatedAnnualLOLP(sample)=mean(SampleAnnualLOLPRecord(1:sample));

    temp=LossLoad(1:sample,:);

SampleCoV_LOLP(sample)=(std(temp(:))/(sqrt(sample*size(LoadData,1))))/
(EstimatedAnnualLOLP(sample));

    end

    %Compute LOLP and EUE (final estimates)

    HourlyLOLP{count}=mean(LossLoad,1);
    HourlyEUE{count}=mean(UnservedEnergy,1);

    AnnualLOLP(count)=mean(HourlyLOLP{count})
    AnnualEUE(count)=mean(HourlyEUE{count})

    %Compute coefficient of variation

CoV_LOLP(count)=(std(LossLoad(:))/(sqrt(NumberSamples*size(LoadData,1)
))) / AnnualLOLP(count)

EstimatedCoV_LOLP(count)=std(EstimatedAnnualLOLP) / AnnualLOLP(count);

    %Compute rampings of the renewable power outputs and controllable
load
    %(that is, load - aggregated renewable power output)
    RenewableRamps{count}=diff(WeightedTimeSeries{count});
    ExpectedRamps(count)=mean(abs(RenewableRamps{count}));
    StdRamps(count)=std(abs(RenewableRamps{count}));

    ControllableLoadRamps{count}=diff(LoadData-
WeightedTimeSeries{count})

ExpectedControllableLoadRamps(count)=mean(abs(ControllableLoadRamps{co
unt}))

```

```

StdControllableLoadRamps(count)=std(abs(ControllableLoadRamps(count)))

end

savefile='Samples';
save(savefile, 'GenState');

figure(1)
plot([ExpectedOutputRangeMin:increment:ExpectedOutputRangeMax], AnnualL
UE), hold on;

% figure(2)
%
plot([ExpectedOutputRangeMin:increment:ExpectedOutputRangeMax], AnnualL
OLP), hold on;
%
plot([ExpectedOutputRangeMin:increment:ExpectedOutputRangeMax], StdRene
wablePowerOutputMinusLoad, 'r');

if EnableLoad==1

    figure(2)

    [AX, H1, H2]=plotyy([ExpectedOutputRangeMin:increment:ExpectedOutputRang
eMax], AnnualLOLP, [ExpectedOutputRangeMin:increment:ExpectedOutputRange
Max], StdRenewablePowerOutputMinusLoad)
    set(get(AX(1), 'Ylabel'), 'String', 'Annual LOLP');
    set(get(AX(2), 'Ylabel'), 'String', 'Standard deviation of (renewable
power output minus load) in MW');
    xlabel('Expected renewable output in MW');

    figure(2)

    [AX, H1, H2]=plotyy(RenewableInstalledCapacity, AnnualLOLP, RenewableInsta
lledCapacity, StdRenewablePowerOutputMinusLoad./RenewableInstalledCapac
ity)
    set(get(AX(1), 'Ylabel'), 'String', 'Annual LOLP');
    set(get(AX(2), 'Ylabel'), 'String', 'Standard deviation of (renewable
power output minus load) in MW');
    xlabel('Expected renewable output in MW');

end

figure(3)
plot(RenewableRamps{1}), hold on
plot(diff(LoadData), 'r'), hold off

```

```
CorrelationRampRenewableLoad=corr(RenewableRamps{1},diff(LoadData));

clear TimeSeries AreasTimeSeries LoadData FuelData Region RegionTxt
RegionRaw Fuel FuelTxt FuelRaw Load LoadTxt LoadRaw TimeSeries txt1
raw1 WindPowerOutput Windtxt1 Windraw1 Capacities CapacitiesTxt
CapacitiesRaw GenState LossLoad UnservedEnergy AvailableCapacity
GenState GenSamples TimeSeriesOnly ExtendedTimeSeries
PeakLoadTimeSeries

SaveFile=[SaveFileName,datestr(now,'yyyy-mm-dd HH;MM;SS')];
save(SaveFile);

toc
```

REFERENCES

- [1] B. Drake and K. Hubacek, "What to expect from a greater geographic dispersion of wind farms?—A risk portfolio approach," *Energy Policy*, vol. 35, no. 8, pp. 3999-4000, Aug. 2007.
- [2] F. Roques, C. Hiroux and M. Saguan, "Optimal wind power deployment in Europe—A portfolio approach," *Energy Policy*, vol. 38, no.7, pp. 3245–3256, Jul. 2010.
- [3] Y. Rombauts, E. Delarue and W. D'haeseleer, "Optimal portfolio-theory-based allocation of wind power: Taking into account cross-border transmission-capacity constraints," *Renewable Energy*, vol. 36, no. 9, pp. 2374–2387, Sep. 2011.
- [4] Y. M. Atwa, E.F. El-Saadany, M.M.A. Salama, R. Seethapathy, M. Assam and S. Conti, "Adequacy evaluation of distribution system including wind/solar DG during different modes of operation," *IEEE Trans. on Power Systems*, vol.26, no.4, pp.1945-1952, Nov. 2011.
- [5] Y. Degeilh and C. Singh, "A quantitative approach to wind farm diversification and reliability," *Intl. J. of Electrical Power & Energy Systems*, vol. 33, no. 2, pp. 303-314, Feb. 2011.
- [6] F.J. Ardakani, G. Riahy and M. Abedi, "Design of an optimum hybrid renewable energy system considering reliability indices," in *18th Iranian Conference on Electrical Engineering, ICEE 2010, Tehran, Iran*, May 10-13, 2010. pp. 842-847.
- [7] A. Testa, S. De Caro and T. Scimone, "Optimal structure selection for small-size hybrid renewable energy plants," in *14th European Conference on Power Electronics and Applications (EPE 2011)*, Aug. 30, 2011-Sept. 1, 2011, pp.1-10.
- [8] S. R. Dahman, S. Grijalva, K. J. Patten and A. M. Visnesky Jr., "Large-Scale integration of wind generation including network temporal security analysis," *IEEE Trans. on Energy Convers.*, pp. 181-188, vol. 22, no. 1, Mar. 2007
- [9] G.R. O'Connor, "Benefits of spatial smoothing for the integration of wind power," M.S. thesis, Dept. of Elect. Eng, Univ. of Pittsburgh, PA, 2011.
- [10] M. Carolin Mabel, R. Edwin Raj and E. Fernandez , "Analysis on reliability aspects of wind power," *Renewable and Sustainable Energy Reviews*, vol. 15, no. 2, pp. 1210–1216. Feb. 2011.

- [11] G. Tina and S. Gagliano, "Probability analysis of weather data for energy assessment of hybrid solar/wind power system," Presented at the 4th IASME/WSEAS International Conference on Energy, Environment, Ecosystems and Sustainable Development (EEESD'08), Algarve, Portugal, Jun. 11-13, 2008.
- [12] G. Sinden, "Characteristics of the UK wind resource: Long-term patterns and relationship to electricity demand," *Energy Policy*, vol. 35, no. 1, Jan. 2007, pp. 112-127.
- [13] B. Kirby, 2007, "Evaluating transmission costs and wind benefits in Texas: Examining the ERCOT CREZ transmission study," The Wind Coalition and Electric Transmission Texas, LLC, Texas PUC Docket NO. 33672, Apr. 2007.
- [14] S. Wilcox and W. Marion, "National solar radiation database, 1999–2005: User's manual," Natl. Renewable Energy Lab., Golden, CO, Tech. Rep., TP-581–41364, 2007.
- [15] J.O. Paumeir and R.O. Brode, "User's guide for the AERMOD meteorological Preprocessor (AERMET)," U.S. Environmental Protection Agency, Research Triangle Park, NC, EPA-454/B-03-002, Nov. 2004.
- [16] V.Thapar, G. Agnihotri and V.K. Sethi, "Critical analysis of methods for mathematical modelling of wind turbines," *Renewable Energy*, vol. 36, no.11, pp. 3166-3177, Nov. 2011.
- [17] J. Schwartz and D.Puffer, "Home Power's 2007 Solar-Electric module guide," *Home Power 121*, pp.70-78, Oct.-Nov., 2007.
- [18] C.S. Supriya and M. Siddarthan, "Optimization and sizing of a grid-connected hybrid PV-wind energy system," *Intl. J. of Eng. Science and Tech.*, vol. 3, no. 5, pp. 4296-4323, 2011.
- [19] Y.Degeilh, "Wind farm diversification and its impact on power system reliability," M.S. thesis, Dept. of Elect. Eng, Texas A&M Univ., TX, 2009.
- [20] R.D. Prasad, R.C. Bansal and M. Sauturaga, "Some of the design and methodology considerations in wind resource assessment," *Renewable Power Generation, IET*, vol. 3, no.1, pp. 53-64, Mar. 2009.
- [21] G. McNerney and R. Richardson, "The statistical smoothing of power delivered to utilities by multiple wind turbines," *IEEE Trans. Energy Convers.*, vol. 7, no. 4, pp. 644-647, Dec. 1992.

- [22] J. Dunlop, "Modern portfolio theory meets wind farms," *J. of Pvt. Equity*, vol. 7, no. 2, pp. 83-95, Spr. 2004.
- [23] MATLAB version 7.12.0. Natick, MA, The MathWorks Inc., 2010.
- [24] F. Olsina and C. Larisson, "Optimization of spinning reserve in stand-alone wind-diesel power systems," in *Wind Power*, 1st ed. Vukovar, Croatia, InTech, June 2010, ch.19, pp. 437-465
- [25] P.Hu, "Reliability evaluation of electric power system including wind power and energy storage," Ph.D. dissertation, Dept. of Elect. & Comput. Eng, Univ. of Saskatchewan, Saskatoon, Canada, 2009.
- [26] G. Taljan, M. Pantoš and A. F. Gubina, "Unified approach to reliability assessment in an isolated RES System," presented at the 9th International Conference on Probabilistic Methods Applied to Power Systems KTH, Stockholm, Sweden , Jun. 11-15, 2006.
- [27] M. Bosatra, F. Fazi, P. F. Lionetto and L. Travagnin, "Utility-Scale PV and CSP solar power plants: Performance, impact on the territory and interaction with the grid," Presented at Power-Gen Europe 2010, Amsterdam, The Netherlands, June 8 –10, 2010.
- [28] A. E. Curtright and J. Apt, "The character of power output from Utility-Scale Photovoltaic systems," *Progress in Photovoltaics: Research and Applications*, vol. 16, no. 3, pp. 241–247, May 2008.
- [29] R. D. Wright and T. E. Ramsay, Jr., "On the effectiveness of common random numbers," *Management Science*, vol. 25, no.7, Jul., 1979, pp. 649-656.
- [30] M. Haugh, "Simulation methods and an introduction to variance reduction methods," IEOR E4703: Monte Carlo Simulation, Dept. of IE&OR, Columbia University, NY, Fall 2004, unpublished.
- [31] M. Milligan and K. Porter, "Determining the capacity value of wind: A survey of methods and implementation," presented at Windpower 2005, Denver, CO, May 15–18, 2005
- [32] A. Noguee, J. Deyette and S. Clemmer, "The projected impacts of a national renewable portfolio standard," *The Electricity Journal*, vol. 20, no. 4, pp. 33-47, May 2007.