Wind Energy Forecasting: The Economic Benefits of Accuracy

John Zack, Bruce Bailey and Michael Brower
AWS Truewind, LLC
255 Fuller Road, Suite 274
Albany, New York USA 12203
jzack@awstruewind.com
+1-518-283-5169 voice
+1-518-283-5357 fax

Abstract

Despite the many desirable attributes of wind energy, the fact that wind is an intermittent resource has been a source of concern for utility system operators and managers. State-of-the-art wind power production forecast systems have demonstrated that they can significantly enhance the value of wind generation by increasing system reliability and reducing operating costs. The best forecasts are made with a sophisticated multi-component forecast system such as AWS Truewind's eWind system, which is composed of both physics-based and statistical models that take advantage of a wide range of local and regional meteorological and power production data. Such systems typically yield a mean absolute error for hourly power production forecasts of about 2% to 5% of installed capacity for a one-hour-ahead forecast, about 9% to 14% for a 4-hour ahead prediction, and about 13% to 19% for a day-ahead forecast. A study by GE Energy estimated that this level of forecast performance, if applied to 3,300 MW of wind generation in New York State, would reduce utility system operating costs by about $125 million per year compared to no forecasts. The study also indicated that the current level of forecast performance would attain 80% of the cost reduction that could be achieved with a perfect wind forecast.

1. Introduction

The utility-scale generation of electricity from the wind has a number of desirable attributes – such as no air pollution and low operating costs. Despite these attributes, operators of electric power systems remain concerned that wind, as an intermittent resource, can harm system reliability and raise operating costs. There is no doubt that the addition of large amounts of wind generation can impose a burden on system operators, who must dynamically schedule other types of generation to respond to changes in wind generation. In response, wind plants are often required by grid rules, contractual obligations, or market incentives to accurately forecast their output hours or days ahead.

These requirements motivate the development and deployment of wind forecasting systems capable of producing accurate forecasts of the wind power production in time steps ranging from minutes to days and for forecast time horizons extending from a few minutes to many days ahead. These forecasts enable system operators to serve system loads while minimizing both the probability of failure and the costs of providing electricity. Recent studies have indicated that accurate wind energy forecasts bring significant cost savings. This paper provides a summary of the state-of-the-art approach to wind power forecasting, an overview of the performance typically achieved, and an estimate of the value of these forecasts.

2. Forecasting Methods

State-of-the-art forecast systems typically have three components: physics-based models, statistical models and plant output models. These components operate upon a wide range of data types to produce wind power production forecasts. The state-of-the-art in wind power production forecasting systems is exemplified by AWS Truewind's eWind system. A schematic chart of the system components and the types of data typically used by the system is presented in Figure 1.

2.1. Physics-based Models

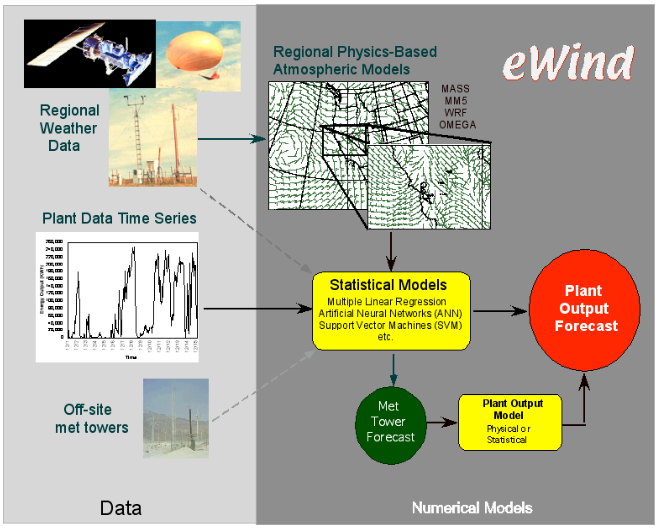

Figure 1. A schematic depiction of the data and models used by AWS Truewind's eWind forecast system to produce state-of-the-art wind power production forecasts.

The physics-based atmospheric models are essentially computational fluid dynamic (CFD) models that have been configured to efficiently and accurately simulate atmospheric processes. These models are formulated from the fundamental principles of physics (i.e. conservation of mass, momentum and energy and the equation of state for the constituents of air), which yield a set of differential equations that are typically solved on a three-dimensional grid. The size of the grid elements and the extent of the computational domain in these models determine the scales of atmospheric processes that can be simulated by a specific configuration of a model. A user of the eWind system has the option to use one or more of several models and to configure them for a particular application. The primary physics-based model in the eWind system is the Mesoscale Atmospheric Simulation System (MASS) model [1]. Other physics-based models available in the system are the Weather Research and Forecasting (WRF) model [2], Mesoscale Model 5 (MM5) and the OMEGA model [3]. The availability of several different models in the eWind system facilities the generation of forecast ensembles as discussed in section 2.4.

2.2. Statistical Models

The statistical atmospheric models are basically empirical quantitative relationships between atmospheric variables or between atmospheric variables and non-atmospheric variables. It is important to differentiate between the algorithms used to derive a set of empirical relationships and the relationships themselves. The relationships represent the model of a particular process or phenomena. The algorithms represent a way to quantify those relationships. The eWind system employs one or more of several different statistical algorithms to develop the statistical models for a particular application. The algorithms include techniques such as Screening Multiple Linear Regression (SMLR), Artificial Neural Networks (ANN), and Support Vector Regression (SVR) as well as other methods such as fuzzy logic clustering that can be employed to pre-condition training samples to enable the training methods to find stronger empirical relationships. These statistical tools can be used to develop relationships between the output of the physics-based models and specific parameters at the wind plant such as the wind speed and direction or the plant power production. This type of application is often referred to as Model Output Statistics (MOS). The same statistical tools can also be used to develop relationships between sets of measured parameters. The autoregressive prediction of a plant's power production from a time series of that production is an example of such an application. The eWind system has a high degree of flexibility in how the statistical tools can be used to generate wind power production predictions.

2.3. Plant Output Models

The third type of model employed within the eWind prediction system is a plant output model. The plant output model is a relationship between the meteorological parameters at or in the vicinity of a wind plant and the concurrent power production. The relevant meteorological parameters are typically the wind speed and direction at turbine hub height and the air density. In general, plant output models can be physics-based or statistical models. The eWind system employs only statistical plant output models. This is because physics-based plant output models require much information about the details of a plant's layout, turbines and the properties (roughness height, etc.) of the earth's surface in the vicinity of the wind plant. In the eWind system it is possible to avoid the use of an explicit plant output model and generate a power production prediction directly from the output of physics-based atmospheric prediction models or meteorological measurements from on-site, off-site or remote sensing systems. In that case the plant output model is implicit in the statistical relationships between the predictors and the resulting power output. In general, better results have been achieved by using an explicit plant output model in most applications.

2.4. Forecast Ensembles

The use of an ensemble of forecasts has been demonstrated to yield higher quality forecasts and forecast uncertainty estimates in a number of meteorological forecast applications [4]. The basic concept is that a set of forecasts is generated by perturbing the input data and the model configuration parameters within their respective ranges of uncertainty, producing a new forecast with the perturbed input data or model parameters. In theory, this provides a set of forecasts that bracket the ultimate realized value of the predicted variables. A composite (e.g. mean, median, etc.) of the set of forecasts typically provides an explicit prediction than any individual forecast and the dispersion (i.e. the spread) of the ensemble provides information about the forecast uncertainty. Since there is an enormous number of input data variables and model parameters, it is not practical to generate forecasts with all of the possible perturbations. Thus, in practice, one must select a subset of input data or model parameters to perturb to generate a forecast ensemble. The objective is to select the input data or model parameters that are responsible for most of the uncertainty in the forecast system. This can be quite difficult since the data or parameters responsible for the uncertainty typically will vary from one forecast cycle to another due to differences in weather regimes and other factors. The eWind system is capable of generating an ensemble of forecasts by using different sets of input data or physics-based or statistical model configurations. The resulting set of forecasts can be combined into a ensemble composite forecast by training a statistical model to effectively weight the different forecasts according to their performance characteristics in the training sample or by combining them in a user-specified manner (e.g. averaging all the members of the ensemble). Measures of the ensemble spread (range, standard deviation, etc.) are then used to estimate the forecast uncertainty.

2.5. Forecast System Operations

The relative importance of the various inputs and models depends upon the look-ahead period of the forecast as well as other factors such as the characteristic weather regimes, surface properties in the vicinity of the wind farm and the amount and type of available data from the plant and other sources. The skill of short-term forecasts is typically more dependent upon the time series data from the wind plant as well as recent data from nearby off-site locations or nearby remote sensing systems (such as Doppler Radars or wind profilers) and the performance of the statistical models. However, even 1 to 2 hour ahead forecasts can benefit from the intelligent use of output data from a customized high-resolution physics-based model. The performance of day-ahead forecasts does not have much dependence on the current data from the wind plant or nearby locations. These forecasts are based predominantly on the output from the physics-based atmospheric models that has been adjusted by a MOS procedure to remove systematic errors that are common in the output of physics-based models. Although current data from the wind plant is not crucial to day-ahead forecast performance, historical meteorological and plant production data is crucial to the successful utilization of the MOS procedure and the construction of a high quality statistical plant output models.

3. Forecast Performance

An assessment of the quality of a particular forecasting application depends upon the metric that is employed and the particular parameter that is forecasted (e.g. 5-min production, hourly production, daily production, etc.). There are a number of other factors such as the characteristics of the wind plant and the associated wind resource, the quality of data from the plant and surrounding areas, and types and scales of phenomena responsible for the variability in the wind. Therefore, it is difficult to compare forecast performance among wind farms or even among several months of the year for the same wind plant. Therefore, any discussion of forecast performance must be phrased in terms of relatively broad ranges unless one is referring to a specific plant for a particular period of time.

3.1. Single Wind Farm

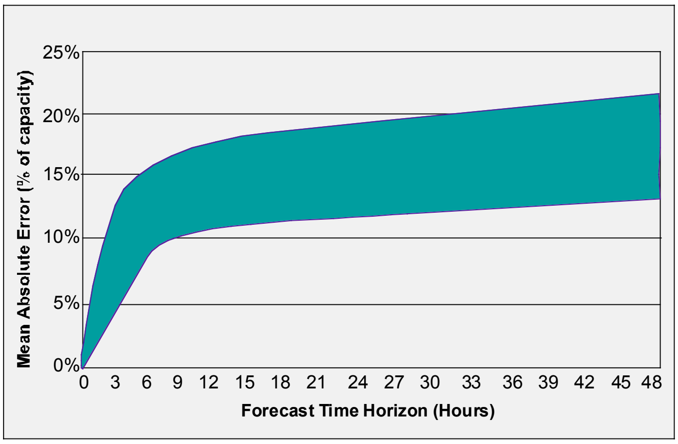

Figure 2. The approximate range of mean absolute error (MAE) as a percentage of a plant's installed capacity for forecasts of the hourly power production for an individual wind plant by forecast time horizons for eWind's next-hour (0-8 hours) and next day (9-48 hours) forecast services.

Results from many applications of the eWind system in a wide range of meteorological regimes indicates that forecasts of the hourly wind power production of an individual wind plant typically are in the range shown in Figure 2 during the first 48 hours of the forecast period. Day-ahead forecasts generally have a mean absolute error between 13% and 19% of the installed capacity of the plant. This range is consistent with the day-ahead forecasting results obtained in a recent wind power production forecasting evaluation project [5] sponsored by the California Energy Commission, which compared the performance of the eWind system and the Prediktor system developed by Risoe National Laboratory of Denmark for two wind plnats in California over a one year period. The errors typically grow at a rate of about 1.0% to 1.5% of installed capacity per day through the first 5 days of the forecast period. After 5 days, hourly predictions of wind power production have very little skill over a climatological forecast (long-term average production by hour of the day) and the errors cease to grow with the time and the forecast quality is the same as that of a climatological forecast. However, the performance of forecasts of the daily power production is often better than climatology significantly beyond 5 days. The variation in the performance as a function of the forecast time horizon is much greater for hours-ahead type forecasts. A benchmark for state-of-the-art (eWind) performance for 1-hour-ahead forecasts of the hourly power production (of an individual plant) is an MAE of 2% to 5% of installed capacity. The error growth is relatively rapid for very-short term forecasts and the MAE for state-of-the-art 4-hour-ahead forecasts of the hourly power production is typically in the range of 9% to 14% of installed capacity.

3.2. Regional or System-Wide Forecast Performance

The performance of forecasts for a collection of wind plants will typically be much better than that for individual plants. This is because the errors in the predictions for individual plants tend to offset each other. The mean absolute error for forecasts of the aggregated power production for a collection of wind plants is sensitive to the number of wind plants and the inter-plant error correlations (related to the geographic dispersion of the plants). An example of the improvement in forecast performance associated with the dispersive growth of wind power production on a power system was compiled in a study for the New York State Independent System Operator (NYISO). In this study, a total of 3300 MW of wind generation capacity in areas likely to see wind plant development was added to the system over a 10-year period. The wind conditions for the year 2002 were used to estimate day-ahead forecast errors for each of the projected wind plants as well as day-ahead forecast errors for aggregated power production of all of the wind farms. The results of these calculations are shown in Figure 3. The table on the left side of the figure lists the number of wind farms and the associated installed capacity that was assumed to be in operation during each year. The chart on the right depicts the year-by-year MAE.

The average MAE for the wind farms is about 14.5% at the start of the period and rises slightly to just over 15% during the middle of the period and then slightly decreases towards the end of the 10-yr period. These fluctuations are related to the variations in wind conditions among the locations of the plants. In the first year of the period, the MAE of the aggregated forecast is slightly less than 12.5% or about 2% of installed capacity below the average MAE for the 3 individual plants. The improvement of 2% represents the impact of offsetting random errors among the three plants. In the following year (2005) the MAE of the aggregate forecasts sharply rises to just over 14%, which is only about 0.5% below the average MAE of the four operating wind plants. Thus, the impact of the aggregation effect has substantially decreased from 2004 to 2005. This decrease is related to the fact that one wind plant with a size of 300 MW was added to the system from 2004 to 2005. This significantly reduced the geographic diversity of the system since the aggregate error was dominated by the errors at the 300 MW wind plant's location (vs. 48 MW elsewhere). The MAE of the aggregate forecasts drops back to near 12.5% in 2007 as 3 more wind plants in different locations in the western and central part of the state with a capacity of another 200 MW are added to the system. After 2007, the MAE of the aggregate forecasts gradually decreases to about 11% (about 4% below the average MAE) as more wind plants are added in the western and central part of the state. The rate of MAE reduction decreases in the 2007-2012 period as more plants are added in the same geographical region, which results in a fairly high correlation of hourly errors between the new plants and the existing plants. A substantial decrease in the aggregated MAE to near 10% occurs between 2012 and 2013 as 600 MW are added in an anticipated offshore project in the extreme southeastern part of the state. The forecast errors for this site have a low degree of correlation with the majority of the plants in the western and central part of the state and hence the aggregated MAE is significantly reduced. At this point the magnitude of the "aggregation effect" is about 5%.

Year	Plants	Capacity (MW)
2004	3	48
2005	4	348
2006	7	549
2007	12	960
2008	16	1,396
2009	20	1,750
2010	24	2,029
2011	28	2,264
2012	32	2,577
2013	36	3,348

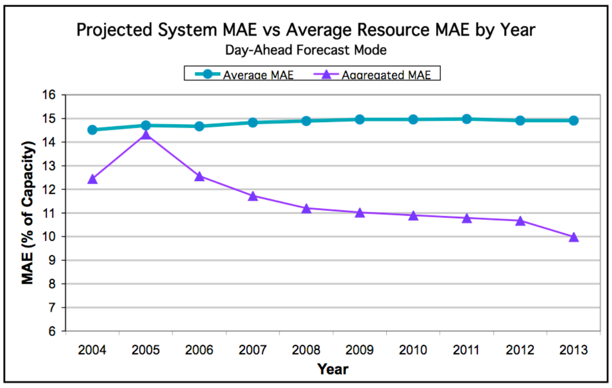

Figure 3. A comparison of the annual average wind power production forecast mean absolute error (MAE) for individual wind farms and the annual MAE of forecast for the aggregated production from all wind farms for the specific scenario of projected growth in the installed capacity wind power production in New York State for the period 2004 to 2013 that is shown in the table on the left.

4. Economic Benefits

The economic benefits of wind forecasting were assessed in a recent study [6] done by GE Energy for the New York State Energy Research and Development Authority (NYSERDA) and the NYISO. The study was a comprehensive evaluation of the effects of integrating wind power on transmission system planning, reliability, and operation. A portion of this study examined the overall economic impact of integrating a hypothetical 3,300 MW of wind power production capacity into the New York State system. The 3,300 MW of projected wind generation represents approximately 10% of the system's current peak load and almost two orders of magnitude more wind power capacity than the 48 MW in operation at the time the report was prepared.

The GE Energy study calculated energy prices based on a model of functioning day-ahead and hour-ahead commercial wholesale markets. This is different from most previous analyses of the economic impact of wind power production, which have considered operating costs only. In the GE Energy study it was assumed that the wind power produces are "price-takers". This means that the supply-demand balance in the market determines the price and that wind generators are paid the market price.

The overall conclusion of the study was that the New York State power system can reliably accommodate at least a 10% penetration of wind relative to the system's peak load with just minor adjustments to the planning, operation and reliability practices. In addition, this study estimated that the annual variable operating costs (fuel, plant start-up costs, etc.) of the system would be reduced by $350 Million with the addition of wind. It was further estimated that $125 Million, or 36%, of the cost reduction is associated with state-of-the-art wind power forecasting. This is about 80% of the estimated cost reduction that could be achieved with a perfect wind power production forecast.

5. Summary and Conclusions

State-of-the-art wind power production forecast systems have demonstrated that they can significantly enhance the value of wind generation by increasing system reliability and reducing system operating costs. AWS Truewind's eWind system exemplifies such systems. The Mean Absolute Error (MAE) of hourly power production forecasts for an individual wind plant grows from about 2% to 5% of installed capacity for a one-hour ahead forecast to about 9% to 14% of installed capacity for a 4-hour ahead forecast. The typical MAE of forecasts of the hourly power production for the next calendar day is 13% to 19% of installed capacity. After the next calendar day the MAE typically grows at a rate of 1% to 1.5% of installed capacity per day. The forecast performance for the power production from aggregations of wind plants is typically much better because uncorrelated errors from the individual plants tend to offset each other. A study by GE Energy has demonstrated that 3,300 MW of wind capacity in New York State would reduce total system operating costs by about $350 Million, of which $125 Million would be due to wind power forecasting. This indicates that state-of-the-art forecasts have an immense value for utility systems compared to their cost.

References

[1] Kaplan, M. L., J. W. Zack, V. C. Wong and J. J. Tuccillo, 1982: Initial results from a mesoscale atmospheric simulation system and comparisons with an AVE-SESAME I data set., *Mon. Wea. Rev.*, **110**, 1564-1590.

[2] Michalakes, J., J. Dudhia, D. Gill, T. Henderson, J. Klemp, W. Skamarock, and W. Wang, 2004: "The Weather Research and Forecast Model: Software Architecture and Performance,"to appear in proceedings of the 11th ECMWF Workshop on the Use of High Performance Computing In Meteorology, 25-29 October 2004, Reading U.K. Ed. George Mozdzynski.

[3] Bacon, D. P., N. N. Ahmad, Z. Boybeyi, T. J. Dunn, M. S. Hall, P. C. S. Lee, R. A. Sarma, M. D. Turner, K. T. Waight, S. H. Young, and J. W. Zack, 2000: A Dynamically Adapting Weather and Dispersion Model: The Operational Multiscale Environment model with Grid Adaptivity (OMEGA). *Mon. Wea. Rev.*, 128, 2044-2076.

[4] Sivillo, J.K, J.E. Ahlquist, and Z. Toth, 1997: An ensemble forecasting primer. *Wea. Forecasting*, **12**, 809-818.

[5] California Wind Energy Forecasting System Development and Testing Phase 2: 12-Month Testing, EPRI, Palo Alto, CA, and the California Energy Commission, Sacramento, CA: 2003, EPRI Report # 1007339. Available at: http://www.epri.com/OrderableitemDesc.asp?product_id=1007339.

[6] "The Effects of Integrating Wind Power on Transmission System Planning, Reliability, and Operations:, report prepared for New York State Energy Research and Development Agency, March 4, 2005. Available at: http://www.nyserda.org/publications/wind_integration_report.pdf.