THESIS

STATISTICAL ANALYSIS OF THE CHALLENGES TO HIGH PENETRATION OF WIND ENERGY

Submitted by

Matthew O'Connell

Department of Mechanical Engineering

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Summer 2014

Master's Committee:

Advisor: Anthony Marchese
Co-Advisor: Daniel Zimmerle

Peter Young

ABSTRACT

STATISTICAL ANALYSIS OF THE CHALLENGES TO HIGH PENETRATION OF WIND ENERGY

Grid penetration of renewable energy technologies, especially wind power, is higher than ever and continues to increase. The inherent stochastic variability of wind makes predicting wind, and thus power generation difficult. Generating companies usually don't openly share power output predictions or historical generation data which increases the level of complexity when determining new wind plant locations or estimating delivered grid level power. This work focuses on statistical data analysis and advanced data modeling related to wind power forecasting and generation.

The first part of this thesis uses power output logs from several wind plants and a well-known forecasting method to determine energy storage requirements for individual wind plant contract firming. Forecasts of varying accuracy are used to characterize storage requirements based on contract period length, forecast lead time, and forecast accuracy. Results show that forecast error distributions are effected more by forecast accuracy and lead time than wind plant size and location. The biggest reductions in produced power deviations can be achieved by increasing forecast accuracy and decreasing forecast lead time.

The second part of this work develops a statistical analysis which allows estimation of contract firming requirements for a specific wind plant location without the need for time series wind and forecast data. The developed method requires only a wind speed and forecasting error distribution. Using these distributions, deviations between forecast to produced power and energy

can be estimated. Results from comparing to historical time series data show this method is accurate to within 10% of actual amounts. Since distributions are much more easily attained than historical time series data, this analysis is useful for developers when evaluating potential new locations.

The third part of this work uses a pattern matching algorithm to recognize wind ramp events and separate the forecasting error due to timing from the forecasting error due to magnitude. Wind ramp detection is achieved by developing a pattern matching algorithm which is also shown to work in identifying start and stop transients in electrical device current draw. The analysis confirms wind ramp events can be detected by calculating a bimodal ranking value from a histogram of power data, and the effects of forecast timing and magnitude can be separated from overall forecasting errors. The results of this analysis show magnitude errors contribute more in large wind ramp events, while timing errors contribute more in small ramp events.

TABLE OF CONTENTS

ABSTRACT ii

INTRODUCTION 1

Renewable Energy 1

Wind Power 1

Forecasting Methods 2

Wind Ramps 3

THESIS OVERVIEW 4

PART I: 6

SUPPLEMENTAL ENERGY NEEDED FOR WIND INTEGRATION 6

1. INTRODUCTION 6

2. DATA AND PROBLEM DEFINITION 7

3. REPRESENTING FORECAST ERRORS 8

3.1. Quantify a Reference Forecast Error 8

3.2. Reliability Metric 10

3.3. Sensitivity Simulations 11

4. SENSITIVITY ANALYSIS 12

4.1. Wind Plant Configuration 12

4.2. Contract Period 13

4.3. Measurement Period 14

4.4. Forecast Accuracy 15

5. CONCLUSIONS 16

PART II: 18

STATISTICAL ESTIMATION OF WIND FORECAST ACCURACY USING WIND SPEED AND FORECAST ERROR DISTRIBUTIONS 18

1. INTRODUCTION 18
2. DATA AND ANALYSIS 23
2.1. Method 1 24
2.2. Method 2 27
2.2.1 Error Modeling 28
2.3. Method 3 30
2.3.1. Forecast Comparison 34
3. RESULTS 34
4. CONCLUSIONS 37
PART III: 39
ADVANCED WIND RAMP DETECTION METHODLOGY USING BIMODALITY STATISTICS 39
1. INTRODUCTION 39
2. DETECTION METHODOLOGY 42
2.1. Data 42
2.2. Detection Methods 43
2.2.1 Bimodality Coefficient 45
2.2.2 Hartigan's Dip Test Statistic 46
2.3 Ramp Detection 49
3. TIMING AND MAGNITUDE ANALYSIS 52
3.1 Power and Energy Errors 53
3.2 Timing and Magnitude 54
4. CONCLUSIONS 59
OVERALL CONCLUSIONS 60
REFERENCES 62

INTRODUCTION

Renewable Energy

Energy is available in two forms: renewable and non-renewable. Renewable energy is energy generated from natural resources which are replenished on a typical human's life timescale, or within ~100 years. Non-renewable sources are naturally replenished over thousands of years and once completely deplete, the current or immediate future generations of humans can no longer access them. Non-renewable sources include coal, oil, natural gas, etc. Types of renewable energy include solar, wind, geothermal, hydro, etc. Typically, energy is removed from non-renewable sources through combustion which produces byproducts such as CO2 and methane, both greenhouse gasses. The use of non-renewable fuels for transportation and electricity production accounts for the most CO2 emissions worldwide, and the use of motor vehicles alone in the US account for up to 30% of its CO2 emissions [1]. Studies continue to conclude greenhouse gases contribute to global warming [2]. Due to the link among non-renewable energy, global warming, greenhouse gases, and climate change; renewable energy is an area of high interest worldwide. In addition to environmental benefits, countries also seek renewable energy sources for energy independence. Wind energy in particular is very well established and one of the principle sources of renewable energy [3]. According to forecasts of the Global Wind Energy Council, it could provide more than 20% of worldwide electricity by 2030 [4].

Wind Power

Installed wind capacity in the United States and worldwide has been following an upward trend, which is expected to continue throughout the next several decades [5]. Continuous improvement of wind turbine technology, governmental policy encouragement, and reduction of

costs has helped allow this to occur [6]. The increased prominence of wind power magnifies the inherent problems with high penetrations of wind power, e.g. its stochastic nature [7]. Unpredictable weather patterns and unpredictable wind behavior cause unpredictable power output from wind power plants. Unanticipated fluctuations in power output must be mitigated at the grid level through load balancing techniques such as natural gas firming or electric energy storage [8]. Natural gas firming diminishes the environmental benefit of renewable energy technologies, and electric energy storage systems require costly materials, installation, and maintenance [8]. Improving forecasts is a simple solution to reduce firming requirements and carbon emissions of wind power while increasing its potential for high penetration [9]. Although higher forecast accuracies will still require firming, it will be at a reduced level.

Forecasting Methods

A significant barrier to large scale wind power integration is the inaccuracy of wind forecasts [10]. Wind forecasts must predict both timing and magnitude to be effective. However, atmospheric conditions are chaotic and there exists a limit to predictability which degrades as prediction time increases [11], [12]. Prediction time as used here is the duration of time between forecast generation and the time period for which it is predicting wind activity. Errors in forecasts translate to costs in the energy market, thus more accurate forecasts can translate to savings for utilities and customers [13]. Short term forecasts, typically up to one hour, are predominantly persistence based. As forecast time increases, complex forecasting models such as the National Center for Atmospheric Research (NCAR) Dynamic Integrated Forecast System (DICast) and Variational Doppler Radar Assimilation System (VDRAS) forecasting systems analyze meteorological conditions, wind plant data, and previous forecast inaccuracies to predict wind plant power output [14]. These predictions are used by wind plant and grid operators such as Xcel

Energy to predict power output for specific times. While forecasting errors can be mitigated with operating reserves or electrical energy storage, neither is as advantageous as optimizing forecast accuracy.

Wind Ramps

Large changes in wind speed over relatively short times are known as wind ramps [13]. Since wind speeds govern turbine power output, wind ramps cause power ramps. Large unpredicted ramps are especially problematic as they are sources of large deviations between forecasted and produced power. It is relatively easy to balance electrical loads when a small wind plant experiences a ramp since the magnitude of the ramp, even at 100% plant capacity, is still small. However as wind penetration and wind plant sizes continue to increase, ramp event size also increases. These larger ramp events can make up a significant percentage of grid power, and balancing grid load to generation starts to become problematic [15]. Therefore, in addition to improving overall wind forecasting techniques, improving recognition and prediction accuracy of wind ramps is of utmost importance to the future of wind power.

THESIS OVERVIEW

This thesis is divided into three parts: I) Supplemental Energy Needed for Wind Integration, II) Statistical Estimation of Wind Forecast Accuracy Using Wind Speed and Forecast Error Distributions, and III) Advanced Wind Ramp Detection Methodology Using Bimodality Statistics.

Part I focuses on understanding and estimating the impact of forecast accuracy on contract firming requirements for individual wind plants. Actual forecasts were not accessible for the power output logs available. Instead, statistical distributions and power output logs from several wind plants are used to create artificial forecasts. A forecast accuracy parameter, γ, is created to adjust artificial forecast accuracy, and forecasts are created with varying levels of inaccuracy. A "perfect" forecast is used as a reference best-case comparison. These forecasts are used to relate contract firming requirements to forecast inaccuracy, as well as evaluate impacts of other parameters such as regional aggregation, contract period length, and wind plant size on firming requirements. The results provide guidance on which factors of wind plants (forecast accuracy, size, contract length, etc.) are the most influential on contract firming requirements

Part II develops a statistical method using limited data to estimate contract firming requirements for an individual wind plant. This is important since power and forecast data are usually not available for wind plant developers. Unlike Part I, actual forecasts corresponding to the power output logs for one plant are available, allowing measurement of the performance of the estimation method. Part II outlines three separate methods of firming requirement estimation and concludes in a comparison of all three. Method 1 uses a simple persistence forecast to estimate contract firming requirements. Persistence forecasting is a method where forecasts are made equal

to what is currently or previously observed for the same location. For short time scales one hour or less, persistence still performs better than the most advanced forecasting models. Method 2 uses time series power data and a forecast error distribution instead of time series data. It samples the forecast error distribution and compares against the time series power observation data to generate contract firming requirements. Method 3 requires no time series data since it calculates estimated contract firming requirements using a wind power distribution and a forecast error distribution. All three methods are shown to estimate similar firming requirements. As a final step, Method 1 was repeated for the wind plant with actual forecasts available. These results are compared with contract firming requirements calculated from actual forecast data.

Part III investigates an important facet of wind forecasting: wind ramps. Wind ramps are rapid, large changes in wind speeds that significantly change power output of wind plants. Unpredicted ramps lead to large gaps between forecast and produced power and energy. Part III develops an advanced wind ramp detection model that filters out potential false triggers of other detections methods by monitoring the bimodality of a histogram of wind power levels. It develops a method of detecting wind ramps using an aggregated bimodality metric derived from several methods and a novel detection methodology. Part III then separates the effect of timing and magnitude on forecast ramp errors from the overall forecast error to determine the individual contribution of each variable. This research can direct further research into the area of the factor considered more important.

PART I:

SUPPLEMENTAL ENERGY NEEDED FOR WIND INTEGRATION

1. INTRODUCTION

Wind energy is well established as one of the principal sources of renewable electricity [3]. With multiple political jurisdictions calling for increasing the proportion of electricity generated from renewable resources [16], [17], penetration of renewables, including wind, will likely continue to increase [18]. A key concern with renewable resources is the variability in output due to changes in metrological conditions – i.e. wind speed for wind turbines – that are uncorrelated with load. To reduce the variability, multiple studies have proposed methods to smooth wind power output allowing plants to commit to firm power contracts for given contractual periods, a result often called "firming" [7].

Studies have recommended firming wind plants at a regional level by exploiting the natural smoothing that occurs as wind power is aggregated regionally [19]. However, regional smoothing is not always practical. There may be economic or structural constraints that prevent regional investment, and therefore an interest in requiring individual plants to firm their own output. In other cases, technical constraints, such as weak transmission links, may make it advisable to firm power locally. Independent of these specific business considerations, the question of firming individual wind plants remains active, and in the corresponding author's experience, is most often driven by three questions. First, developers of storage technologies are interested in the wind firming applications, and no ready design guidance exists to determine if a particular storage technology is well suited to this application [17], [20]. Second, changes in market operations, such

as reduced contract windows and later commitment periods [21], have reduced energy and power requirements for wind firming, returning to the question of whether it is practical and/or desirable. Third, teams working on improved wind and wind-power forecasts are interested in how improved forecasts will improve the predictability of wind power [22], [23], and the energy needed to firm a wind plant is good measure of the net effect of forecast errors.

Therefore, the purpose of this study is to provide design guidance for the three questions above. To start, we illustrate an easily understood and replicated bounding analysis that can be applied to wind plant data and compared to the present data set. To do so, we utilize the energy required to firm a wind plant as a means of quantifying the variability in the plant, which is also one of two metrics of key interest to storage technology providers. The other key metric being peak power requirements. Then, we analyze how this "required energy" metric varies with changes in conditions including contract period, forecast accuracy, measurement window.

2. DATA AND PROBLEM DEFINITION

The data utilized in this study are power output logs for 13 wind plants, provided by the National Renewable Energy Laboratory (NREL). The data for nine plants was provided at one-minute resolution, and the one-second resolution data from an additional four plants was averaged to produce one-minute resolution. The plants cover three geographic areas throughout the western United States and all plants had data for a minimum of two years.

This analysis considers two nested time intervals for evaluation of the wind plants' performance. First, for "firm" power contracts, generating plants estimate and contractually agree to a certain energy production for a "contract period", which can range from five minutes to an hour or more. While currently rare in wind power purchase agreements, we assume that penalties

may be applied for either over- or under-producing relative to the contracted amount. Further, several measurements of the power output may be made during the contract period, which we term the "measurement period." For example, a one-hour contract may be measured at ten minute intervals. Penalties could be applied when power output from the plant deviates from the contracted output during measurement periods, as such errors impact the firming resources required from the system operator. We are interested in, first, the energy required to firm a wind plant for a range of contract periods, and second, how that energy is impacted if several binding power measurements are made during the contract period.

3. REPRESENTING FORECAST ERRORS

There exists no standard definition to quantify the deviation between forecasted power output and actual power output for a wind plant. This section describes how deviations were quantified for this study, reduced to a small set of specific metrics, and then utilized to compare a wide range of wind plant and wind forecast characteristics. This analysis consists of three operational steps:

3.1. Quantify a Reference Forecast Error

For given contract and measurement periods, the deviation between forecast power and power output was computed using a hypothetical power forecast. The power deviations were integrated to produce energy deviations. The histogram of energy deviations were then fit to a suitable probability density function producing a parameter(s) which can be compared for subsequent analysis. Following [24], a Cauchy probability density function (PDF) was utilized, and fits were performed on the cumulative form given in (1), where μ is the mean and γ is the

scale parameter of the distribution. For the wind power forecasts analyzed, the mean is forced to zero by the sampling process, therefore allowing the distribution to be represented solely by the scale parameter.

$$P(x) = \frac{1}{2} + \frac{1}{\pi} * \tan^{-1}\left(\frac{x-\mu}{\gamma}\right) \quad (1)$$

Since persistence forecasts are widely used and easily simulated [23], Fig. 1, this analysis uses persistence as a reference forecast for all subsequent work. When a persistence forecast for a specified contract period and wind plant is fit to the Cauchy distribution, a representative scaling parameter is produced, $\gamma_f(F, C)$, where F identifies the wind plant and C is the length of the contract period. Looking across all plants and periods, the mean of all scaling parameters is 0.014, and γ_f is shown for all plants for a range of contract periods in Fig. 2. Measurement periods in Fig. 2 are equal to contract period for each case. Results from [24] and [25] also show that as contract period increases, the error distributions have fatter tails, which are represented by Cauchy distributions with larger scale parameters, γ.

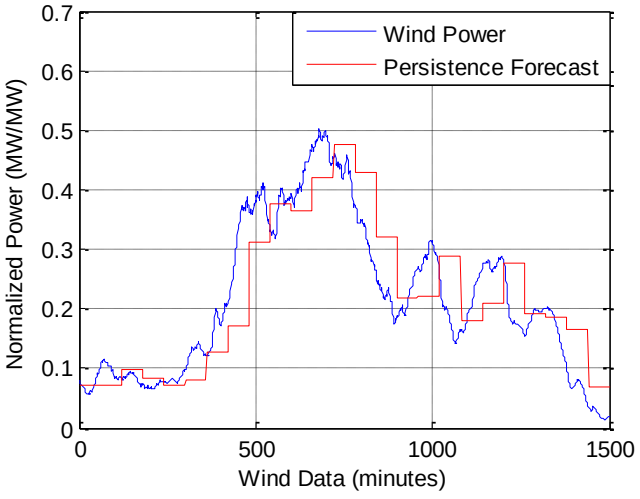


Fig. 1: Wind power data shown with persistence forecast model

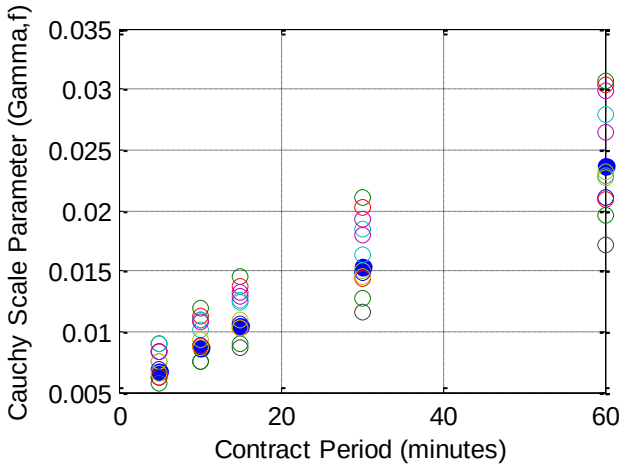


Fig. 2: Cauchy best fit scale parameters as a function of contract period length using persistence forecast and measurement period equal to contract period. Differently colored circles at each contract period represent each of the 13 wind plants.

3.2. Reliability Metric

An example of the probability density function is shown in Fig. 3. From this distribution, we wish to discuss the required energy – provided by storage or other generation sources – required for firming. The maximum energy requirement for firming (tails of the PDF) is typically driven by rare events where the required energy approaches the size of the wind plant for the contract period. A full power shutdown would be an example of this. Such events are typically handled by contingency reserves similar to conventional power plants. Therefore, for this analysis we discard 1% of the high & low events – the tails on the PDF – in effect defining achievement as successfully firming the wind plant for 98% of the contract periods during the simulation. The results of an analysis are therefore reduced to a single metric: the maximum energy required for a 98% probability of firming the wind plant for all contract periods. If using a 60 minute contract period, this is 175 contract periods per year that are not successfully firmed. If using a 10 minute contract period, this is 1051 contract periods not successfully firmed per year.

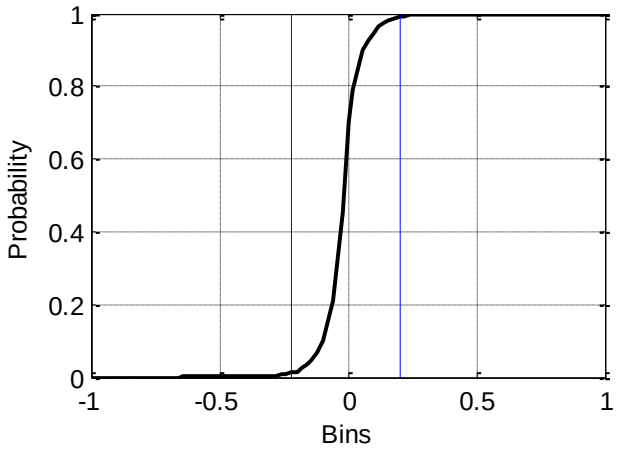


Fig. 3: Probability distribution histogram of persistence forecast errors for contract period of 60 minutes for one wind plant.

3.3. Sensitivity Simulations

Monte Carlo was utilized to simulate firming behavior. For each contract period, a "perfect forecast" was simulated: the energy generated in the contract period exactly equals the forecast energy. In addition, imperfect forecasts were simulated for a range of scale parameters. For a given γ, a set of 20 imperfect forecasts was synthesized by drawing a power deviation for each contract period from a Cauchy PDF. The plant was simulated and an energy requirement, $E = E(\gamma, F, C)$ was measured as described in section *3.2*. The simulated behavior assumes a "perfect controller" where the controller responded perfectly to minimize the firming energy given the forecast error. Therefore, results presented here should be considered a lower bound on firming energy requirements. All results here are presented as the ratio of the energy required (in MWh) to the size of the plant (in MW), producing a measure in time units of hours.

To facilitate comparison, it is often useful to ratio results to the reference γ_f calculated in section *3.1*. Scale parameter ratio is calculated by:

$$\gamma_R = \gamma/\gamma_f \qquad (2)$$

and energy requirement ratio by:

$$E_R = E(\gamma)/E(\gamma_f) \tag{3}$$

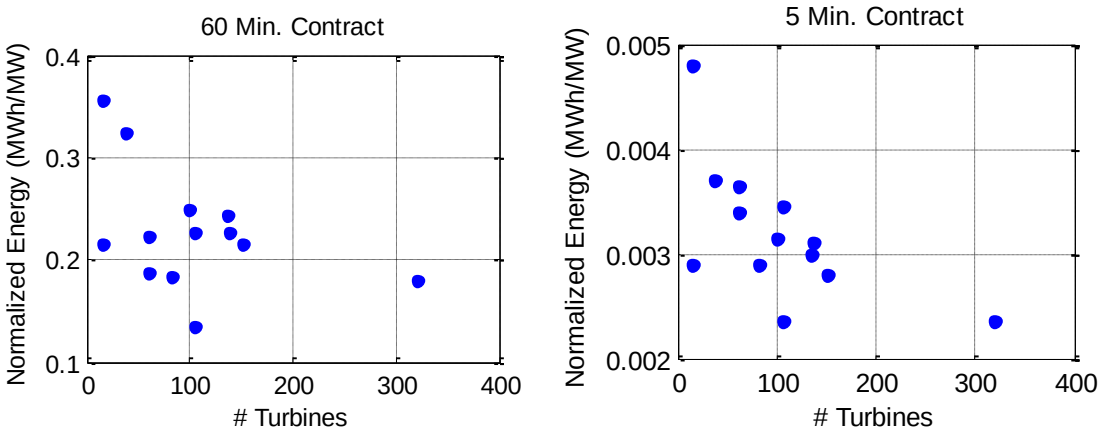
$E_f = E(\gamma_f)$ is the simulated energy requirement when $\gamma_R = 1$. For convenience, we plot the perfect forecast at $\gamma = 0$, although this is not strictly correct.

4. SENSITIVITY ANALYSIS

This section discusses the variation in required energy to firm a wind plant driven by plant configuration, forecast accuracy and measurement period.

4.1. Wind Plant Configuration

Fig. 4 shows the variation in firming energy requirement (E) for all of the plants, by turbine count. Turbine count is a measure of the plant's diversity and inherent smoothing over practical contract periods: In general, plants with more turbines tend to have smoother power profiles and require less energy to firm, relative to the plant size. However, significant variation exists between plants with similar numbers of turbines due to changes in turbine size, dispersion, and other factors.



(a) 60 Minute Contract Period (b) 5 Minute Contract Period

Fig. 4: Energy needed to firm contract period using a persistence forecast, plotted as a function of the # of turbines in each wind plant. Contract period and measure period are 60 minutes (left) and 5 minutes (right).

If variability in the output from individual units in the plant is uncorrelated, variability for the entire plant will decrease at $V_1/\sqrt{n}$, where V_1 is the variability in one turbine, and n is the number of turbines in the plant. Since firming energy is a form of variability, completely uncorrelated parameters should illustrate this behavior. For two plants with different turbine counts, variability should reduce as:

$$\frac{E_1}{E_2} = \frac{\sqrt{n_2}}{\sqrt{n_1}} \qquad (4)$$

where E_1 is firming energy of plant 1, E_2 is firming energy of plant 2, n_1 is the number of turbines in plant 1, and n_2 is the number of turbines in plant 2. Comparing the plants to the plant with the highest turbines count in Fig. 4, the 60 minute contract period has a mean ratio difference of 0.33 with one plant less than 0.1. The 5 minute contract period has a mean ratio difference of 0.14, with 6 plants having a ratio difference less than 0.1. This indicates that for 60 minute periods, output across the wind plant is significantly correlated. While some correlation exists for five minute contract periods, the correlation is much less significant. This effect explains why firming energy increases significantly as the contract period increases.

4.2. Contract Period

Fig. 5 shows firming energy when measurement period is equal to the contracted period. This study considered contract periods from five to sixty minutes. Sub-hour contract periods reduce the firming energy required as well as the range of required energy. It is logical to assume larger contract periods are harder to forecast, and will yield larger deviations between generated and forecasted power.

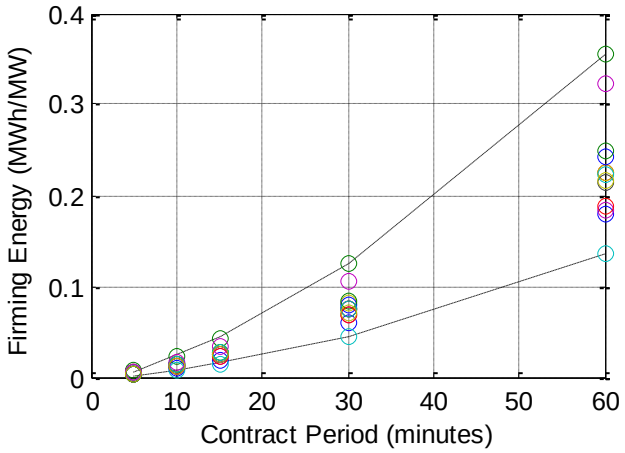


Fig. 5: Firming energy required for different contract periods when measure period is equal to contract period, using a persistence forecast. Differently colored circles at each contract period represent each of the 13 locations.

The black dashed lines are the upper and lower bounds of values, calculated by a second order polynomial fit of the plants with the highest and lowest firming energy requirements. These lines show that the range of firming energy required increases as contract period increases. Larger firming energy ranges demand larger energy storage reserves due to the unpredictability of firming energy requirements. Fig. 5 shows firming energy required for a 30 minute contract period is around 8-9% of plant size with a persistence forecast.

4.3. Measurement Period

A perfect forecast's generated energy exactly equals the forecasted energy for each contract period. However, despite the fact that no error exists over the contract period, as measurement period is reduced below contract period, energy deviations may exist in each measure period. For example, over a contract period of one hour, energy and deviations will be zero; but, in each 10 minute measurement period during the hour, there may be significant deviations between desired and actual power and energy. To successfully firm the wind farm, the firming energy supply must drive the error for every measurement period to zero. Table I shows the impact of measurement

period on the energy required for a perfect forecast. This represents the "best case" energy requirement: the standard for comparison. Table II shows the energy requirements when the forecast error has a scaling parameter of γ_f, for the same plant.

Table I
Firming Energy for Contract and Measure Periods for Perfect Forecast – $\gamma = 0$. Percent of rated plant size.

		Contract Period				
		60	30	15	10	5
Measure Period	5	4.1%	1.2%	0.4%	0.2%	0%
	10	4.0%	1.2%		0%	
	15	3.9%	1.2%	0%		
	30	3.8%	0%			
	60	0%				

Table II
Firming Energy for Contract and Measure Periods for Reference Forecast – $\gamma = \gamma_f$. Percent of rated plant size.

		Contract Period				
		60	30	15	10	5
Measure Period	5	42.9%	16.5%	6.3%	3.6%	0.8%
	10	42.8%	16.7%		2.0%	
	15	43.9%	16.6%	3.6%		
	30	43.4%	10.1%			
	60	28.5%				

Once measurement period decreases from the contract period, firming energy drastically increases. Both tables show firming energy not strongly dependent on measurement period once it is less than contract period, but a significant difference exists between "equal to contract period" and "not equal to the contract period." Results are similar for other values of γ. This result illustrates that firming is extremely difficult if intra-period measurements are made, and these should be avoided if firming is desired.

4.4. Forecast Accuracy

Forecast accuracy in the timescales evaluated does affect firming energy requirements, and deviation from a perfect forecast increases as γ of the simulated imperfect forecasts increases.

Fig. 6 shows increasing firming energy as the forecast deviates from the "perfect forecast" at 0.

Fig. 6 demonstrates the relation between forecast accuracy and firming energy required for each contract period for all 13 wind plants. The x-axis scale parameter ratio is calculated from the simulated imperfect forecasts (γ) divided by reference forecast (γ_f). The y-axis ratio of firming energy required in the range of imperfect forecast is calculated the same way: $E(\gamma)$ divided by $E(\gamma_f)$. These ratios in Fig. 6 show the relation of required firming energy and forecast accuracy are independent of wind plant size and configuration, and nearly linear. A 50% improvement in forecast accuracy represented by a 50% reduction in forecast ratio γ/γ_f, will reduce firming energy required by about 45%.

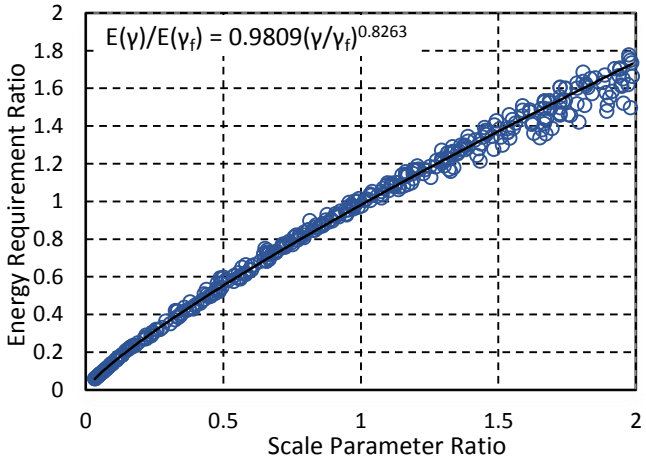


Fig. 6: Scale parameter ratio and energy requirement ratio normalized by the reference forecast. Data displayed is for all contract periods evaluated, when measurement period equals contract period.

5. CONCLUSIONS

This study is based on the Cauchy distribution, which is an excellent fit to the forecast error distributions. Other work, such as [24] and [26] suggests a normal distribution is a good fit to the forecast error. However, [25] shows that this is true only for large contract periods of 24 hours and

above which were not part of this scope. In view of the results, as contract period increases, Cauchy scale parameter increases and the error distribution approaches a normal distribution shape. For those analyses, a Gaussian distribution may be acceptable; but, for work here, a Cauchy distribution is the best fit.

Wind plants with higher turbine counts tend to have smoother power generations, but significant differences exist between plants with similar numbers of turbines. Turbine count is not an indicator of firming energy requirements. As contract periods decrease, wind plant correlation also decreases. Decreasing contract periods reduces firming energy required as well as the range of firming energy values across multiple wind plants. When measure period is less than contract period, even a perfect forecast has some deviation between generated and forecasted energy. To minimize firming energy, measurement periods need to equal contract periods.

Persistence forecasts have large deviations and high firming energy requirements. Improving forecast accuracy for any contract period or wind plant evaluated reduces the firming energy required by almost the same magnitude without the need for additional wind plant analysis. Improvements to forecast accuracy will reduce firming energy for any wind plant equal to the curve fit in Fig. 6.

This work looks at firming energy and the magnitude of forecast errors only. Further work should compare the differences of firming power requirements with firming energy for individual plants, and look at effects of the timing of forecasts in addition to magnitude.

PART II:

STATISTICAL ESTIMATION OF WIND FORECAST ACCURACY USING WIND SPEED AND FORECAST ERROR DISTRIBUTIONS

1. INTRODUCTION

Wind energy is a prominent energy resource due to its economic advantages and technical feasibility in generating electricity [27]. In some areas, a significant penetration of wind energy has already been installed and wind power production represents a high proportion of electrical load at peak production. Current trends in the wind energy penetration are expected to grow at a steady pace. In Europe, goals are in place for up to 37% of electricity to be generated by wind and solar by the year 2020 [27]. The unavoidable variability in wind velocities causes variable power generation uncorrelated with grid load. In a deregulated market, wind plants participate in the primary energy market where they commit specific power per unit time termed as "contract period" [28]. Unit commitments of wind plants have an element of uncertainty due to the deviation between forecasted and observed production, termed "forecast error". This may hinder the ability of wind plants to meet the firm contracts. Output of a wind plant being predicted extremely accurately by the forecasting model is a rare occurrence. Normally there exists marked deviation between predicted and actual output of wind plants. Natural smoothing due to aggregation of outputs of wind plants at a regional level is an effective technique to firm wind contracts [19]. As the physical distances between plants increases, correlation between the outputs of those plants decreases, resulting in more consistent total power independent of local wind behavior. However, regional aggregation may not always be effective since it is governed by correlation between the outputs of

those plants. Implementation of regional smoothing is subjected to challenges such as availability of suitable transmission infrastructure and flexible governance. Alternatives to regional smoothing of wind power have to be pursued. Studies have proposed local storage of excess energy to help meet these contracts, and this process is known as "firming" [19]. In this study, firming energy is defined as the energy required to meet a wind plant's contract. Firming power is the average power required to firm a plant's contract within the period of measure. A previous study has concluded firming requirements are minimized when measurement period is equal to contract period [29]. Contract firming has gained increased importance with renewable energy resources being integrated in the modernized electricity grids. In this report, firming requirements will refer to the difference between power and energy forecasted and produced at each wind plant. Negative firming requirements correspond to more power or energy produced than forecasted, while positive requirements correspond to less being produced than forecast. For a plant participating in local contract firming through storage, negative firming requirements require energy to be stored, while positive firming requirements require energy be supplied to the plant's output. As discussed in [26], firming requirements needed for individual or multiple plants are not easily determined.

Generation deficits in the range of a few kW to several MWs within a short time span frequently occur in bulk interconnections as experienced firsthand by system operators. These generation deficits may be associated with generation outages, line outages, deficits in renewable energy resources, etc. Generation deficits are typically compensated using operational spinning and non-spinning reserves (at the level of a balancing authority or independent system operator) or storage (supply locally) coordinated with the system operator. One storage technology used to counter variability of wind power plants is pumped hydro. However, limitations of environmental concerns, geographical availability for adding sites, and other demands on water supplies have led

to the increased importance of other contract firming technologies [30]. For success of other energy storage systems (ESS), a balance between charging and discharging times has to be maintained. Studies have been performed evaluating the feasibility of ESS. A reliability assessment of using ESS with wind plants in the Roy Billinton Test System (RBTS) has been performed [31]. The system reliability indices under different scenarios of expected wind energy and load with the ESS being controlled by the wind farm or the system operator is explored. For small islanded power systems, a significant improvement in the reliability with ESS implementation is presented [32]. The improvement in system reliability is inferred from the improvements in the loss of load expectation (LOLE) index. Also described is the synthesis of effective load carrying capacity (ELCC) for a wind plant with ESS, which indicates that the increase in ELCC is due to ESS. Hence, the importance of ESS has been well documented in the literature and is an accepted norm of achieving increased system reliability, especially with increased wind energy penetration.

When evaluating potential wind plant locations, developers benefit from accurate firming requirement predictions. To accurately estimate this, several years of wind or power observation data for a specific location are required [33]. If only wind observations are available, turbine power curves can be used to calculate potential power. To calculate firming requirements at a specific wind plant, the forecast power production is compared to the observed production, and the firming requirements are the deviation between forecast and observed. However there is a large problem with this method, i.e., wind and forecast data are not readily available to developers, and wind data requires several years of data acquisition to attain [34]. Frequently, developers must set up a dedicated measurement tower at the location to record wind speeds. Wind probability distributions, though, can be estimated from nearby areas or from atmospheric models and are easier to find. The analysis presented here will demonstrate a method for determining firming

requirements for wind plants of different sizes and geographic areas without the need for time series wind or power data. We will develop a statistical method using only a wind probability distribution and a forecasting error distribution that estimates firming requirements of 10 minute contracts. The data will be presented according to a 99.5% firming success rate. This means the resulting firming requirements are calculated for successful firming of 99.5% of contract periods, leaving 262 contract periods throughout the duration of one year not firmed. The reason 0.5% of contract periods are left unaccounted for is to limit the impact of shutdown to trip events during periods when the plant is running near capacity, as these are handled by contingency planning by the grid operator, and represent normal grid operations independent of the forecasting errors being studied here. If firming requirements were calculated for 100% of contract periods, firming requirements would be at the rated capacity of each plant and this analysis would be irrelevant. The required error distributions can be acquired from wind plants with similar characteristics to the potential plant of interest or from locations proximal to the one of interest. Optimally, forecast error distributions would be supplied by the same company that will operate the potential wind plant, assuring similar forecasting techniques will be used and the error distribution used for the analysis will be similar to what is seen during operation. For this analysis, forecast data was not available for any locations except one, so the results are compared with results using time series power measurements and a persistence forecast to confirm accuracy. At short time scales less than one hour, forecasts are almost entirely persistence, so persistence forecasts are an accurate representation of actual forecasts [14]. Lastly, persistence forecasts for varying lead times of one hour and greater are created at the location where actual forecasts are available. These persistence forecasts are compared with the actual forecasts to evaluate the accuracy of using purely

persistence forecasts for longer forecast lead times, where advanced forecasting technologies should normally provide a benefit over persistence.

The statistical method described here has 3 main parts:

1. Time series power observation and time series forecast data comparison
    1.1. Generate a forecast from time series wind or power data. If data is wind speed it must be converted to turbine power, which is accomplished with turbine power equations available from manufacturers [35] or more sophisticated methods developed by plant operators.
    1.2. Compare forecast and power observation data to determine forecast errors.
    1.3. Calculate power and energy required for stated success rate of contract firming.
2. Time series power and forecast error distribution comparison
    2.1. Characterize distribution of errors between forecast and observed power from step 1.
    2.2. Apply error distribution characterizations to the time series wind power data used in step 1 to create new observations. Determine the error between original power data and power data with applied forecast errors.
    2.3. Calculate power and energy firming requirements from error in step 2.2.
3. Power distribution and forecast error distribution comparison
    3.1. Generate a power probability distribution from observed power for each location. Sample the distribution to create theoretical power observations equal to the original data time span.
    3.2. Apply error distributions used in step 2 to the theoretical observations created from the power distributions. Determine error between the theoretical observations and error-adjusted theoretical observations
    3.3. Calculate power and energy firming requirements from error in step 3.2.

Method one is an ideal case where both observed power data and the respective forecasts are available. This allows the exact firming requirements to be calculated. However, having both sets of data is not very likely, and method two or three are more likely to occur. Method two requires wind or power observation data but not actual forecasts. Since many wind forecasters already use persistence forecast for short timescales, this method may only have marginal improvement or value over the first method. It is most likely that neither wind nor power data would be available to a developer. In this case, method three would be of great value and would allow firming requirements of a location of interest to be calculated.

2. DATA AND ANALYSIS

The data used for this analysis was provided by the National Renewable Energy Laboratory (NREL), the National Center for Atmospheric Research (NCAR), and Xcel Energy. The data from NREL consists of logs of power output for 11 wind plants in the western United States for the time span of three consecutive years, in one minute resolution, while the NCAR and Xcel Energy data is logs of both power forecasts and observations for one location for the time span of 2 consecutive years in 5 minute resolution. Although the power observations are in 5 minute resolution, the power forecasts are averaged over 1 hour. If the available data were wind speeds it would have first been converted to potential wind power based on the turbine model at each location, but wind power observations were available so this was not necessary. For all analyses, energy and power are normalized to each wind plant's rated generation capacity in MW.

Depending on the market structure, firm contracts are obtained from the market data ahead of when the actual contract period begins. Thus, wind plants use the outputs of forecasting models for commitments which are associated with errors that are a function of the forecast horizon. Wind

plants participate in markets with time spans varying between commitments one day ahead, to five minute commitment periods. If the wind plants are unable to meet the commitment for a contract period, energy needs to be supplied via reserves or spot market [28]. The analysis presented here calculates the power and energy needed to firm contracts in these cases [25], [36].

2.1. Method 1

The first step was to generate a forecast based on the supplied data. Persistence forecast is the preferred forecasting method for short forecast windows and is commonly used as a reference with which to compare other forecasting methods. It was chosen to represent a reference forecast from which forecasting errors were calculated for this study [22], [23]. Longer lead time forecasts are not purely persistence, but for short time spans the actual forecasts are heavily weighted by persistence and are almost identical [13]. Persistence forecasting is preferred on account of its simplicity of application and accuracy over short periods. It can be described as a perfect forecast for the current period with a time delay to a future period. It works by having forecasted power for any contract period being equal to average production of the previous period with no delay. However, in practice when persistence forecasts are applied there is some delay. A one-period delay was chosen in this study, so the forecast for each period is actually the average of the single period two periods prior to the one being forecasted. The persistence forecast method is illustrated in Fig. 7. A ten minute contract period and forecast lead time was chosen as good representation of forecasting and contracting periods as industry-leading electricity markets are moving to 10 or 15 minute scheduling.

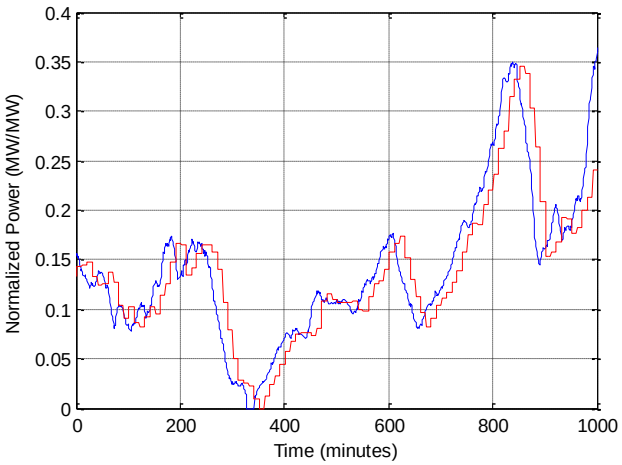


Fig. 7: Wind power data with persistence forecast model and 10 minute contract period. Blue line is actual power, red line is persistence forecast.

Ideally, wind forecasts and wind or power observations would also be available for analysis. These would allow historic firming requirements to be calculated. However, actual forecast data was not available so a realistic forecast had to be simulated. Using the reference forecast, power errors, e_p, were computed by subtracting the forecasted power, Φ, from the observed power, p, for the i^{th} minute of a contract.

$$e_{p,i} = p_i - \Phi_i \qquad (5)$$

Forecast power each contract period, Φ, is equal to the average observed power two periods prior calculated by dividing the sum of observed power by the length of the contract period. Forecast power is the constant throughout each contract period.

$$\Phi = \frac{\sum_{i=1}^{N} p_i}{N} \qquad (6)$$

N is the number of minutes in each contract period. Forecast power is constant throughout the contract period. Since Φ is the average power two periods prior, the first two contract periods do

not have forecasts and are not included in the analysis. Likewise, energy values for the respective contract periods were computed by subtracting the observed energy each contract period by the forecast energy. Firming requirements for each plant were calculated every contract period for each year. Fig. 8 plots the distribution of forecast power errors for one plant for each year, and Table III displays the upper and lower bounded maximum and minimum power and energy errors for the same plant. Table IV displays the same information for a second, smaller capacity plant. All error values are percent of wind plant capacity.

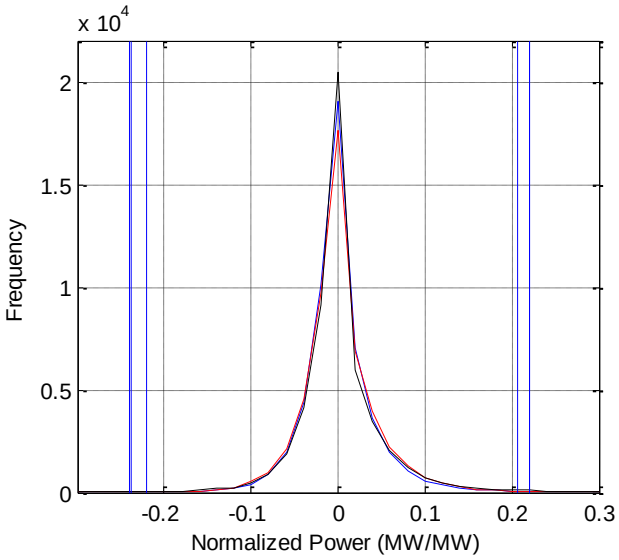


Fig. 8: Distribution of forecast errors for each year for one wind plant. Vertical lines are the upper and lower bounds. Positive values mean power must be supplied for firming, negative values means power must be absorbed.

Table III
Power firming requirements for a large wind plant in percentage of capacity of the wind plant. Max and min power are in units of MW/MW. Max and min energy are in units of MWh/MW. Positive numbers require energy to be supplied to firm a contract while negative numbers mean more energy is produced than predicted.

Plant 1	Max Power (MW/MW)	Min Power (MW/MW)	Max Energy (MWh/MW)	Min Energy (MWh/MW)
Year 1	20.6%	-22.0%	3.4%	-3.7%
Year 2	22.0%	-24.0%	3.6%	-4.0%
Year 3	22.0%	-23.8%	3.6%	-4.0%

Table IV
Firming requirements for a smaller rated capacity wind plant. This table shows the same information as Table III for a smaller rated capacity plant.

Plant 2	Max Error (MW/MW)	Min Error (MW/MW)	Max Energy (MWh/MW)	Min Energy (MWh/MW)
Year 1	31.2%	-34.2%	5.2%	-5.7%
Year 2	30.2%	-31.0%	5.0%	-5.2%
Year 3	31.6%	-33.0%	5.2%	-5.5%

In Table III and Table IV, Max Error is the largest deviation between predicted and produced power during a contract period when the power produced is less than predicted. Min Error is the largest deviation between predicted and produced power during a contract period when the power produced is more than predicted. Max Energy is the largest amount of energy deviation in one contract period between predicted and produced when less energy is produced than predicted. Min Energy is the largest deviation during one contract period of predicted and produced energy when more energy is produced than predicted. The results from this method are an ideal case where both wind power and wind forecast data are available, which allows the exact firming requirements to be calculated. As seen in Fig. 8 and similar with other locations, forecast error distributions are primarily governed by the forecasting timescale regardless of other parameters [24]. The next method uses time series wind power data and a forecast error distribution to estimate firming requirements. While it is a more general estimation, it is more valuable to developers who may not have access to wind power forecasts.

2.2. Method 2

The next statistical method for estimating firming requirements involves time series power data, and distributions of the forecast error from method 1. Rather than creating a persistence forecast as in method 1, method 2 uses the forecast error distributions to estimate firming requirements assuming a similar forecast error distribution is seen in practice. Observations from

the data set utilized here, and in agreement with [24], show the distributions for similar contract timescales are similarly shaped and can be represented by a Cauchy distribution. First, Cauchy distributions were fitted to each year's cumulative probability distribution of forecast power errors for each location, yielding a unique γ and forecast error distribution fit for each case. Each distribution is limited at -1 and 1, which correspond to an error 100% of the rated capacity of the wind plant. Next these error distributions are applied to the time series power observations to synthesize forecasts that can be compared with the observations to calculate firming requirements.

2.2.1 Error Modeling

Forecast errors are used to calculate firming requirements. They can be modeled using standard distributions such as Cauchy, normal, and beta [24]. Selection of a suitable distribution to model the forecasting error is dependent on the forecast horizon. For the data sets and contract period used in this analysis, it has been observed that a Cauchy distribution fits the errors better as compared to the normal distribution which has been employed for longer contract and forecast periods [25]. Fig. 9 shows a Cauchy distribution fit to forecasting power errors for a persistence forecast and a ten minute contract.

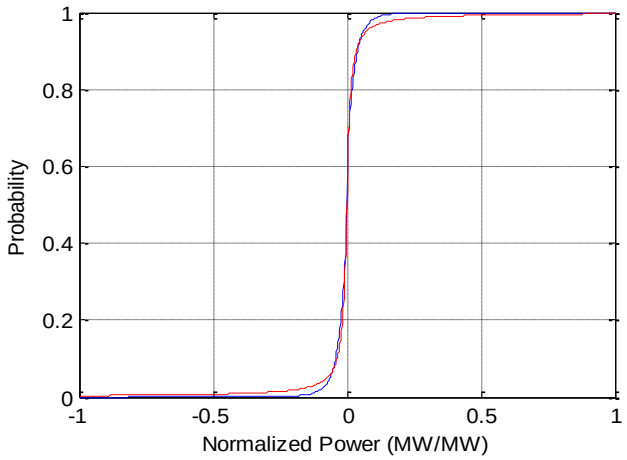


Fig. 9: Cauchy cumulative distribution (red) fit to persistence forecast errors (blue) for a ten minute contract period.

A Cauchy distribution was fit to the forecasting error distribution for each wind plant and year of data. The choice of Cauchy distribution is also supported by the persistence error forecasts as analyzed in [24]. In a Cauchy distribution, the parameter γ defines the shape of the probability distribution function (PDF). It is equal to the half-width at half-maximum, or half the interquartile range. A smaller γ value produces a taller and narrower PDF than a larger one. The results from the data and analysis in [29] report Cauchy distributions for forecasting errors corresponding to the contract period of 10 minutes have γ values between 0.0075 and 0.0125. In that study larger values of γ indicate larger magnitude forecast errors.

Once the error distributions are created, the average power observation of each contract period is determined utilizing the same process as in method 1. The forecast error distribution is randomly sampled the same number of times as the number of contract periods, and each period's average power observation, *P*, is multiplied by the randomly sampled error, *e*, to produce a new vector of observations.

$$\mathrm{P}' = \mathrm{P} * \mathrm{e} \qquad (7)$$

This new vector, P' is considered to be a new set of power forecasts. For further calculations it is treated the same as the persistence forecast in method 1. The new forecasts are compared to the observations following equation 5, and the error between observed and forecasted power and energy each contract period is calculated. The differences in firming requirements between the results in method 1 and method 2 are shown in Table V and Table VI. The results are calculated for the same locations with the same data.

Table V
Difference in firming requirements from the plant in Table III, in percent of rated plant capacity. Requirements are calculated from a generated forecast based on power observations and the persistence forecasting error in section 2.1. Positive numbers represent larger absolute max and min errors from method 2.

Plant 1	Max Power (MW/MW)	Min Power (MW/MW)	Max Energy (MWh/MW)	Min Energy (MWh/MW)
Year 1	4.9%	4.1%	0.8%	0.7%
Year 2	6.3%	5.4%	1.3%	0.9%
Year 3	1.3%	-0.2%	0.3%	-0.2%

Table VI
Difference in firming requirements for the smaller plant described by Table IV. Forecasts are generated from power observations and persistence forecast error distributions.

Plant 2	Max Power (MW/MW)	Min Power (MW/MW)	Max Energy (MWh/MW)	Min Energy (MWh/MW)
Year 1	1.2%	-1.4%	0.2%	-0.2%
Year 2	3.9%	3.2%	0.6%	0.5%
Year 3	1.4%	0.6%	0.3%	0.1%

2.3. Method 3

The final method of estimating firming requirements eliminates the need for wind or power observations. Wind distributions can be converted to power distributions with turbine power curves [35]. With only a wind probability distribution for a specific location and a forecasting error distribution, firming requirements are estimated. This method is valuable for evaluating potential wind plant locations because it eliminates the need for wind or power observation data which are not always easily obtained. Fig. 10 displays the power distribution for one location for one year of power data.

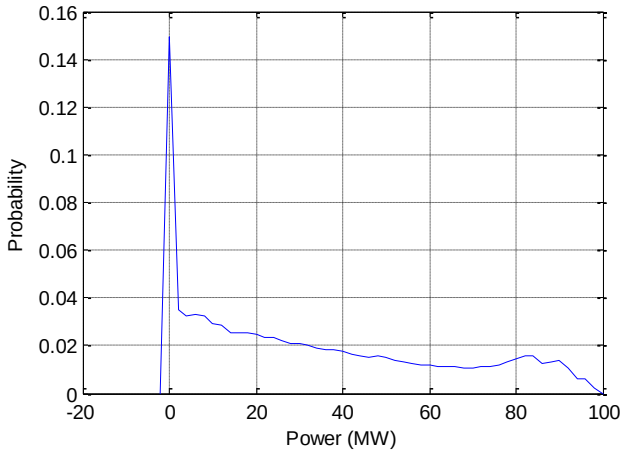


Fig. 10: Wind power distribution for a wind plant with a capacity near 100 MW.

Implementing this estimating method is similar to method 2 (section 2.2), but there is an additional initial two steps required. Step one is to convert a wind speed distribution to a wind power distribution. For this study, the available data was provided in terms of time-series of power. Therefore, no conversion from wind to power is necessary. The second step is to randomly sample the wind power distribution the same number of times as there are contract periods in a year to synthesize wind power observations. Next, the same forecast error distributions from method 2 are randomly sampled an equal number of times to create an average forecast error corresponding to each contract period. These two steps result in newly created observations and forecasts for each contract period of the year. The newly created observations correspond to hypothetical average power observed each contract period. These synthesized observations are multiplied by the contract forecast errors following equation 7 to produce error-adjusted observations corresponding to each contract period. The error-adjusted observations are treated as forecasts for firming requirement calculations. As the final step, the resulting forecasts (error-adjusted observations) are compared to the originally created observations to calculate error according to equation 5. The difference between observed and forecast power and energy each contract period is calculated, and

these differences are the firming requirements for each location, calculated from only power and forecast error distributions. The firming requirements for this estimating method for each location compared to the results of method 1 are displayed in Table VII and Table VIII.

Table VII
The same wind plant is shown as in Table III and Table V. Difference in firming requirements in percent of rated plant capacity between the final estimation method where only a wind and forecast error distributions are used and method 1. Positive numbers represent larger absolute max and min errors.

Plant 1	Max Power (MW/MW)	Min Power (MW/MW)	Max Energy (MWh/MW)	Min Energy (MWh/MW)
Year 1	8.7%	7.3%	1.5%	1.2%
Year 2	6.0%	4.6%	1.0%	0.8%
Year 3	0.1%	-0.5%	0.1%	-0.1%

Table VIII
Difference in firming requirements for the same plant as Table IV and Table VI, between method 3 and method 1, calculated according to method 3 described by section 2.3.

Plant 2	Max Power (MW/MW)	Min Power (MW/MW)	Max Energy (MWh/MW)	Min Energy (MWh/MW)
Year 1	0.7%	-2.3%	0.1%	-0.4%
Year 2	4.1%	3.1%	0.7%	0.5%
Year 3	1.6%	0.6%	0.3%	0.1%

On the next page Table IX displays a summarized comparison of the results from each method. Firming results are shown as percent of rated plant capacity for each year at 2 different locations. Both maximum and minimum power and energy requirements are displayed. Positive numbers mean firming power or energy must be supplied, negative numbers mean it must be curtailed.

Table IX
Summary of firming requirements calculated using each of the three methods.
Results are shown as percent of rated plant capacity. Results for two plants shown.

Power (MW)

Plant 1		Method 1	Method 2	Method 3
Year 1	Max	20.6%	25.5%	29.3%
	Min	-22.0%	-26.1%	-29.3%
Year 2	Max	22.0%	28.3%	28.0%
	Min	-24.0%	-29.4%	-28.6%
Year 3	Max	22.0%	23.3%	22.1%
	Min	-23.8%	-23.6%	-23.3%

Energy (MWh)

Plant 1		Method 1	Method 2	Method 3
Year 1	Max	3.4%	4.2%	4.9%
	Min	-3.7%	-4.4%	-4.9%
Year 2	Max	3.6%	4.9%	4.6%
	Min	-4.0%	-4.9%	-4.8%
Year 3	Max	3.6%	3.9%	3.7%
	Min	-4.0%	-3.8%	-3.9%

Power (MW)

Plant 2		Method 1	Method 2	Method 3
Year 1	Max	31.2%	32.4%	31.9%
	Min	-34.2%	-32.8%	-31.9%
Year 2	Max	30.2%	34.1%	34.3%
	Min	-31.0%	-34.2%	-34.1%
Year 3	Max	31.6%	33.0%	33.2%
	Min	-33.0%	-33.6%	-33.6%

Energy (MWh)

Plant 2		Method 1	Method 2	Method 3
Year 1	Max	5.2%	5.4%	5.3%
	Min	-5.7%	-5.5%	-5.4%
Year 2	Max	5.0%	5.6%	5.7%
	Min	-5.2%	-5.7%	-5.7%
Year 3	Max	5.2%	5.5%	5.5%
	Min	-5.5%	-5.6%	-5.6%

2.3.1. Forecast Comparison

Once these statistical estimation methods were complete, the last step was to compare persistence with actual forecasts for lead times where persistence should be less accurate than the actual forecasts. Method 1 (section 2.1) was repeated with the wind plant which had actual forecast data available. Power forecasts were available in 1 hour resolution, so a persistence forecast was constructed for 1, 3, 6, 12, and 24 hour lead times. These forecasts were compared to the power observations, and persistence forecast error distributions were created. Next, the actual power forecasts for the same lead times were compared with the power observations, and actual forecast error distributions were created. Firming requirements for the real forecast error distributions were calculated based on a 99.5% success rate. The results are compared in the next section.

3. RESULTS

The locations studied have various capacities and turbine counts. The number of turbines ranges from 15 to 320, and the MW capacity ranges from 20 to 300. No obvious patterns governing the shape of the wind power distributions could be distinguished between them. There was also no correlation between available wind plant characteristics and persistence forecasting error distributions. The most influential factor of forecast error distributions appears to be the forecast timescale, which agrees with [24], [25].

Fig. 11 displays a comparison of each year for two plants, for each of the three firming requirement estimating methods. Plant one is a larger wind plant in both capacity and number of turbines when compared to plant 2. Firming requirements shown in Fig. 11 are largest absolute values of both negative and positive requirements. Firming requirements for the larger plant

fluctuated more between years than the second. Additionally, the larger plant's persistence forecasting error distributions fluctuated more between years compared to a smaller plant. Plant 2 had Cauchy curve γ values for each year within 1% of each other, while plant 1's Cauchy curve γ values were only within 10%. The largest percentage difference of Cauchy γ values between all wind plants studied was 49%.

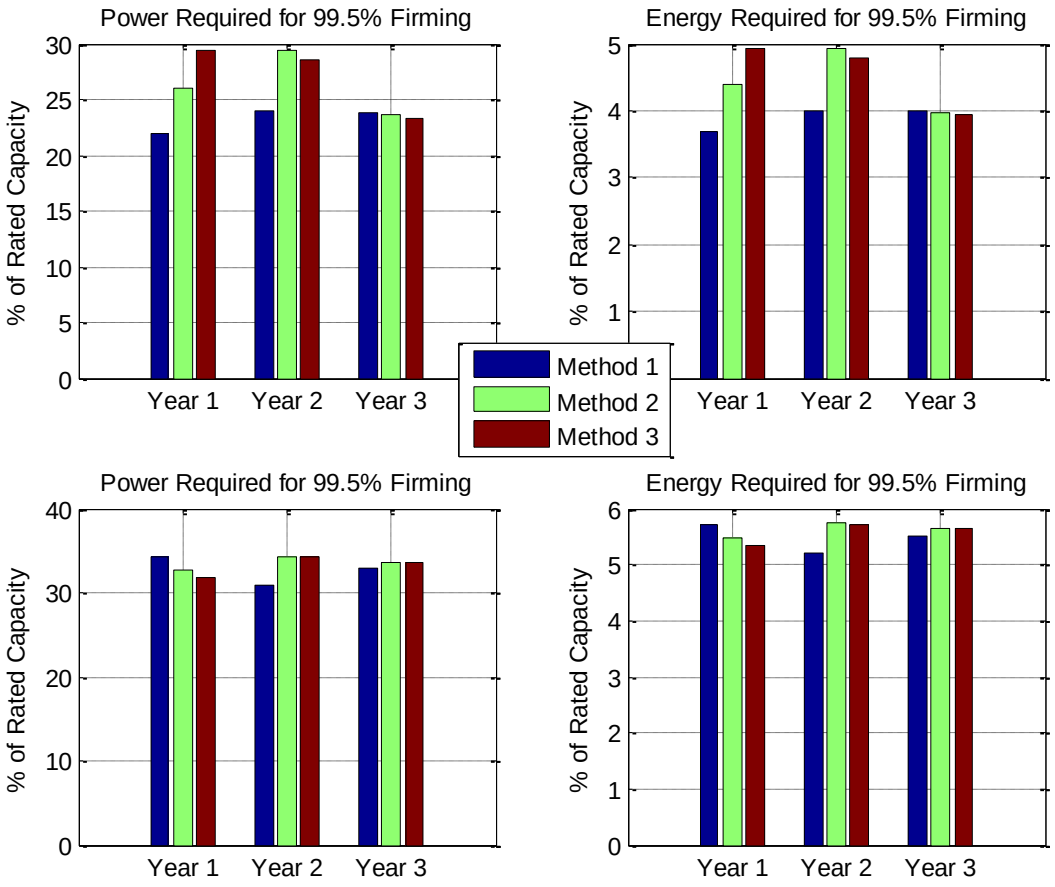


Fig. 11: Largest absolute power (left) and energy (right) firming requirements for each year of two wind plants, to successfully firm 99.5% of contracts. Top row are results from plant 1, bottom row are results from plant 2. Plant 1 is both larger in capacity and number of turbines than plant 2.

The firming requirements calculated with the described analyses were very similar to each other. Tables 3-8 compare each individual method. While there is significant variation between

plants, there is much less variation between years at the same wind plant. Methods 2 and 3 both estimate conservatively, or slightly overestimate firming requirements. Firming power required due to forecasting errors calculated in method 3 (section 2.3) is within 10% of actual in the worst case, and within 1% in the best case. Not surprisingly, power and energy follow a similar pattern. These results show that firming requirements can be successfully calculated from wind and forecasting error distributions without requiring actual forecasts or time series data.

In addition to this analysis, forecasting error distributions of long lead times with both persistence and actual forecasts were compared. Compiling error distributions for both persistence and actual forecasts compared to observations, firming requirements for a 99.5% success rate were calculated. It was expected that actual forecasts would be more accurate at these long lead times than persistence forecasts. This is confirmed by the results in Table X. Table X shows the difference in firming requirements between persistence and actual forecasts for forecast lead times of 1, 3, 6, 12, and 24 hours. Absolute value of actual firming requirements are subtracted from persistence so positive numbers represent the reduction, or improvement from persistence to actual forecast firming requirements. The results are from a wind plant not used in earlier analysis.

Table X
Reduction in firming requirements between actual and persistence forecasts for several different lead times in percent of rated capacity of the wind plant.

Lead Time (hours)	Max Power	Min Power	Max Energy	Min Energy
1	24.6%	29.6%	2.1%	2.5%
3	23.0%	17.0%	1.9%	1.4%
6	25.2%	20.4%	2.0%	1.7%
12	18.6%	20.6%	1.5%	1.7%
24	16.4%	23.0%	1.3%	2.0%

Comparing forecast methods doesn't show a relationship between forecast lead time and improvement from persistence. Firming power requirements for actual forecasts are much less than the firming requirements if purely persistence forecasts were used. This result confirms the advanced forecasting system employed by NCAR is more accurate than persistence forecasts for wind and power forecasting at times of 1 hour and greater.

4. CONCLUSIONS

It has been shown that firming requirements for 10 minute contract periods can be accurately estimated to within 10% using a wind or power distribution and forecasting error distribution. Firming requirements for 99.5% of contracts throughout the year are reported. If a plant participates in local contract firming through storage, this results in 262 contracts where the contract was not firmed using available local storage. Increasing the percentage of firm contracts will increase the reported firming requirement amounts. These results would be close to the results seen in actual operation if the same contract periods were used. If different contract periods or forecast timescales are used the results will be different and the scope of this work does not investigate other combinations of operating parameters. While method three still requires access to a wind forecast distribution, as stated earlier these distributions are similarly shaped for each contract period timescale regardless of other factors and can at worst be estimated.

Next, by comparing persistence forecasts with actual forecasts, it was shown that persistence forecasting is effective below an hour lead time, but longer times require advanced forecasting methods, such as the forecasting provided by NCAR. The actual forecast results confirm the advanced forecasting available from NCAR is more accurate than persistence forecasts at these forecast lead times. Table X compares firming requirements of persistence and actual

forecasts. These values show the reduced firming requirements necessary when forecasts are more accurate, and lead to the assumption that increasing forecast accuracy will decrease required contract firming, allowing higher renewable energy market penetration. Firming energy improvements are small, about 2% reduction in MWh when normalized to plant capacity. However firming power improvements are much larger. Firming power is reduced by an average of 22% of plant capacity by using actual forecasts instead of persistence at lead time 1 hour and greater.

If available, historical observation and forecast data for a particular location are always preferred. However, frequently they are not available and distributions that characterize different locations must be implemented. After following this statistical analysis method, the calculated information can be used in many advantageous ways. Developers can use the estimations to size required energy storage systems if a plant is participating in local contract firming, or to inform potential operators of the expected deviation between forecast and produced power and energy. This can be a valuable tool to save time and money in future development projects.

PART III:

ADVANCED WIND RAMP DETECTION METHODLOGY USING BIMODALITY STATISTICS

1. INTRODUCTION

Installed wind capacity in the United States has been following an upward trend, with over 40,000 MW of installed capacity at the end of 2010 [37]. This trend is expected to continue worldwide throughout the next several decades [5]. The continually increasing prominence of wind power throughout the world magnifies the inherent problems with high penetrations of wind power, e.g. its stochastic nature [7]. This highly variable nature makes accurate forecasting techniques of power output from generation systems critical [13]. Large changes in wind speed, called "ramps", can cause large changes in wind power which are difficult for power system operators to manage [9]. Wind ramps consist of two factors, a specific magnitude change and a specific time. Correctly forecasting timing but not magnitude results in larger errors than expected at the expected ramp time. If magnitude is underestimated, energy reserves may be unavailable to provide additional energy when needed. Correctly forecasting magnitude but not timing creates the same situation and errors, except errors occur at unexpected times. Accurately forecasting both ramp timing and magnitude is critical to improving ramp forecasts, but requires complicated statistical analyses [9], [14]. Ramp events are highly influenced by regional geography and weather patterns, and forecasting them is an area of high interest to commercial providers of wind power. Some operators have partnered with national labs to research ramp behavior and develop sophisticated wind prediction systems [9]. Separating the individual effects of timing and

magnitude from overall forecasting errors would quantify which factor contributes more to overall forecasting errors. Focusing resources on the factor of greater importance would result in a better overall improvement in ramp forecasting accuracy.

Ramp events are especially problematic for wind plants committing to specific power output per unit time, termed "contract periods". Forecasting "errors" are power or energy deviations between forecasted and observed wind power output during a contract period. Fig. 12 explains the difference between power and energy forecasting errors.

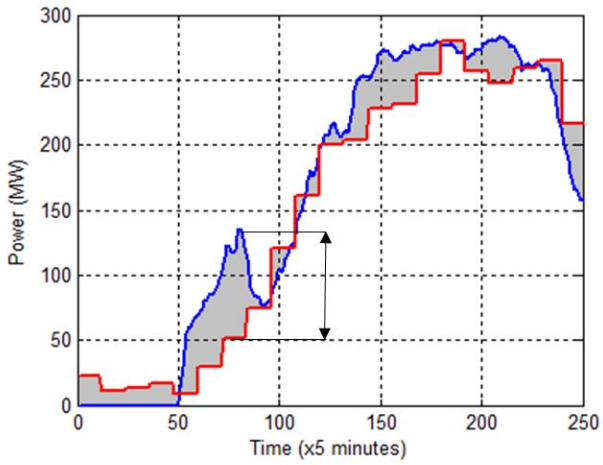


Fig. 12: Schematic representation of power and energy forecast errors. The blue line is actual power data and the red line is the corresponding forecasts. Power error is the difference between the observed and forecasted power at each point. The maximum power difference for the pictured period is shown by the arrows. Energy error is the sum of all power errors, or the area of the shaded region in the figure.

To lessen the effect forecast errors, "contract firming" is a solution where power is supplied through system or local reserves when forecast errors are large, e.g. spinning reserves or energy storage systems that can deliver immediate power [38]. Energy reserves diminish the environmental benefits of renewable energy technologies due to low efficiencies, emissions, material and energy requirements for manufacturing, capital investments and maintenance costs. [8]. Improving forecasts is a universally accepted solution to reduce supplemental energy

requirements and carbon emissions of wind power, while increasing its potential for high penetration [9].

Currently, forecast lead times less than one hour are predominantly persistence, and resemble pure persistence forecasts more as these lead times decrease [13]. As forecast lead times increase beyond one hour, advanced forecasting techniques such as the National Center for Atmospheric Research's (NCAR) Dynamic Integrated Forecast System (DICast) and Variational Doppler Radar Assimilation System (VDRAS) become more accurate than persistence [14]. These models weight future predictions with previous history and current atmospheric data to tune their outputs, which are continuously changing wind power production forecasts. Wind plant operators use these forecasts to contract specific production per unit time duration before contracted periods start. Separating the influence of timing and magnitude on forecast errors will allow better prioritization of wind forecast research and advancement efforts. One widely accepted ramp detection method is to compare change in power output over a specific time to a percent of operating capacity of the location. If the change in output is larger than this threshold, it is identified as a wind ramp event.

The method constructed in this paper is a statistical approach to determine if and when ramp events occur. A similar method has been used by the authors to detect electrical devices and identify unique signatures through current draw, and start / stop transients. The ramp detection algorithm must detect ramp events in clean and noisy data while filtering out trends that do not qualify as ramp events. It will detect both if a ramp occurs, and estimate a consistent temporal center of the event. The ramp center is defined as the point of highest rate of change in power during the event. After ramp detection we will compare forecast data for the same points in time

to determine the individual influences of timing and magnitude on forecast power and energy errors.

In this paper we will discuss several ramp detection methods and develop a weighted detection coefficient. Next we will calculate power and energy errors between detected ramps and corresponding forecast data, then optimize forecast timing and recalculate these values with the forecast adjusted to this timing. Last, timing and magnitude influences on the forecast errors are extracted and analyzed.

2. DETECTION METHODOLOGY

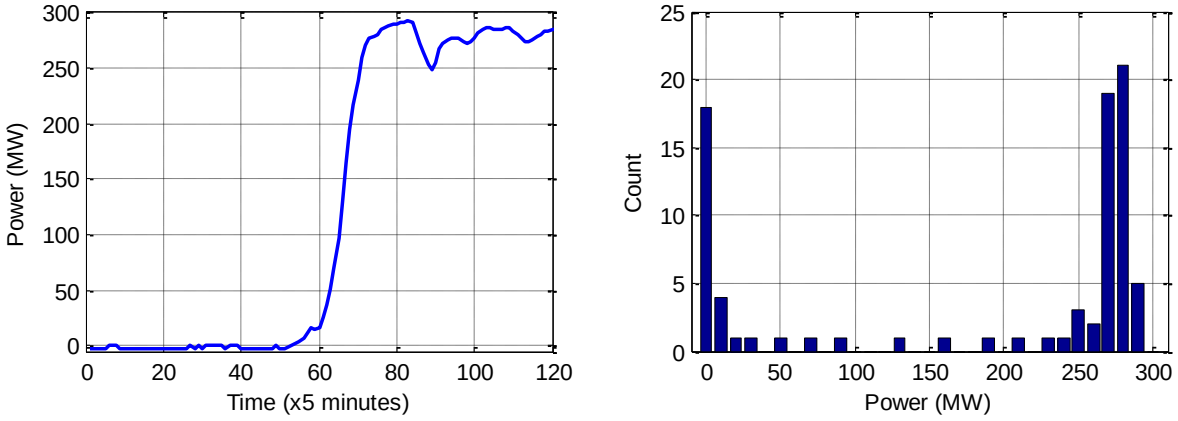
2.1. Data

The data used for this analysis was provided through Xcel energy and NCAR. It consists of wind forecasts made available through NCAR's advanced wind forecasting models, and wind power output observations for a 300MW capacity wind plant in northern Colorado from November 2009 until July 2011. Forecasts were hour ending, for 3 hours ahead, in 1 hour intervals. *Hour ending* means the forecasted value is predicted for the hour ending at that time. Power observations were available every 5 minutes, each an average of 10 second observation data over the 5 minute period. Out of 111780 available observations 6787 were missing, and out of 9315 hour long forecasts, 1257 were missing. Many missing values were single points, so these points were interpolated to the nearest available neighbors. If more than 4 consecutive points were missing there was no interpolation for these points. Any missing forecasts or observations were excluded from ramp detection and power and energy error calculations.

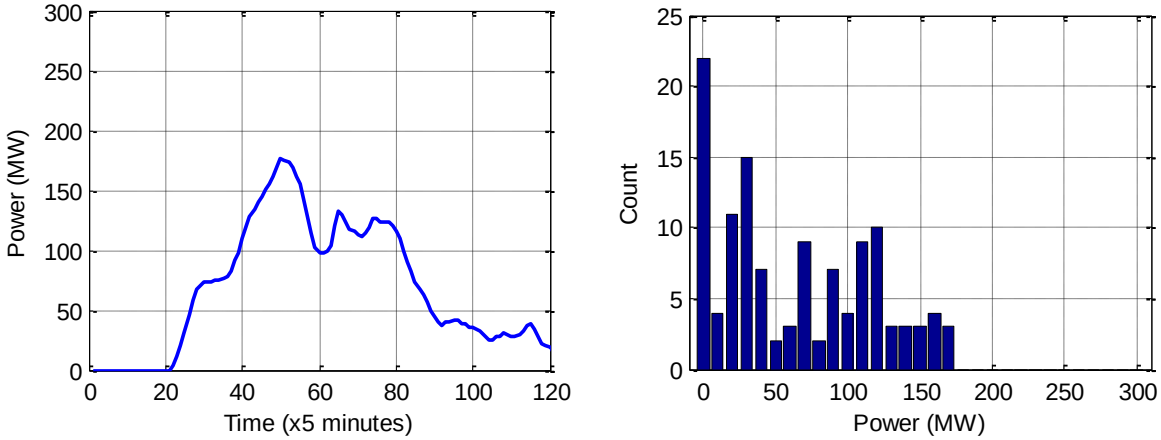
2.2. Detection Methods

The simplest method of detecting ramps is monitoring the maximum change in power output per unit time, and any difference above a threshold value signifies a wind ramp. While this method is easily implemented, it is prone to false positives. Any short spike in wind regardless how short or in what direction, or erroneous data points can increase the difference between minimum and maximum levels to above the threshold. A very short spike in power should not be considered a true ramp event since it would not have a significant impact on the forecast energy error. If an obvious wind ramp occurs but the maximum change is just below the threshold value, it will not detect it as a ramp event. Instances like this are still of interest, but because they fail to meet the minimum threshold value, a ramp is not detected. A more intelligent method of ramp detection is required to filter out these falsely triggered ramps from true, sustained ramp events and to detect both large and small ramps. If a ramp event occurs and the data window is several times the duration of the event, the ramp will resemble a step change in power output. If a pure step change occurs in the middle of a data window and that data is shown as a histogram, the histogram will have two peaks, each one corresponding to the operating level before and after the step. This histogram shape is known as a bimodal distribution and is the basis of the ramp detection method that is constructed in this paper. An example section of data and corresponding histogram for a ramp event and no ramp event is shown below in (b) Non-ramp Event

Fig. 13. The algorithm iterates through the observation data, updating the histogram each step. The process of adding and removing data for each iteration is described in detail later. In the figure the top graph produces a histogram with two peaks corresponding to the horizontal lines, while the bottom has a histogram with one peak.



(a) Ramp Event



(b) Non-ramp Event

Fig. 13:Comparison of histograms from a ramp (top) and non-ramp event (bottom). Top: Example showing a ramp event taking about 100 minutes and the corresponding bimodal histogram. Bottom: Example of a power transient over a longer time and not qualified as a ramp, producing a unimodal histogram.

Differentiating between bimodal or unimodal distributions can be evaluated by calculating *bimodality strength*. The bimodality strength of a distribution represents how closely the distribution resembles a purely bimodal distribution. There are three widely accepted methods to calculate this value. These are the bimodality coefficient (BC) [39], Hartigan's dip test (HDT) [40], and Akaike's information criterion (AIC) [41]. Several works have been published that compare either BC and HDT, or all three methods [42], [43], [44]. Bimodality coefficient and HDT

detect ramps more often than AIC, and HDT fluctuates less than BC from small changes in the histogram shape, resulting in less false positives. Although there are many comparisons, none of these attempt to create a composite, weighted value, nor do they apply to detecting changes in the data driving the histograms from which bimodal statistics are calculated. AIC is less desirable for this work because doesn't calculate bimodality value directly, rather it calculates the goodness of fit between data and a distribution. This analysis will focus only on BC and HDT.

2.2.1 Bimodality Coefficient

Bimodality coefficient, b_c, is calculated as follows [39]:

$$b_c = \frac{s^2 + 1}{k + \frac{3(n-1)^2}{(n-2)(n-3)}} \quad (8)$$

Where n is the number of points in the sample, s is the skewness, and k is the kurtosis of the sample. The bimodality coefficient ranges from 0 to 1, where a lower value corresponds to a more bimodal distribution. For the detection algorithms here, it is more convenient to consider a "bimodality strength" measurement defined as:

$$\mathrm{b_c}' = 1 - \mathrm{b_c} \quad (9)$$

Which causes b_c' to increase as the distribution becomes more bimodal. Recall that variance, denoted by s, is the second moment about the mean, μ, and is described by the equation:

$$s = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2 \quad (10)$$

Standard deviation, denoted by σ, is calculated as the square root of variance, s:

$$\sigma = \sqrt{s} \tag{11}$$

The 3rd and 4th moments about the mean, called skewness and kurtosis, respectively, are calculated by:

$$s = \frac{\sum_{i=1}^{n}(x_i - \mu)^3}{\sigma^3} \tag{12}$$

$$k = \frac{\sum_{i=1}^{n}(x_i - \mu)^4}{\sigma^4} \tag{13}$$

Using these equations the bimodality strength value is calculated from the bimodality coefficient.

2.2.2 Hartigan's Dip Test Statistic

Hartigan's dip test value is calculated and output from a MATLAB function made available to the public by Mechler [45]. These functions require a distribution of the data of interest as an input. The Hartigan dip test value ranges from 0 to 0.25, and a higher value corresponds to a more bimodal distribution. To match the range of the bimodality strength value, a "HDT strength" value is calculated by:

$$h' = 1 - 4 * h \tag{14}$$

where h is the Hartigan dip test value and h' is the HDT strength parameter which describes how bimodal the distribution is. Both b_c' *and* h' are calculated from a continuously updated histogram as the ramp detecting algorithm iterates through the data. The ramp detecting algorithm analyzes 30 power observation data points at once, starting at the beginning of the array, adding and deleting one point each step until all data is exhausted. This process is depicted in Fig. 14.

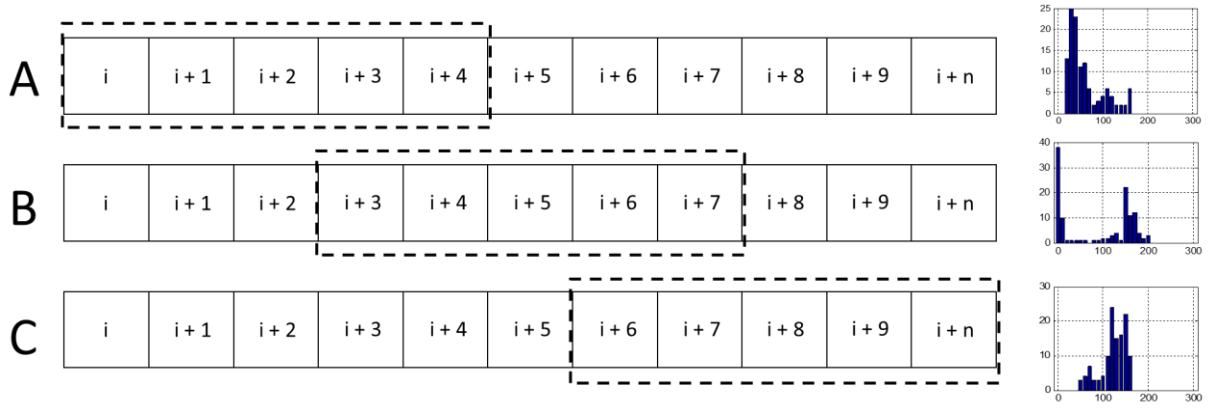


Fig. 14: Starting at the beginning of the array (step A) a histogram is created from the current data window (dotted line). The detection algorithm steps through the power data, each time adding one data point on the end and removing the first point at the beginning. This moves the current data window through the array. This continues until a bimodal histogram is detected (step B). At this point the location in time is recorded and the algorithm continues iterating until the histogram is no longer bimodal (step C). The ending value is indexed, and a ramp event is detected and saved for further analysis.

We consider a histogram bimodal when the bimodality value being evaluated is above a chosen threshold. When there are two peaks and the histogram is bimodal, the bimodality strength value will pass a certain threshold and a ramp event is recognized. The ramp event data is indexed and saved until the ranking value falls below the threshold value, and the histogram is no longer considered bimodal. Fig. 15 shows an example of normalized power observation data, the corresponding bimodality value at that point in time, and the section of data that was identified as a ramp event. Bimodality value is impartial to the direction of change in power and detects ramps both increasing and decreasing, as is apparent by ramps in both directions in Fig. 15.

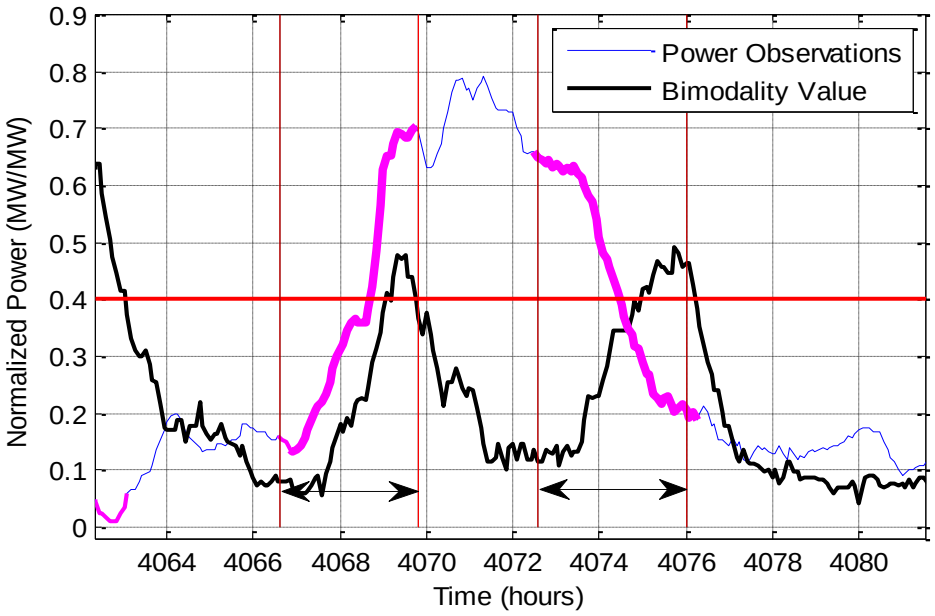


Fig. 15: Power observations with detected ramps. Horizontal red line is the ramp threshold while the bold magenta areas are sections where a ramp was detected. Vertical red lines and arrows show the beginning and end of each ramp event. The sections where a ramp was detected correspond to areas where the bimodality value is above the detection threshold. The corresponding power data for a ramp detected at time i are from i-n to i which is why the bold areas precede the bimodality value. Power observations are normalized to 1 for ease of displaying over bimodality values.

This process continues until all power observation data is exhausted and all ramp events that triggered detection are indexed and saved. This method of measuring bimodality detects ramp events while minimizing false triggers during short spikes in power, or during small non-monotonically increasing or decreasing regions as seen in the first detected ramp of Fig. 15. If the detecting algorithm is working correctly short spikes in the power observation data will not produce a peak on the histogram. Fig. 16 shows an example of a short spike in power that does not trigger ramp detection. The bimodality value increases during the rapid change in power output, but does not stay at the new level and quickly retreats.

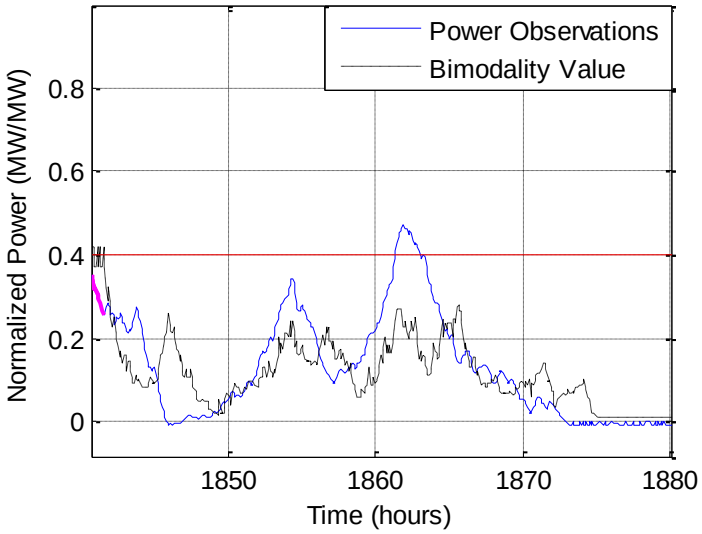


Fig. 16: Example of a peak in power not triggering a ramp event. Horizontal line is the ramp threshold. The dashed line is the bimodality value and the solid line is normalized power observation data.

2.3 Ramp Detection

All three ramp detecting methods were performed on the same power observation data using a detection window of 2.5 hours, or 30 five minute power observation data points, for a range of detection thresholds. Table XI displays the results. BC and HDT values have a range from 0 to 1, while the maximum power differences can range from 0 to 300 (MW), the rated capacity of the wind plant. So the simple ramp detection method has an output range from 0 to 1, each maximum power difference is divided by the rated capacity of the wind plant.

Table XI
Results of ramp detection methods on power observation data. HDT detected the largest number of ramps.

	Threshold								
	0.1	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**
BC	1325	2640	1037	189	34	4	0	0	0
HDT	3011	2187	1460	860	466	189	0	0	0
ΔPower	902	1128	900	615	404	257	152	86	34

The results from Table XI show the three methods yielded very different results. Many of the same ramps were detected across 2 or even all three methods, however none fully encompassed another one's results and each one detected ramps unique to that method. Fig. 17

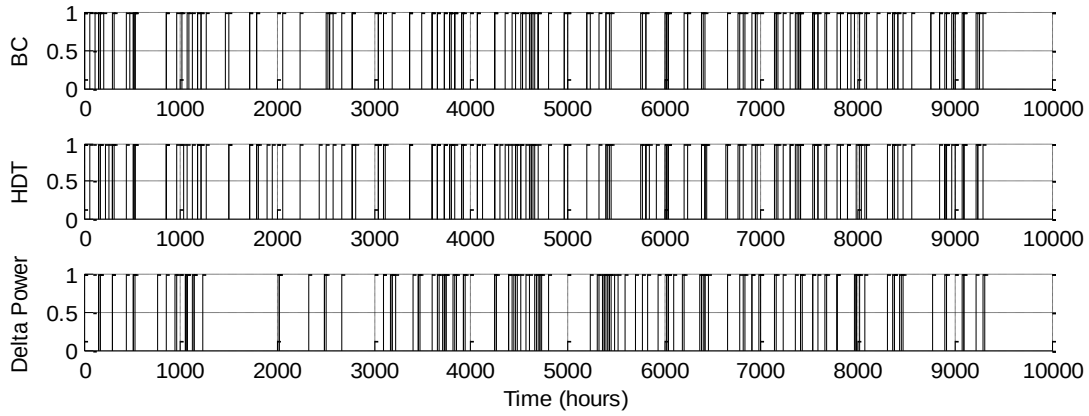


Fig. 17: Ramp Detections indexed in time for each method. Thresholds for each method where adjusted until all methods detected a similar number of ramps. BC threshold is 0.4, HDT is 0.6, and ΔPower is 0.7 which resulted in 189, 189, and 152 detections, respectively.

Because each method produced unique detects, there is no clear indication that any one method is redundant and could be discarded as unimportant. Instead a composite, weighted bimodality strength parameter was created from the results of each individual method. Comparing the number of detects shows Hartigan's dip test is the most sensitive. A previous study by [43] also recommends it as the most accurate way to measure bimodality of a distribution. For these reasons it was chosen to be the highest weighted parameter, contributing 60% to the weighted bimodality strength value. The BC was chosen to contribute 30%, and the simple detecting method of maximum change in power contributes 10% of the weighted factor. The simple method only makes up 10% because it's easily triggered by any short spike or decline in output, and is limited to detecting ramps over a certain magnitude. The BC and HDT methods are not limited to a certain change in power, and are less affected by very short duration changes in power output. Since the

range of bimodal values for each method is from 0 to 1, the weighted bimodality value, b_w, is calculated by:

$$b_w = 0.6h' + 0.3b_c{}' + 0.1\Delta P \qquad (15)$$

Where ΔP is the normalized maximum change in power of the current data window that h' and b_c' are calculated from. The new, weighted bimodality strength value also has a range from 0 to 1. The same detection analysis of bimodality values was performed on b_w as each individual detection method. Each value is compared to a threshold value and any values above the detection threshold are considered ramp events for the duration of time the values are above the threshold. Table XII shows similar results as Table XI but for the weighted bimodality value.

Table XII
Ramp detections using a weighted bimodality strength value.

	Threshold								
	0.1	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**
Weighted	2196	1967	1265	627	296	97	17	0	0

It is interesting to note that the number of detections for each method, multiplied by their respective percent contribution to b_w and then summed together, are very close to the actual number of weighted ramp detections shown in Table XII. Through visual inspection of the power observation data and detected wind ramps, a weighted threshold of 0.4 was chosen as a reasonable value to detect wind ramps while avoiding false triggers. All further analysis is performed with the 627 detected ramps found using the weighted threshold of 0.4.

3. TIMING AND MAGNITUDE ANALYSIS

Once ramps were detected using the weighted bimodality value the contribution of timing and magnitude on forecast inaccuracy could be evaluated. Power observations corresponding to each indexed ramp event were extracted and the center of the ramp was determined using a step function correlation. While not immediately needed for this analysis, the center of each ramp event was determined for reference and possible future analysis. The first and last several points of each detected ramp were averaged to find the starting and ending value for the ramp event, and a step function with these values was created. The step function was passed across the power data for the ramp, and the difference between each point of the two curves summed each time the step function was moved (i.e. an L_1 norm). The point where the sum of the differences of the curves is at a minimum corresponds to the center of the ramp event. Fig. 18 gives a visual example of how this process works.

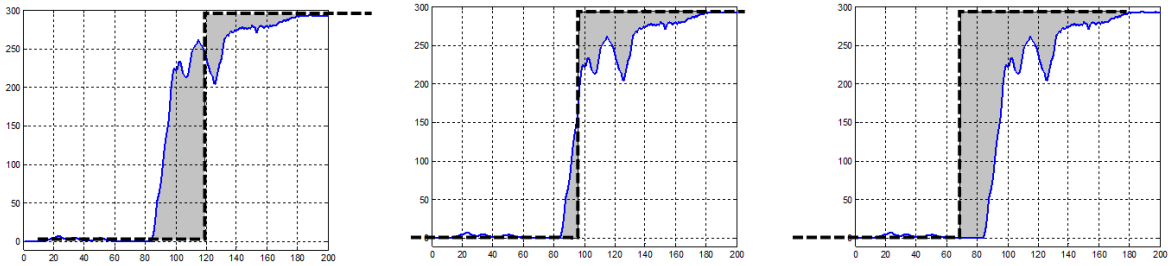


Fig. 18: Process of moving a step function over the detected ramp to determine the center. The location of the step function with the smallest amount of shaded area (center) between the two curves is the center of the ramp event.

By adding the distance from the start of the individual ramp to the indexed ramp start value, the exact location in the power observation data can be found. This technique works well as an automated method for detecting the location of the ramp that is stable over variations in both data and method, and does not require subjective input. A segment of power observation data showing two ramps and the center of the ramp events found using this method is shown in Fig. 19.

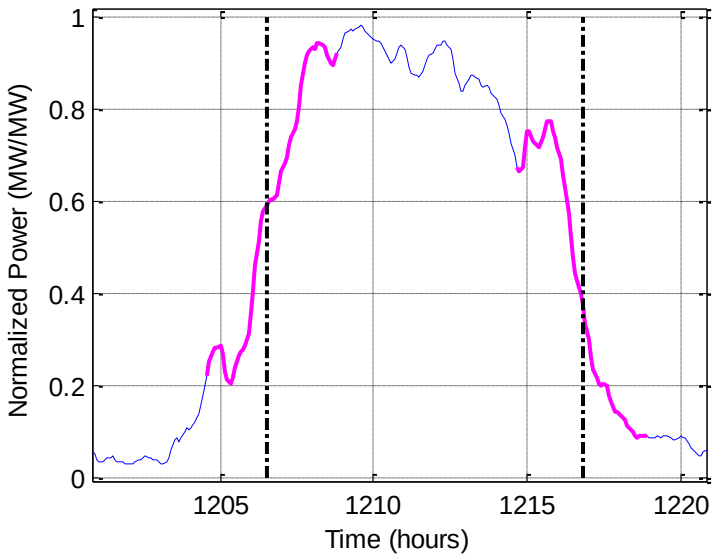


Fig. 19: Sample of normalized power observations with detected ramps and center of ramps using the described methods. Ramps are detected for both increasing and decreasing values.

3.1 Power and Energy Errors

To calculate power and energy errors, forecasts must be compared to observations. After all detected ramp events were indexed and extracted from power observations, forecasts corresponding to the same times were extracted from the forecast data. The forecasts and observations were directly compared to each other and maximum power error and total energy error for the duration of each observed ramp event was calculated. Observed energy error, e_o, is calculated by:

$$e_o = \frac{1}{5}\sum_{i=1}^{n}\left(P_{a,i} - P_{f,i}\right) \qquad (16)$$

Where n is the number of points in the ramp event, P_a is observed power, and P_f is forecasted power. The total sum is divided by 5 since each point is a 5 minute average so the output is in units of MWh. A positive power or energy error occurs when the forecasted value is less than produced,

while negative values occur when forecasted values are higher than produced. Any forecasts with missing data were excluded and that ramp event skipped.

3.2 Timing and Magnitude

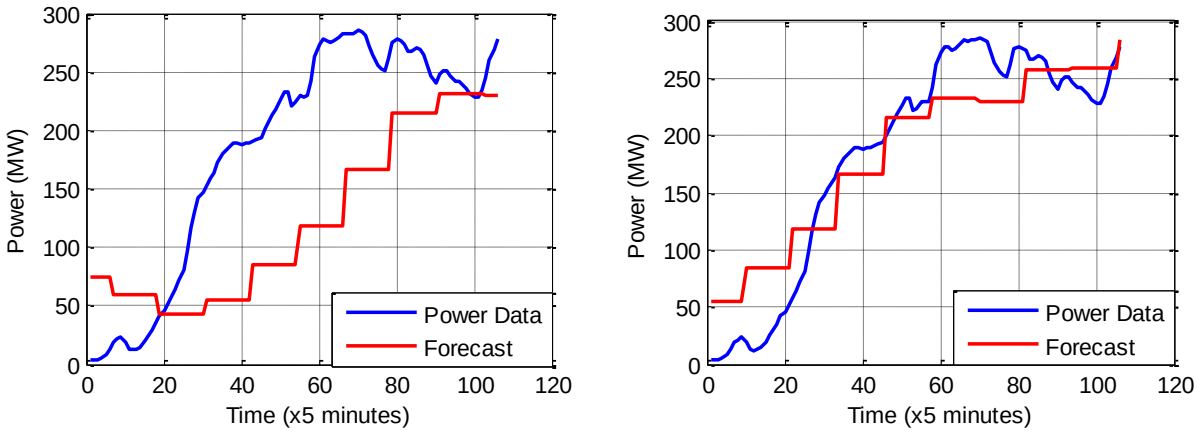
To determine the individual effects of timing and magnitude related to forecasting ramp events, each must be separated from the overall forecast errors. This is achieved by determining the optimal timing of the forecast regarding the observed ramp, then recalculating the forecast errors at the optimal timing. The difference between the recalculated errors and the original errors before the forecast was shifted in time is the contribution to the forecast error due to timing. The power and energy errors at the optimal forecast timing are the contribution to the forecasting error due to magnitude.

The optimal timing is determined by a sliding correlation similar to finding the center of each ramp event. The original observed ramp event is held stationary while the forecast data is moved across it and energy error at each step, e_i, is calculated. Forecast timing shift is evaluated to a possible ~8 hours ahead or back in time, which was determined visually by comparing forecast and observation data. Once every possible timing shift is evaluated, the point in time, i, that minimizes e is the optimal timing shift. e_i is calculated by:

$$e_i = \sum_{j=1}^{N_r} \left(P_{a,j-i} - P_{f,j-i}\right) \quad for\ i = -N_k \ldots N_k \tag{17}$$

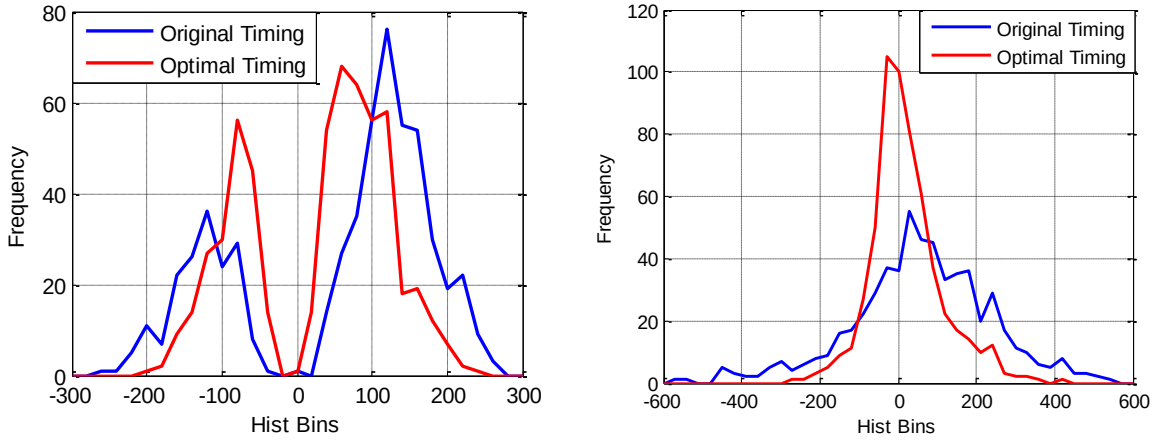
Where N_r is the number of points in the ramp event, P_a is actual power, P_f is forecasted power, and N_k is the maximum number of periods the signal is shifted. For this study, $N_k = 100$, which is ~8 hours since each point is represents 5 minutes of data. e_i is a minimum at the point of optimal timing between the forecast and observed values. Fig. 20a depicts an example of an observed

ramp event and corresponding forecast data with the original forecast timing, and Fig. 20b shows the same with the forecast shifted forward in time to the optimal time. Histograms of power and energy errors at the original and optimal timing are shown in Fig. 21a and Fig. 21b.



(a) Original Timing (b) Adjusted Timing

Fig. 20: Left: Wind ramp data with original timing forecast. Right: Optimal timing forecast. Original forecast was shifted by 33 points to the left to optimize forecast timing.



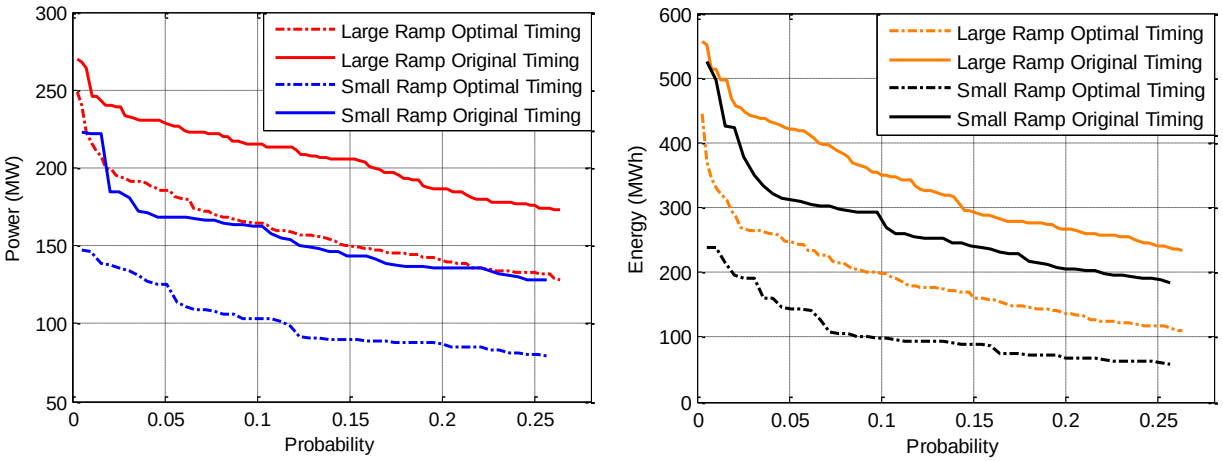
(a) Forecast power errors (b) Forecast energy errors

Fig. 21: Left: Histogram of forecast power error with original timing and optimal timing. Right: Histogram of forecast energy error with original and optimal timing (right). Data is for 627 ramps detected by using a weighted bimodality ranking value threshold of 0.4.

Wind turbines do not produce power linearly with wind speed. Instead, there are three key wind speed regimes of a turbine power curve. Regime 1 is the region above the cut-in speed when

the turbine begins to produce power, but below its rated power output. Regime 2 is the region where the turbine outputs its rated power, and regime 3 is the region above the cut-out speed which the turbine produces no power [35]. No power is produced if wind speed is less than the cut-in speed or greater than the cut-out speed. This makes the relationship between wind speed and power output simpler within certain wind speed ranges. Sustained wind speeds within regime 2 result in predictable power output for varying wind speeds, while lower wind speeds within regime 1 produce power related to the cube of the wind speed. Here, slight errors in wind speed prediction result in forecast errors while slight errors in wind speed in regime 2 do not. Due to this behavior, power and energy errors were separated into two groups whether the ramp size was greater or less than 150MW, or 50% of the capacity of the wind plant. Ramp size is defined as the maximum change in power during the wind ramp event. Out of 627 detected ramp events, 139 corresponding forecasts were missing data. After removing forecasts with missing data there was 320 large ramps and 168 small ramps to compare.

The influence of incorrect forecast timing on overall power and energy errors is the difference between errors with no timing change and errors at the optimal timing. The influence of incorrect magnitude forecasts on overall power and energy errors is the errors at the optimal forecast timing. Fig. 22a shows a section of an absolute value probability distribution of forecast power errors with the largest 0.25% of errors removed, and Fig. 22b shows the same but for energy. Removing the largest 0.25% of errors removes the influence of very low probability outliers in the data that may skew the results. Note that larger magnitude events correspond to lower probabilities, while smaller events correspond to higher probabilities of occurrence. Table XIII summarizes the results of Fig. 22.



(a) Power Errors (b) Energy Errors

Fig. 22: Comparison of power errors (left) and energy errors (right) for optimal and original timing. Left: Power errors for large and small ramp events before and after timing change. Right: Energy errors for large and small ramp events before and after timing change.

Table XIII
Qualitative summary of the results from the figures. Interestingly, for low probability events (large magnitude errors), behavior for small ramps (<50% of plant size) is opposite of that for large ramps (>50% of plant size).

	Probability of Occurrence	Small ramps (< 50% of plant capacity)	Large ramps (≥ 50% of plant capacity)
Power Errors	<0.025	Timing error increases with increase in error size	Timing error decreases with increase in error size
	>0.025	Timing error approximately independent of error size	
Energy Errors	<0.05	Timing error increases with increase in error size	Timing error decreases with increase in error size
	>0.05	Timing error approximately independent of error size	

Fig. 23 below compares the timing and magnitude influences on forecast errors by looking at the fraction of the energy or power errors attributable to timing. Depicted on the vertical axis is the ratio of timing error to magnitude error. This ratio is the ratio of each solid line of Fig. 22a and Fig. 22b to its dashed line counterpart. The horizontal axis is the same section of the absolute value probability distribution function from 0.25% to 26%.

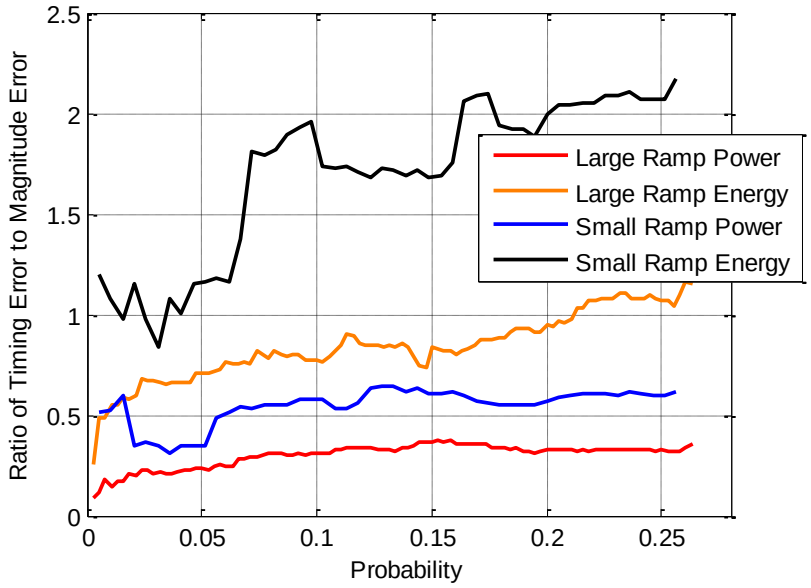


Fig. 23: A lower timing to magnitude ratio means of the overall power or energy error, a larger portion was due to magnitude.

Although the left side of Fig. 23 is less likely to occur, it consists of larger magnitude power and energy errors. The higher probability region to the right is more likely to occur but consists of lower magnitude errors. Fig. 22a shows there isn't a large difference in power errors between low and high probability regions, while Fig. 22b shows there is more variation in energy errors.

As probability increases, forecast timing has a larger role in the overall error, more in small ramp events than large ramp events. The power observation data used for this study had definite daily oscillations of wind increasing and decreasing. Most large ramp events are likely these daily fluctuations, where small wind fluctuations throughout the day likely lead to small ramp events. Explanations for this behavior could be smaller wind ramps have less time between them, making exact timing more difficult to predict. Conversely, large ramp events with longer time intervals are easy to detect and predict the timing of with atmospheric data, but predicting terminal ramp speed may be more difficult. In other words, it is easier to forecast a large wind front approaching, which increases power output from near zero to 80, 90, or 100% operating capacity approaching, but difficult to exactly predict the final wind speed, and thus power output.

4. CONCLUSIONS

Ramp events were successfully detected in power observation data using a bimodality strength value, which was calculated from a histogram of power observation data. For each observed and detected ramp event, the largest difference in power and total energy between forecasts and observations was calculated. Next, for each detected ramp event the best case timing was determined. Power and energy were recalculated after adjusting the forecast to the optimal timing. Using these results the influences of timing and magnitude were separately estimated in the power and energy errors and categorized by large and small ramp events (whether the ramp size was larger or smaller than 50% of the capacity of the plant).

Results indicate that efforts to improve wind power forecasts should concentrate more on better estimation of power output magnitude than on improving the timing of ramp detections for large wind ramp events. Timing errors tend to contribute more to forecast errors in small ramp events. However small ramp events shouldn't be discarded as unimportant because poor forecasting of small events can cause energy and power errors similar or larger in size to accurate forecasting of large ramp events

Further research in this area could look at other methods to categorize ramp events, looking for specific trends such as time of day or year that large and small ramps occur. This study defined ramp size as the largest difference in magnitude of power data for the duration of the ramp event. Other methods may choose to use the terminal power level, or rate of change. Further economic analysis of the tradeoff between low likelihood, large power and energy errors and high likelihood, small power and energy errors could be beneficial to grid operators.

OVERALL CONCLUSIONS

Since wind generation is stochastic in nature, there will always be forecasting errors and a need to smooth grid power, especially as wind power increases in prominence worldwide in the future. Likewise, improving forecasting models is a necessity. This study has shown improved forecasts greatly improves contract firming requirements. It is widely accepted that improving forecasts is the best solution for high wind power penetration [9].

Part I showed forecasting errors for short time scales can be represented by a Cauchy distribution. It determined wind plant size and number of turbines are not always good indicators of contract firming requirements. Instead, forecast lead time has the biggest significance on forecast accuracy. It developed a normalized curve to relate forecast accuracy to energy firming requirements.

Part II created a statistical analysis method to estimate forecast errors and firming requirements using a wind distribution and forecasting error distribution. The estimation method was compared to results using time series power observation and forecast data and shown to be accurate. Forecast errors between actual forecast and power observations were then compared for several forecast lead times. These results confirm persistence forecasting is effective for lead times less than 1 hour but longer times require the use of more advanced forecasting models.

Part III looked closer at wind forecasting methodologies, specifically targeting wind ramp events. Wind ramps cause significant forecast errors and more accurately predicting them greatly improves forecast accuracy and reduces firming requirements. Part III develops an advanced ramp detection method using a novel combination of bimodality detection methods. It also uses a unique method of finding the center of a ramp event. It compares detected ramps to forecasts to calculate

forecast errors for the ramp events, and removes the individual contributions of both timing and magnitude errors on overall forecast errors. Data is categorized according to ramp size. These results suggest magnitude errors significantly contribute to forecasting errors in large wind ramp events, and timing errors contribute more in small ramp events.

Several research focuses have been identified as good future work in these areas. Part II compares actual forecast errors to persistence at a 1 hour forecast lead time. To more accurately compare to its earlier results 10 minute forecast data should be obtained or each of the three methods performed with 1 hour wind power observation data instead of 10 minute. Part III could benefit from future research looking at other ways to categorize ramp events such as by time of day or year. An economic analysis of the benefit of reducing low frequency large magnitude ramp forecast errors or high frequency small magnitude errors would offer guidance to where research should focus.

REFERENCES

[1] DeLuchi M. A., 1991, "Emissions of Greenhouse Gases from the Use of Transportation Fuels and Electricity. Volume 1: Main Text," National Technical Information Service, Alexandria, VA.

[2] Lashof D., Ahuja D., 1990, "Relative contributions of greenhouse gas emissions to global warming," Nature, **344**, pp. 529-531.

[3] Makower J., Pernick R., Wilder C., 2009, "Clean Energy Trends 2009," Clean Edge Inc., San Francisco, CA.

[4] Global Wind Energy Council, 2012, "Global Wind Energy Outlook Report, 2012," GWEC, Brussels, Belgium.

[5] Tuohy A., O'Malley M., 2009, "Impact of pumped storage on power systems with increasing wind penetration," *2009 IEEE Power and Energy Society General Meeting*, IEEE, Calgary, AB.

[6] Wiser R., Bollinger M., Barbose G., Belyeu K., Hand M., Heimiller D., 2008, "Annual Report on U.S. Wind Power Installation, Cost, and Performance Trends," US Department of Energy, Lawrence Berkeley National Laboratory.

[7] Tuohy A., Meiborn P., Denny E., O'Malley M., 2009, "Unit Commitment for Systems with Significant Wind Penetration," IEEE Transactions on Power Systems, **24(2)**, pp. 592-601.

[8] Barnhart C., Dale M., Brandt A., Benson S., 2013, "The Energetic Implications of Curtailing *Versus* Storing Solar- and Wind-Generated Electricity," Energy and Environmental Science, **6(10)**, pp. 2804-2810.

[9] Ela E., Kemper J., 2009, "Wind Plant Ramping Behavior," NREL/TP-550-46938, National Renewable Energy Laboratory (NREL), Golden, CO.

[10] Wu Y.K., Hong J.S., 2007, "A literature review of wind forecasting technology in the world," *Power Tech*, IEEE.

[11] Edward L., 1963, "Deterministic Nonperiodic Flow," Journal of the Atmospheric Science, **20(2)**, pp. 130-141.

[12] Wyngaard J., 2010, *Turbulence in the Atmosphere*, Cambridge University Press, Cambridge, UK.

[13] Haupt S., Mahoney W., Parks K., Troccoli A., Dubus L., 2014, “Wind Power Forecasting,” *Weather Matters for Energy*, Springer, pp. 295-318, Chap. 14.

[14] Mahoney W.P., Parks K., Wiener G., Liu Y., Myers W.L., Sun J., Monache L.D., Hopson T., Johnson D., Haupt S.E., 2012, "A Wind Power Forecasting System to Optimize Grid Integration," IEEE Transactions on Sustainable Energy, **3(4)**, pp. 670-682.

[15] Kamath C., 2010, “Understanding wind ramp events through analysis of historical data,” *2010 IEEE PES Transmission and Distribution Conference and Exposition*, New Orleans, LA.

[16] European Climate Foundation, 2010, “Roadmap 2050: a practical guide to a prosperous, low-carbon Europe,” European Climate Foundation, Europe.

[17] Steinke F., Wolfrum P., Hoffmann C., 2013, “Grid vs. storage in a 100% renewable Europe,” Renewable Energy, **50**, pp. 826-832.

[18] Delmas M., Montes-Sancho M., 2011, “U.S. state policies for renewable energy: Context and effectiveness,” Energy Policy, **39(5)**, pp. 2273-2288.

[19] GE Energy, 2010, “Western Wind and Solar Integration Study: executive summary,” NREL/SR-550-47781, National Renewable Energy Laboratory (NREL), Golden, CO.

[20] Barton J.P., Infield D.G., 2004, “Energy Storage and Its Use with Intermittent Renewable Energy,” IEEE Transactions on Power Systems, **19(2)**, pp. 441-448.

[21] Iler A., 2012, “Midwest Independent Transmission System Operator, Inc. Order No. 745 Compliance Filing; Docket No. ER11-4337-002,” MISO.

[22] Giebel G., Brownsword R., Kariniotakis G., 2003, "The State-Of-The-Art in Short-Term Prediction of Wind Power A Literature Overview," Risø National Laboratory, Roskilde, Denmark.

[23] Nielsen T. S., Joensen A., Madsen H., Landberg L., Giebel G., 1998, “A New Reference for Wind Power Forecasting,” Wind Energy, **1(1)**, pp. 29-34.

[24] Hodge B.M., Milligan M., 2011, “Wind Power Forecasting Error Distributions over Multiple Timescales,” *2011 IEEE Power and Energy Society General Meeting*, IEEE, San Diego, CA.

[25] Pinson P., Kariniotakis G. N., 2003, “Wind Power Forecasting using Fuzzy Neural Networks Enhanced with On-line Prediction Risk Assesment,” *2003 IEEE Bologna Power Tech Conference Proceedings*, IEEE, Bologna, Italy, **2**.

[26] Doherty R., O'Malley M., 2005, "A New Approach to Quantify Reserve Demand in Systems With Significant Installed Wind Capacity," IEEE Transactions on Power Systems, **20(2)**, pp. 587-595.

[27] Mills A., Wiser R., 2012, "Changes in the Economic Value of Variable Generation at High Penetration Levels: A Pilot Case Study of California," Berkeley National Lab.

[28] Kirschen D., Strbac G., 2004, *Fundamentals of Power System Economics*, John Wiley and Sons.

[29] O'Connell M., Zimmerle D., 2013, "Supplemental Energy Needed for Wind Integration," *2013 IEEE Power and Energy Society General Meeting*, IEEE, Vancouver, BC.

[30] Tuohy A., O'Malley M., 2011, "Pumped storage in systems with very high wind penetrations," Energy Policy, **39(4)**, pp. 1965-1974.

[31] Hu P., Karki R., Billinton R., 2009, "Reliability Evaluation of Generating Systems Containing Wind Power and Energy Storage," IET Generation, Transmission and Distribution, **3(8)**, pp. 783-791.

[32] Giorsetto P., Utsurogi K., 1983, "Development of a New Procedure for Reliability Modeling of Wind Turbine Generators," IEEE Transactions on Power Apparatus and Systems, IEEE, **PAS-102(1)**, pp. 134-143.

[33] Milligan M.R., Artig R., 1999, "Choosing Wind Power Plant Locations and Sizes Based on Electric Reliability Measures Using Multiple-Year Wind Speed Measurements," *U.S. Association for Energy Economics Annual Conference*, Orlando, FL.

[34] Dhanju A., Whitaker P., Kempton W., 2007, "Assessing offshore wind resources: An accessible methodology," Renewable Energy, **33(1)**, pp. 55-64.

[35] Wan Y.H., Ela E., Orwig K., 2010, "Development of an Equivalent Wind Plant Power-Curve," NREL/CP-550-48146, *WindPower 2010*, Dallas, TX.

[36] Bathurst G., Weatherill J., Strbac G., 2002, "Trading Wind Generation in Short Term Energy Markets," IEEE Transactions on Power Systems, **17(3)**, pp. 782-789.

[37] Staid A., Guikema S., 2013, "Statistical analysis on installed wind capacity in the United States," Energy Policy, **60**, pp. 378-385.

[38] Ela E., Milligan M., Kirby B., 2011, "Operating Reserves and Variable Generation," NREL/TP-5500-51978, National Renewable Energy Laboratory (NREL), Golden, CO.

[39] SAS Institute, 1990, "SAS-STAT User's Guide, Version 6, 4th Edition," SAS Institute.

[40] Hartigan J. A., Hartigan P. M., 1985, "The Dip Test of Unimodality," The Annals of Statistics, **13(1)**, pp. 70-84.

[41] Akaike H., 1974, "A new look at the statistical model identification," IEEE Transactions on Automatic Control, **19(6)**, pp. 716-723.

[42] Pfister R., Schwarz K., Janczyk M., Dale R., Freeman J., 2013, "Good things peak in pairs: a note on the bimodality coefficient," Frontiers in Psychology, **4**.

[43] Freeman J., Dale R., 2013, "Assessing bimodality to detect the presence of a dual cognitive process," Behavior Research Methods, **45(1)**, pp. 83-97.

[44] Knapp T., 2007, "Bimodality Revisited," Journal of Modern Applied Statistical Methods, **6(1)**.

[45] Mechler, "Hartigan Dip Test MATLAB files", http://www.nicprice.net/diptest/