Draft Report

Overview of Wind Energy Generation Forecasting

Submitted To:

New York State Energy Research and Development Authority

and the

New York State Independent System Operator

Prepared By:

TrueWind Solutions, LLC
Principal Author: Dr. John Zack

AWS Scientific, Inc.
255 Fuller Road, Suite 274
Albany, NY 12203-3656

December 17, 2003

Overview of Wind Energy Generation Forecasting

1. Need for Forecasts

The intermittence of wind energy presents a special challenge for utility system operations. While conventional power plants produce a near constant output – barring rare emergency outages – the output of a wind plant fluctuates. In some parts of the US, such fluctuations can amount to several hundred megawatts in a matter of an hour or two. To the extent the fluctuations cannot be predicted, they create costs for the electricity system and consumers as well as potential risks to the reliability of electricity supply.

One of the principal mechanisms a grid operator, such as an Independent System Operator (ISO), uses to limit unexpected changes in plant output is to charge suppliers a fee for "uninstructed deviations" between their forward schedules – i.e., their predicted output – and actual generation. This policy encourages suppliers to maintain a high level of reliability while also compensating the system for the costs of having either excess or insufficient generation. Typically, the fee is assessed on the deviation in each hour based on the market-clearing price of the Real Time Market. However, considering the volatility of wind plant output, some grid operators recognize that wind energy suppliers could be severely penalized if required to pay for deviations on an hourly basis, and that this would undercut the ability of a state or region to meet its future electricity needs in an environmentally sound manner.

The performance requirements for a forecasting service are dictated by the needs of both the grid operator and the wind generators. From the perspective of wind generators, the priority is to minimize the deviation between forecasted and actual plant output. For an ISO, there are two additional and more demanding priorities. The first is to anticipate changes in wind production as accurately as possible in the very short term (up to a few hours ahead). This is to enable an ISO to manage its grid operations and reserve capacity purchase decisions in an optimal fashion. For this purpose it is natural to consider persistence-type methods. However such methods are inherently limited in that they cannot predict changes in plant output that depart radically from recent trends, as might occur, for example, because of a passing weather front. Thus, in order to achieve the highest possible accuracy the methods should incorporate other data that may signal future trends, such as a conventional weather forecast or data from meteorological stations that lie "upstream" of the wind parks.

The second priority is to forecast the wind generation for the next day. The presumed goal is to enable an ISO to schedule reserve capacity as efficiently as possible. If so, then it may be less important to accurately forecast the *timing* of changes in wind generation than it is to forecast the *minimum* wind plant output during the peak load hours.

In general, a high degree of reliability and accuracy is required by ISOs and utility systems for wind plant forecasts. This is consistent with the usual high standard of reliability applied to all utility system operations. It is particularly important for the next-hour forecasts, because their accuracy declines relatively quickly the older the forecast becomes. The accuracy of next-day forecasts, in contrast, is not as sensitive to the age of the forecast.

2. The Forecasting Problem

The wind energy generation forecasting problem is closely linked to the problem of forecasting the variation of specific atmospheric variables (i.e. wind speed and direction, air density) over short time intervals and small spatial scales for a small volume of the atmosphere (the wind plant) for a variety of time horizons (i.e. look-ahead periods). In general, this is an enormously difficult problem because of the wide variety of spatial and temporal scales of atmospheric motion that play a role in determining the variation of the key parameters within the targeted forecast space-time volume. In order to understand the different issues involved in wind energy forecasting it is useful to divide the problem into three difference time scales: (1) very short-term (0-6 hrs); (2) short-term (6-72 hours), and (3) medium range (3-10 days).

The skill in very short-term forecasting is related to the prediction of small-scale atmospheric features (< 200 km in size) in the vicinity of the wind plant. The major difficulty is that very little data is typically gathered on the scale of these features. This means that it is usually difficult to define the spatial structure and extent of these features. Therefore, the only viable choice is often to infer information about these features from the time series of meteorological and generation data from the wind plant. As a result, real-time data from the wind plant is usually crucial to the performance of very short-term forecasts. In fact, the 0-6 hr time scale has been defined as the period when persistence forecasts will typically outperform wind energy forecasts derived solely from predictions of the regional atmospheric circulation (i.e. the information from the wind plant or the plant's immediate environment is more important than the regional information). Thus, the forecast standard for the very short-term forecast is a persistence forecast.

The ability to forecast the wind energy generation over short-term time scales (6-72 hours) is tied to the skill of forecasting regional scale atmospheric features. These features are often referred to as synoptic scale weather systems and are the ones that are typically depicted in newspaper and TV weather presentations. It is necessary to gather data over a large volume of the atmosphere in order to define the structure of these systems. This is usually done through the use of in situ or remote sensing measurement devices operated by an agency of a national government (such as the U.S. National Weather Service). The importance of measurements at the wind plant drastically decreases at the start of this period. The real-time plant data is able to make some contribution to forecast quality at the start of the period but it has little predictive value after about 12 to 18 hours. This is fundamentally because the information that determines the variations in meteorological parameters for time periods greater than 12 hours comes from locations that are many hundreds if not thousands of kilometers away. As a result, the forecast standard shifts from persistence to climatology (i.e. the average conditions for that location and season) during this period. A climatology forecast will typically outperform a persistence forecast for most locations after about 12 to 18 hours.

The skill of medium range (3-10 days) forecasts is typically tied to the ability to forecast continental, hemispheric and global scale atmospheric circulation systems. Of course, the regional and local features are superimposed upon these large scale features so in order to produce accurate medium range hourly predictions of wind energy production a prediction system must forecast all scales of atmospheric motion from global or hemispheric to local. At medium range time scale it is difficult to accurately predict the evolution of specific local-area or regional features that will affect the forecast target area. Therefore, most of the forecast skill is linked to prediction of general patterns that favor above average or below average winds for a substantial period of time (a day or more). The forecast standard for this time scale is a climatological forecast.

It should be noted that the distribution of atmospheric energy across the space and time scales varies substantially by region, season and atmospheric regime. This has important implications for predictability and forecast performance. If there is very little variability over a specific time scale then the absolute forecast performance is likely to be good but there is likely to be little skill over a simple persistence or climatology forecast on that time scale. Conversely, a situation with large variability over a given time scale will often result in lower absolute performance but a higher level of performance relative to simple persistence or climatology forecasts. The point is that one must be very careful in how one evaluates forecast performance and how performance is compared from location to location and season to season.

3. Forecasting Tools

There are two fundamental types of tools used in the forecasting process: (1) data gathering and (2) data processing. The data gathering tools include the wide range of measuring devices that provide data to the forecast process. These include measurements made at the wind plant itself as well as those in the local-area, regional and even global environment of the plant. The data processing tools serve to transform the measurement data into a forecast for the desired period of time. The data processing tools include physical and statistical atmospheric models as well as models of the relationship between meteorological conditions within the wind plant volume and the plant output (usually referred to as a plant output model). Except for some government agencies forecast providers typically control the data processing tools but not the data gathering tools and therefore they are typically in a position of making the best use out of whatever data is made available to them.

Due to the wide range of spatial and temporal scales that determine the variations in the wind energy power generation, it is necessary to use a diverse mix of data sources and types to achieve the best possible forecast performance. For wind energy forecasting, the most fundamental type of data is the time series of meteorological parameters and power generation from the wind plant itself. The generation data can be for the entire plant or for groups of turbines within the plant. The meteorological data typically consists of wind speed and direction and sometimes temperature, pressure and humidity data from sensors on one or more “met towers” within the plant’s boundaries. This data is typically gathered at the hub height of the turbines deployed in the plant. The additional detail provided by generation data by turbine group and multiple met towers can be very beneficial in developing a more precise and accurate relationship between the meteorological conditions and the plant output. The availability of this time series data alone is sufficient to make a somewhat skillful short-term forecast and at least a climatology-level forecast for the intermediate and long term

However, in order to achieve a higher level of forecast skill it is necessary to utilize data from beyond the plant’s boundaries. Meteorological data from in situ sensors deployed and operated by government agencies have been a traditional source of data for wind energy forecasting. These include sensors on surface-based met towers deployed mostly at airports and sensors carried aloft by weather balloons to provide information about the vertical profile of temperature, humidity, winds and pressure. The main problem with this data is that the spacing between measurements is too large (because of economic constraints) to adequately represent the small or even sometimes medium scale atmospheric features that are responsible for short-term variations in wind energy output. However, these in situ sensor networks do a better job of mapping most of the features that are responsible for most of the variability over 1 to 2 day ahead time scales. Unfortunately, there are large areas (such as the oceans) where very little in situ data is gathered because of the cost of maintaining such systems in those environments. This means that data coverage is far from uniform and that some regions have a lot less data upstream than others. This often results in poorer forecast performance in some areas. An example is the west coast of the United States. Forecast performance is often worse there than in the eastern part of the country because a large data sparse area (the Pacific Ocean) is located in the most frequent upstream direction (to the west) for this area.

The expectation is that remote sensing technology will eventually overcome these limitations of data resolution and coverage. Many types of atmospheric remote sensors have been developed and some have been deployed for operational use. These include Doppler radars, wind profilers (a type of fixed position Doppler radar), lidars, sodars and satellite-based radiometers. While all of these technologies have made contributions to the atmospheric forecasting process, each has had significant limitations that have reduced their level of impact on atmospheric forecast performance. However, remote sensing technology continues to rapidly move forward and there is still an expectation that the next generation of remote sensors to be deployed during the next few years will have a greater impact on forecast performance.

Data processing tools are the other major component of the forecast process. These are the mathematical (often called numerical) models that ingest data and generate predictions. There are four fundamental types of numerical models used in the wind energy forecasting process: (1) physical atmospheric models, (2) statistical atmospheric models, (3) wind plant output models, and (4) forecast ensemble models. There are many types of models within each of these major categories. A particular forecast system may employ one or more of each type of model in its forecast procedure.

Physical atmospheric models are based upon the fundamental physical principles of conservation of mass, momentum and energy and the equation of state for air. These models are actually a type of computational fluid dynamic (CFD) model that has been specially adapted to simulate the atmosphere. These models consist of a set of differential equations that are numerically solved on a three-dimensional data grid which has a finite resolution (i.e. the spacing between grid cells). There are many types of these models. They are based on the same basic physical principles but they differ in how the grids are structured, how the equations are numerically solved and how sub-grid scale processes (i.e. things that happen on scales which are smaller than the grid cells) are represented.

Physics-based atmospheric models fall into two broad categories: prognostic and diagnostic. Prognostic models are formulated to use their equations to step forward from an initial state and make predictions of the future state of the atmosphere. It is necessary to specify an initial state to start this forecast process. An initial state consists of a value for each of the model's variables at each grid cell. This is produced by processing all of the raw atmospheric data from the various sensor systems that were described earlier in this section. There are many three-dimensional prognostic atmospheric models in use. These include the Mesoscale Atmospheric Simulation System (MASS) model developed by MESO, Inc. and the MM5 model developed by Pennsylvania State University and the National Center for Atmospheric Research

(NCAR). Diagnostic models use a similar but often simplified set of physical equations to estimate the values of variables at locations where there is no data from locations where there is data. These models can be used to add more resolution to forecast simulations made with a prognostic model at a lower computational cost than reducing the size of the grid cells of the prognostic model The simplifying assumptions utilized to create the diagnostic model will typically limit its performance below that of a prognostic model run at a similar resolution.

Statistical atmospheric models are simply statistical techniques used for atmospheric applications. They are "atmospheric" models in the sense that atmospheric data is used as input and the output is an atmospheric variable or quantity that is linked to an atmospheric variable (such as wind energy output). Statistical models operate by creating a set of empirical equations from a sample of predictor and predictand (the quantity to be predicted) data called a "training sample". The form of the equations is dependent on the type of model that is used. Typically, the equations have numerical coefficients that must be determined. A modeling procedure uses an optimization scheme to select the coefficient values that yield the best relationship between the predictors and the predictand. The meaning of "best" in this context depends upon what optimization criterion is employed. An example of optimization criteria is the lowest mean absolute error or the lowest mean squared error. Once the coefficients are determined from the training sample, the resulting equations can be used to produce a forecast by inserting the current values of the predictors and calculating the value of the predictand. There are an enormous number of statistical models available for this type of an application. The most popular ones for atmospheric science applications are multiple linear regression and neural networks.

Statistical models are used in a number of different ways in wind energy forecast systems. In one mode, they can be used to adjust the predictions from the physics-based models. This mode is commonly called Model Output Statistics (MOS). However, they also can be used to make predictions directly from measured data. For example, a time series of generation data can be used to train a statistical model and make predictions of the future generation.

Wind plant output models are the relationships between the meteorological variables at the wind plant site and the plant's energy output. They can be formulated as physical models, statistical models or a hybrid of both types of models. In a statistical approach, the parameters measured by sensors on the plant's meteorological towers typically serve as the predictors and the power generation is the predictand. The simplest plant output model is a relationship between the wind speed measured at a meteorological tower and the total plant output. The result is a plant-scale equivalent to the "power curve" of an individual turbine. This simple model can be extended, for example, by developing a separate relationship for ranges of wind directions. This may be useful in accounting for the orientation of the turbine layout relative to the wind direction. For example, the power production may be different when the wind blows along a row of turbines than when it blows across a row.

Wind plant models can also be formulated as physical models. In this approach, the variations in wind flow within the wind plant, the interaction of the wind with the turbines and the effect of turbine wakes on other turbines are explicitly modeled. This requires detailed information about the layout of turbines in the plant, the properties (terrain, roughness etc.) of the earth's surface within the plant and information about the turbine specifications. The physical models have the advantage of being able to produce a power generation forecast without a training sample. They also can explicitly account for changes in the operating structure of a plant such as when turbines go out of service. They also, in theory, account for more of the details of the plant-scale variation in wind and its impact on power production. However, these models are typically much more complex than statistical models and require much detailed data about the plant which may not be easily available. Also, as with almost all physical models, there are likely to be systematic errors in the forecasts due to simplifying assumptions included in the model physics, limited resolution or the inaccuracies in the input data (e.g. inaccurate roughness or terrain etc.). In most applications it is necessary to use a statistical model to adjust the forecasts of a physical plant output model to remove these systematic errors.

The typical use of plant output models in the forecast process is to convert a wind speed prediction for one or more meteorological towers within the wind plant area into a power generation forecast for the plant. However, it is not necessary to have an explicit wind plant output model in a forecast system since it is possible to go directly from external predictors to a power output forecast through the use of an atmospheric statistical model.

Forecast ensemble models are statistical models that produce an optimal forecast by compositing forecasts from a number of different techniques. The use of forecast ensemble models is based on research that has demonstrated that a composite of forecasts from an appropriate ensemble of forecast generating techniques is often superior to those produced by any one member of the ensemble. The method is schematically depicted in Figure 1. The fundamental concept is that if the errors in the forecasts produced by the different methods are unbiased and have a low degree of correlation with one another, the random errors from the individual forecasts will tend to offset each other, with the result that a composite of the forecasts will have a lower error than any individual forecast. If all of the input forecasts are highly correlated the impact of ensembling will be minimal. This means that the underlying forecast methods must be quite different in how they construct the relationships between the raw observational data and their forecasts or the type or amount of input data going into the methods must be significantly different. This "ensemble effect" is a well-known technique used by meteorologists in intermediate and extended range forecasting. The spread of the forecasts produced by the ensemble may also be related to the forecast uncertainty if the differences in the ensemble members are related to the primary factors that introduce uncertainty into the forecasts.

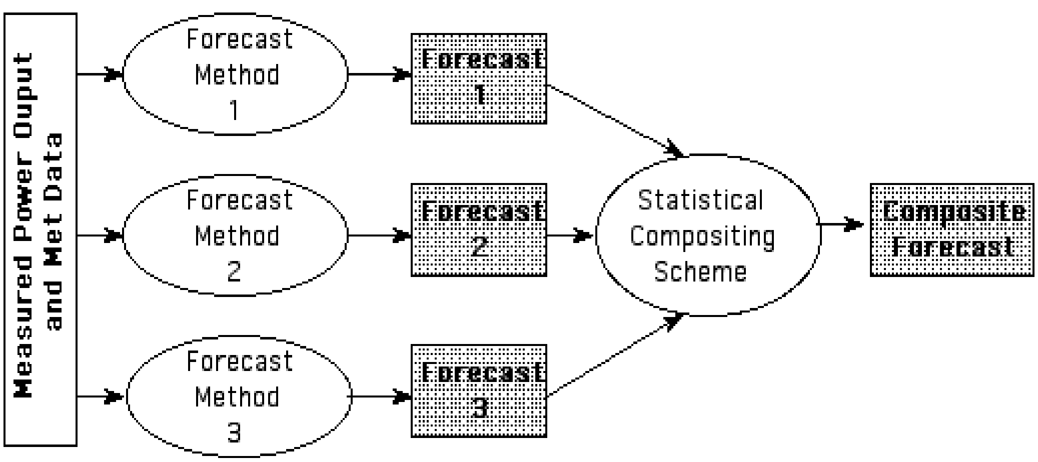

Figure 1. *A schematic depiction of the ensemble technique. This arrangement applies to both short-term and next-day forecasts.*

There are two fundamental strategies that can be used to generate an ensemble of forecasts. One strategy is to use the same forecast models and vary the input data within their range of uncertainty. The other is to use the same input data and to employ different forecast models or different configurations of the same model. The relative value of either strategy depends upon the sources of uncertainty in the forecast procedure. In practice, the sources of uncertainty vary with location, season and other factors, and thus the choice of the ensemble components and the number of members must be determined from experience and experimentation.

This brief overview of forecast tools indicates that there is a large and diverse pool of tools that can be used to generate wind energy forecasts. The challenge is to use the optimal set of tools and configurations for a specific forecast application.

4. Forecast Evaluation

Although it may seem straightforward, there are a number of complex issues associated with the evaluation of wind energy forecasts. The most significant issue is which parameter(s) should be used as the metric(s) for forecast performance. One's choice of metrics can have a significant impact on one's impression of forecast performance.

A wide variety of metrics are in common use and no doubt many more could be devised. One fundamental distinction is absolute vs relative performance. An absolute performance metric provides a measure of the performance of a forecast system that is independent of other forecasts. Examples of absolute performance metrics are root mean square error (RMSE), mean absolute error (MAE) and median error (MDE). A relative performance metric is a measure of the performance of a forecast method relative to another method. Typically the other method is a reference forecast such as persistence or climatology. A popular relative metric is the persistence-based skill score, which is the percentage reduction in the MAE of a persistence forecast that is achieved by a particular forecast method.

A second distinction is the sensitivity to different portions of the error frequency distribution. Some parameters are much more sensitive to outliers, i.e. forecasts with anomalously large or small errors. For example the RMSE is quite sensitive to outliers while the MDE is quite insensitive. The sensitivity of the MAE parameter is between these two extremes.

In addition to the issue of different metrics providing a different picture of performance, there is also the issue that a forecast system can be tuned to produce better performance for a specific metric while possibly degrading the performance for other metrics. This can be done by formulating a statistical technique to minimize the value of a specified optimization (or cost) function. This might be considered to be "gaming" by the forecast provider or it might be viewed as the customization of the forecast system to meet the needs of a specific application. However, the underlying issue is whether the evaluation metric is really linked to the user's cost function. If it is, then it probably makes sense to optimize the forecast system for that metric.

An example of the wide range of perspectives provided by different forecast metrics is provided in Table 1. This table lists the values for a suite of forecast metrics for the performance of 1 to 48 hour forecasts of power output and wind speed for the month of October 2001 for a wind plant in the San Gorgonio Pass of California. Three types of forecasts are compared: eWind forecasts developed by TrueWind Solutions; persistence; and climatology. Different pictures of the absolute and relative forecast performance emerge depending on which metrics are considered. For example, MAE as a percentage of the rated capacity is 14.7% for the first 24-hour period. However, the RMSE is 20.8% and the MDE is 10.3%.

Table 1
Power output and wind speed verification statistics for a wind plant in the San Gorgonio Pass of Southern California.

Verification Statistic	Power Output					
	Month: Oct-01			% Capacity 31.6%		
	Hours 1-24			Hours 25-48		
	eWind	Persistence	Climatology	**eWind**	Persistence	Climatology
MAE %Rated	14.7%	22.3%	28.4%	16.0%	32.4%	28.4%
MAE %Mean	46.4%	70.5%	89.7%	50.5%	102.6%	88.5%
MAE % Std Dev	47.7%	72.4%	92.2%	51.9%	105.5%	90.9%
RMSE-% Rated	20.8%	31.0%	31.9%	22.9%	42.1%	31.6%
Median % Rated	10.3%	16.7%	28.4%	10.9%	27.3%	28.3%
Correlation	0.75	0.47	0.11	0.63	0.00	0.11
Skill-Pers	34.1%	0.0%	-27.3%	50.8%	0.0%	13.8%
Skill-Climate	48.3%	21.5%	0.0%	43.0%	-16.0%	0.0%
Verification Statistic	**Wind Speed - Met Tower**					
	Month: Oct-01			Avg Spd (m/s) 8.83		
	Hours 1-24			Hours 25-48		
	eWind	Persistence	Climatology	**eWind**	Persistence	Climatology
MAE	2.52	3.87	3.86	2.70	5.59	3.86
MAE %Mean	28.5%	43.8%	43.7%	30.5%	63.3%	43.7%
MAE % Std Dev	55.8%	85.8%	86.6%	59.8%	123.9%	85.5%
RMSE	3.13	4.90	4.62	3.58	6.91	4.59
Median	2.10	3.10	3.82	2.00	4.70	3.80
Correlation	0.72	0.51	0.04	0.63	-0.07	0.04
Skill-Pers	35.0%	0.0%	0.4%	51.8%	0.0%	31.0%
Skill-Climate	34.8%	-0.4%	0.0%	30.1%	-44.9%	0.0%

5. State-of-the-art: "Next-Hour" Forecasting

There are a wide variety of methods that have been or are being used to produce very short-term ("next-hour") wind energy generation forecasts. Figure 2 provides a schematic depiction of the many components of the very short-term forecasting process and the ways they can be linked together to produce forecasts.

The simplest type of very short-term forecasts utilize a time series of power generation data from the wind plant and a statistical procedure such as multiple linear regression or a neural network to generate predictions of the future power output. These are often referred to as "persistence" or "autoregressive" models since their only source of information is the history of the plant's power output. These types of models can be enhanced through the use of a time series of meteorological data from the meteorological towers within the wind plant. The addition of this data can be handled in two ways. In the first approach the meteorological data is added to the pool of predictors and the power generation is predicted directly from the statistical model. In the second approach the meteorological data is used to forecast the meteorological inputs into a separate wind plant output model. The wind plant output model then uses these inputs to create an energy generation prediction. The second approach may have an advantage if there is more than one met tower within the plant because it may be possible to capture some of the variability in meteorological conditions within the plant and hence produce a better energy generation forecast. Sophisticated statistical models such as neural networks may be able to find more subtle and complex relationships in the time series data and thereby generate better forecasts than simpler models such as linear regression. However, due to the fact that sophisticated statistical models usually have more adjustable parameters, they are more prone to "overfitting" problems if the training sample is not sufficiently large. Ultimately, all of these methods are limited by the fact that the input information is derived only from the history of conditions at the wind plant.

The next level of sophistication is to use multiple external data sources. The additional data sources can be used as input to the same types of statistical models used in the autoregressive approach. However, the number of predictors is larger. The additional data sources could include meteorological data from nearby meteorological towers or remote sensing systems. Another possibility is to use forecast output from a regional scale physical model. These models provide information about the larger scale trends in meteorological parameters but do not incorporate local area data and typically do not have the ability to resolve the local atmospheric and surface features that are critical to very short-term forecasting. However, some large-scale trends are well correlated with a local-scale response and hence the regional model data can at times add skill to the very short-term forecasts.

A possibility that has yet to be thoroughly tested for very short-term wind power forecasting is to use a physical model with a high resolution grid to produce very short-term forecast simulations for the local-area surrounding the wind plant. In this approach, all of the available local-area data is assimilated into the initial state used to start the physical model simulation. This type of procedure has the potential to simulate the atmospheric features that cause the wind variations in the vicinity of the wind plant. The output data from this local-area simulation is then fed into a MOS procedure. The MOS algorithm selects the best performing predictors from the large volume of physical model data and generates predictions of the wind speed and direction at the wind plant meteorological towers. These predictions are then fed into a wind plant power output model to generate power output predictions. This method is a local-scale analog of the regional scale forecast procedures that has been used quite successfully for 1 to 2 day forecasting.

Another tool that can be used in the very short-term prediction process is a forecast ensemble model. As noted earlier, this is a statistical model that generates a composite forecast from a series of input forecasts generated by different forecast methods.

After one completes an examination of the various methods that have been or could be used in the very short-term forecast process, the obvious questions are (1) what is the typical level of performance that can be expected from very short-term forecast methods? and (2) what is the variation in performance due to differences in methods, locations, seasons and weather regimes? Unfortunately, there has been no comprehensive controlled study of the performance of a broad spectrum of very short-term wind energy forecast methods over a diverse mix of atmospheric conditions. Most of the forecast performance evaluations have been done by forecast providers or researchers and not by an independent third party. It is difficult to draw conclusion from this array of evaluations because the evaluation methods, locations, and times are different.

TrueWind Solutions engages in ongoing forecast system development and evaluation. An example of the performance of several short-term forecasts for July 2003 for a wind plant in the San Gorgonio Pass of California is presented in Figure 3. This performance is typical for this site and season but TrueWind's experience has indicated that there can be large variations in performance from site to site and season to season. In this example, all of the methods yield a small improvement over persistence during the first couple of hours of the forecast period. The methods that use regional physical model data become significantly better than persistence after about 4 hours.

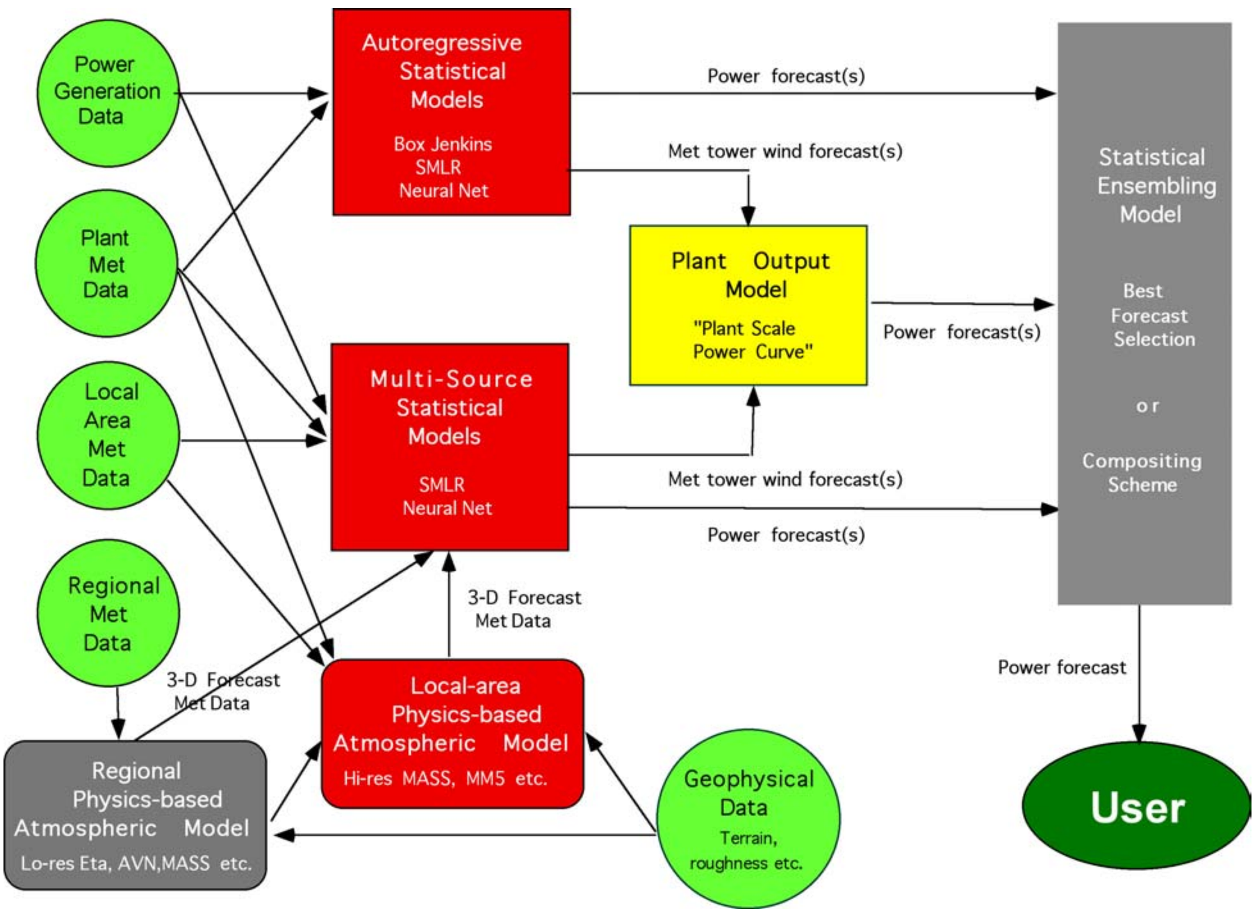

Figure 2. *A schematic depiction of the interrelationship of the components of a very short-term forecast system.*

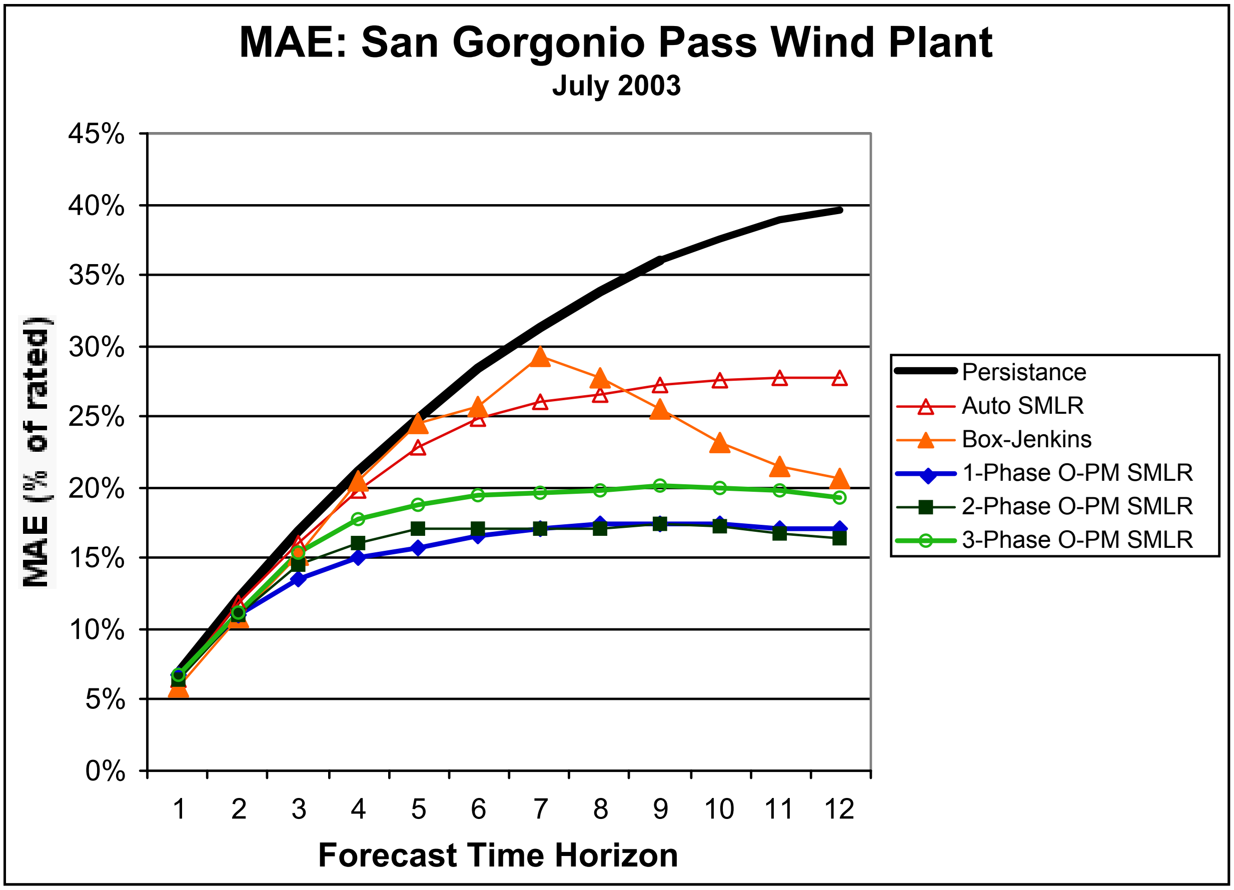

Figure 3. *The mean absolute error by forecast hour for July 2003 for 5 very short-term forecast methods and a simple persistence benchmark forecast for a wind plant in San Gorgonio Pass. The "SMLR" acronym refers to a screening multiple linear regression procedure. "O-PM" refers to the use of both observational and regional physical model data as input to the statistical procedures.*

6. State-of-the-art: Day-Ahead Forecasting

Short-term (day-ahead) forecast methods use essentially the same tools as very short-term forecast techniques. However, there are two important differences: (1) the importance of real-time data from the wind plant and its immediate environment is significantly reduced; and (2) regional and sub-regional simulations with a physics-based atmospheric model play a much more significant role in the forecast process.

Almost all short-term forecast procedures begin with the grid point output from a regional scale physics-based atmospheric model. Typically, these are executed at a national forecast center such as the National Center for Environmental Prediction (NCEP) operated by the U.S. National Weather Service. These models ingest data from a wide variety of sources over a large area and produce forecasts of regional scale weather system for a several day period. However, these models do not resolve the physical processes occurring in the local or mesoscale areas around individual wind plants. The three-dimensional output data from the regional-scale forecast simulations is the basic input into most intermediate-term wind energy forecast systems.

The forecast methods differ substantially from this point. Some forecast procedures attempt to go directly from the regional-scale forecast data to the local scale through the use of either diagnostic physical models, statistical models or a combination of both. The Prediktor system developed by the Risoe National Laboratory in Denmark uses this approach. The main drawback of this approach is that it misses the processes that occur at the sub-regional or mesoscale (i.e. the scale between the large-scale weather systems and the local scale). An alternate approach is to execute sub-regional scale simulations with a physics-based model to account for the mesoscale processes. This is the approach used by TrueWind Solutions in their eWind system and a couple of other North American forecast providers. A schematic depiction of the eWind system is presented in Figure 4. This approach has had considerable success in forecasting the variations in winds attributable to mesoscale processes but it has a much higher computational cost than the regional to local forecast schemes. Both the regional-to-local and mesoscale simulation approaches employ statistical models (typically referred to as MOS) to predict the wind speed and direction at the wind plant's meteorological towers. The predictors are based on either the output from the mesoscale simulations (mesoscale approach) or from the regional or diagnostic physical models (regional-to-local approach). Although it is possible to predict the energy generation directly from the physical model output through the MOS process, most forecast systems are configured to produce wind predictions for the meteorological tower sites from the MOS process and to then use those predictions as input to a wind plant output model to create the energy generation forecasts. The wind plant output model can be either physical or statistical. The Prediktor system has the option to use either a physical model in combination with a second MOS procedure to remove any systematic errors or a purely statistical scheme. The eWind system uses a statistical wind plant output model.

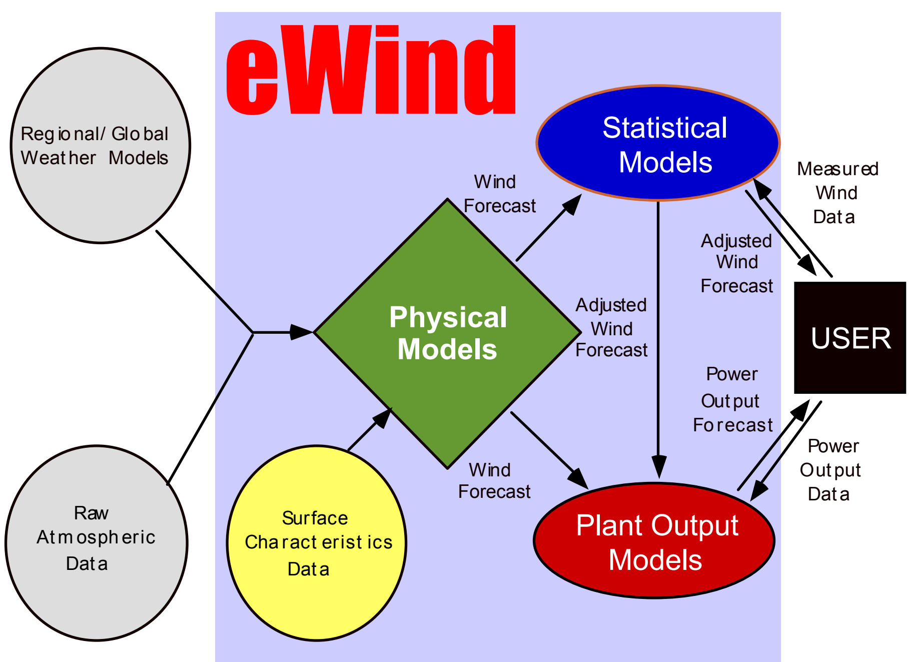

Figure 4. *A schematic depiction of the major components of the eWind intermediate-term wind power forecast system.*

As in the case of very short-term forecast performance, a quantitative assessment of the state of the art in short-term wind energy generation forecast performance is difficult to obtain because most of the evaluations are done by individual forecast providers and researchers. The methods, locations and time periods used in these forecast performance evaluations vary substantially and, therefore, it is difficult to determine the causes of differences in performance.

Fortunately, there is one recently completed investigation of day-ahead type forecast performance that was managed by a third party. This project was funded by the California Energy Commission (CEC) and managed by the Electric Power Research Institute (EPRI). A complete summary of the results from this project can be found in EPRI Report # 1007339 [1]. The objective was to assess the state of the art in wind energy forecasting for the State of California. Two forecast providers participated in the project. Each used their own forecast system to produce 1-48 hour wind energy forecasts for two wind projects in California for a 1–year period extending from October 2001 to September 2002. One forecast provider was Risoe National Laboratory from Denmark. They used their Prediktor system. The other provider was TrueWind Solutions, LLC from Albany, NY. They employed the eWind forecast system. The differences between the approaches used by these two systems were outlined earlier in this section. One of the participating wind projects was the 66 MW Mountain View wind plant in San Gorgonio Pass, which is located just to the east of the Los Angeles Basin in southern California. The other project was a 90 MW plant located in the Altamont Pass, which is just to the east of the San Francisco Bay Area in northern California.

A summary of the forecast performance results from this project is presented in Table 2. The performance statistics in this table are for all forecast hours (i.e. 1-48) and for the entire 12–month evaluation period. The mean absolute error (MAE) as a percentage of installed capacity is in the 14% to 21% range. This is a typical range for 1 to 2 day forecast performance. The percentage MAE of both forecast systems was lower for the Altamont Pass plant. However, the Risoe system showed a greater difference in forecast performance between the two plants than the TrueWind system.

Figures 5 and 6 depict the MAE of TrueWind's persistence and climatology forecasts by forecast hour for each of the plants. It can be seen that persistence forecasts are the best for the first few hours for both plants. This is because no real-time information from the plant or its immediate environment was available for use in the forecast process. After the initial period the TrueWind forecasts outperform the persistence and climatology forecasts by a substantial margin. This is a typical pattern of forecast performance for most sites. These figures also provide an indication of the rate of forecast error growth as the forecast look-ahead period increases. The error growth for the San Gorgonio Pass wind plant (2% of installed capacity per 24 hours) is approximately twice as large as the rate for the Altamont Pass plant. This difference is most likely attributable to the physical properties of the site and its immediate environment and the differences in the weather regimes affecting the two areas over the course of the year.

This study served to document the expected level of performance of short-term wind energy forecast systems. It indicated that state-of-art forecasts systems have considerable skill over climatology and persistence forecasts for 1 to 2 day periods. It also demonstrated that 1 to 2 day forecast performance can vary substantially by location, season and the attributes of the forecast system used to generate the predictions.

Table 2

A summary of the forecast performance results from the EPRI-CEC project

Parameter	Risoe	TrueWind
Mountain View (66 MW rated)		
Mean Error (kWh)	2,888	628
MAE(kWh)	14,305	11,834
MAE(% of rated)	21.7%	17.9%
Skill vs. Persistence (%)	9.5%	32.6%
Skill vs. Climatology (%)	19.8%	33.7%
Altamont (90 MW rated)		
Mean Error (kWh)	702	631
MAE(kWh)	12,985	12,438
MAE(% of rated)	14.4%	13.8%
Skill vs. Persistence (%)	21.6%	30.8%
Skill vs. Climatology (%)	26.2%	29.6%

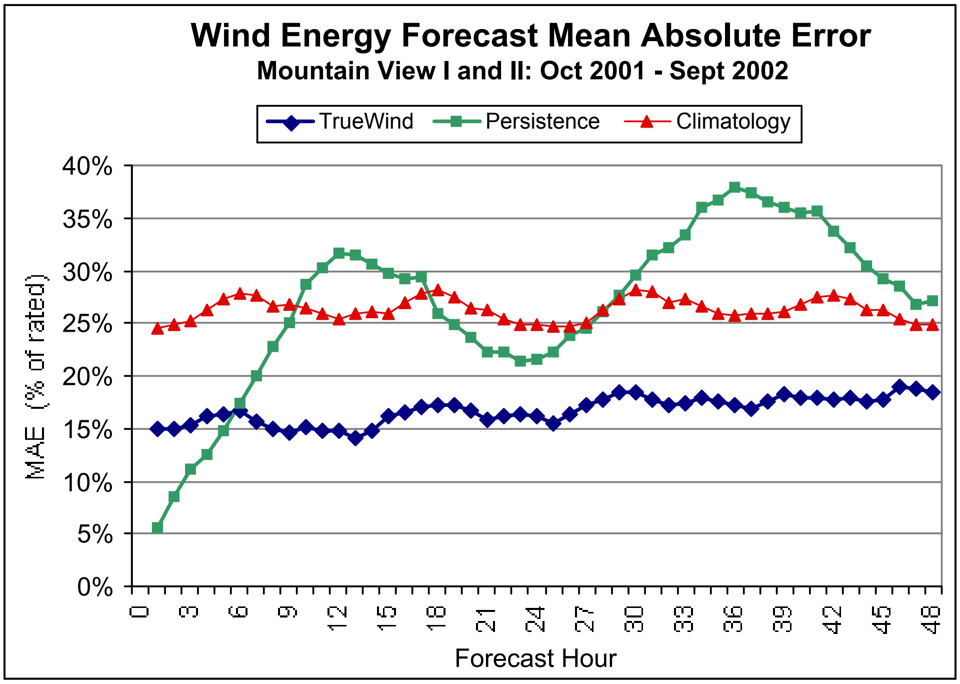

Figure 5. *The mean absolute error by forecast hour for 12 months of TrueWind's (eWind) persistence and climatology energy generation forecasts for a wind plant in San Gorgonio Pass.*

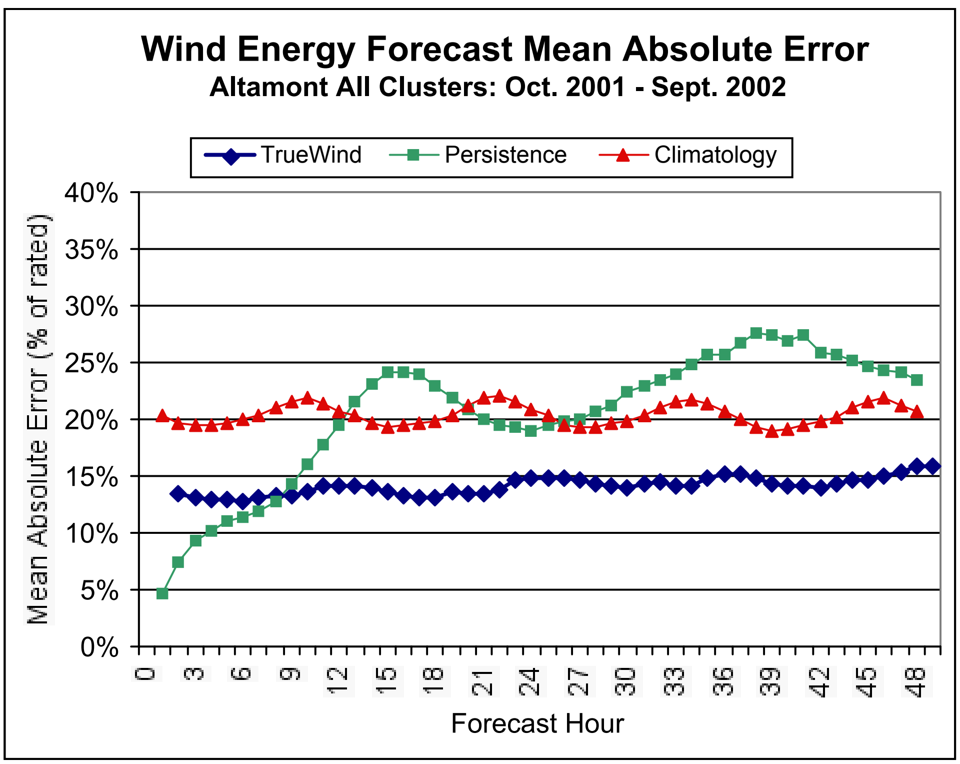

Figure 6. *The mean absolute error by forecast hour for 12 months of TrueWind's (eWind) persistence and climatology energy generation forecasts for a wind plant in Altamont Pass.*

7. Potential for Improved Forecast Performance

Although both very short-term (next hour) and short-term (day-ahead) forecasts made with state-of-the-art forecasting systems currently exhibit considerable skill relative to benchmark persistence and climatology forecasts, there are still many opportunities for forecast improvement on these time scales. There is also an opportunity to extend the range of useful (i.e., those that have skill over climatology) hourly energy forecasts out to 3 to 5 days, which is in the realm of what is called "medium range" forecasting. This section gives an overview of (1) how forecast improvement at each of the three major time scales is likely to be achieved and (2) provides an estimate of the amount of improvement that may reasonably be expected over the next 10 to 15 years. The discussion begins with the prospects for forecast improvement on the medium range time scale (3-10 days) and then addresses the possibilities for improvement on the "day ahead" and "next hour" time scales.

Currently, hourly wind forecasts and the associated energy generation forecasts beyond approximately 3 days have very little skill over a climatology forecast, although forecasts of the wind speed and energy generation for longer time intervals (e.g. daily averages) do have some skill over a climatology forecast out to 6 or 7 days. As forecast technology improves over the next 10 to 15 years, it is likely that forecasts beyond 3 days will become useful to the wind energy community. The chart in Figure 7 provides a perspective on the long-term trend in forecast improvement and what it may mean for future performance. This chart [2] depicts the yearly average S1 score for forecasts of the mean sea level pressure gradients made by several different forecast models run by the United States National Weather Service during approximately the last 50 years of the 20th century. The S1 score is a measure of the relative error of a forecasted spatial gradient of a parameter. Lower S1 scores indicate more accurate forecasts. A forecast that produces an S1 score of about 70 is generally considered useless while an S1 score of about 20 is generally considered to be a perfect forecast for most practical purposes.

The reader would be justified in wondering about the relevance of S1 scores of mean sea level pressure gradients to the discussion of the performance of wind energy forecasts. First, the near-surface wind speed and direction is strongly correlated to the mean sea level pressure gradient for locations that are within several hundred meters of sea level. Therefore, improvements in the forecasts of sea level pressure gradients are closely linked to improvements in near-surface wind speed and direction forecasts. Second, the S1 score is a measure of forecast accuracy that has been used by meteorologists since the late 1940's. The 36-hr S1 forecast verification scores for mean sea level pressure gradients constitute the longest continuous record of forecast verification anywhere [2]. Therefore, it is a metric that can be used to define the trend in forecast performance over a long period of time and provide some guidance about future performance. However, it should be noted that these S1 scores refer to the ability to forecast large scale pressure gradients associated with regional and continental scale weather systems and not the mesoscale or microscale pressure gradients that are responsible for variations in winds in the local area around wind plants.

The half-century of S1 scores that are depicted in Figure 7 clearly indicate tremendous progress in the ability to forecast sea level pressure gradients. The line labeled "36 hr (AVN)" denotes the S1 score for what was considered to be the best National Weather Service forecast procedure during each time interval of this 50-year period. In the late 1940's the forecasts were made through a manual procedure based on extrapolation and subjective pattern recognition techniques. The S1 scores for a 36-hour forecast at that time were in the mid 60's, which is slightly better than useless. The first numerical weather prediction models went into operational use in the middle 1950s and the S1 scores began to steadily improve thereafter. The improvements in forecast performance after the 1960s have been attributed to: (1) better observed data; (2) better methods for incorporating data into atmospheric models; and (3) better atmospheric models themselves. The improvement was persistent if not dramatic from the early 1960s through the end of the 1990s. The S1 scores were routinely in the low to mid 30's by the end of the century. The line labeled "72 hr AVN" depicts the S1 scores for 72-hour forecasts. These forecasts were not even attempted by the National Weather Service in the 1950s and 1960s. The first scores appear on the chart in the late 1970s and they are near the "useless" line. However, a significant improvement occurred in the mid-1980s and the S1 scores for a 72-hour pressure gradient forecast were typically near 50 by the mid 1990s. This was the same level of skill achieved for the 36-hour forecast around 1980. Thus, in the mid 1990s the 72-hour forecast of the mean sea level pressure gradient was typically about as good as the 36-hour forecast was in 1980.

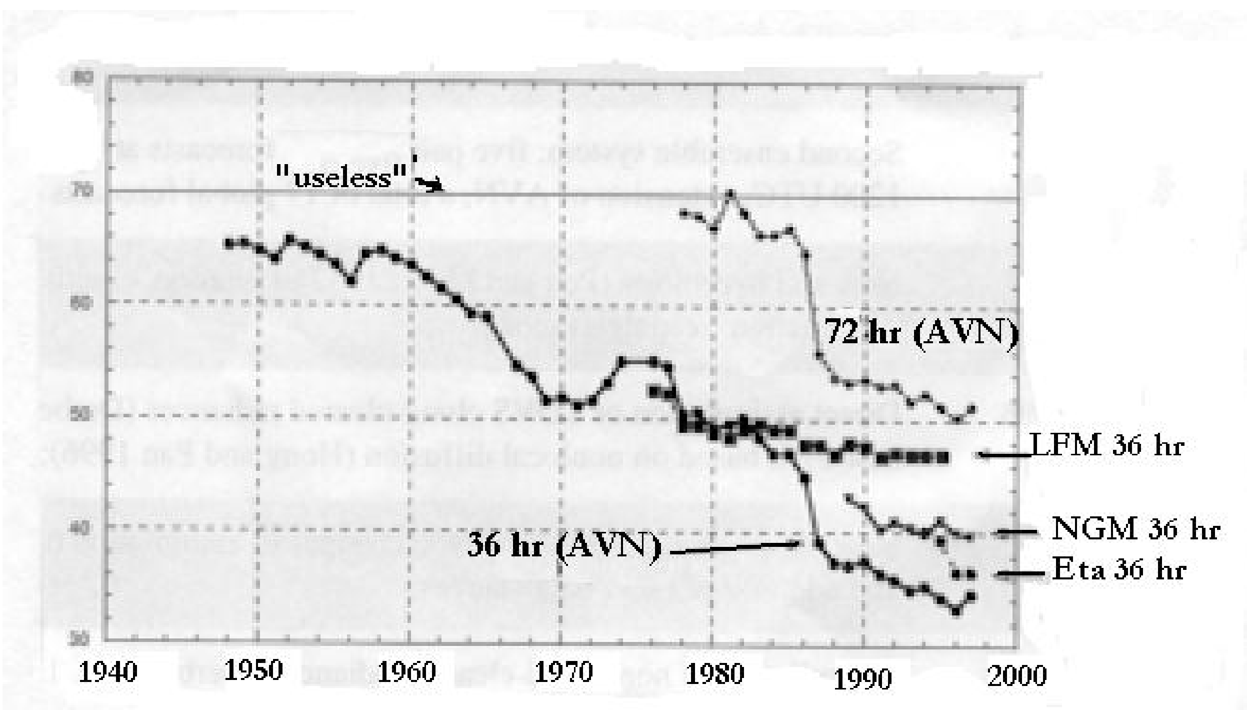

Figure 7. *A depiction of the yearly average S1 scores for forecasts of the mean sea level pressure gradient over North America produced by several different United States National Weather Service models (AVN, LFM, NGM and Eta) during the second half of the 20th century from [2]. The S1 score is a measure of the relative error of the gradient of a parameter over a specified region. A lower score indicates a more accurate forecast. The mean sea level pressure gradient is strongly correlated with the near surface wind speed at most locations within several hundred meters of sea level.*

One can extrapolate these rates of forecast performance improvement to obtain an estimate of the likely improvement in wind forecast skill over the next 10 to 15 years. The rate of forecast improvement inferred from the S1 data in Figure 7 suggests that the performance level of a 36-hour forecast in the 2000-2003 era would be achieved for a 72-hour forecast by approximately 2015 and the performance level of the 72-hour forecast in the 2000-2003 period might be achieved for an 108-hr (4.5 day) forecast by 2015. What does this mean for wind energy forecasts? Currently, a typical 36-hour forecast of the hourly energy generation (note Figures 5 and 6) has a mean absolute error of about 13-18% of a plant's installed capacity and a skill score (% reduction in mean absolute error) of about 30% over a climatolgy forecast. Therefore, this level of performance is likely for a 72-hour forecast by 2015. At present, a 72-hour forecast of the hourly energy output of a wind plant is near the end of the time period for which a forecast has skill over a climatology forecast. At this range the typical MAE is between 20 and 25% (note the climatology forecast in Figures 5 and 6) and the skill over climatology is a few percent. This level of performance is a reasonable expectation for a 108-hour (4.5-day) forecast by 2015.

There doesn't seem to be a reason to expect that these extrapolated improvements in forecast performance will not be achieved, since research and innovation continues to be very active in all three of the previously mentioned areas which have been driving forces behind the improvements depicted in Figure 7: (1) better observed data; (2) better methods for incorporating data into atmospheric models; and (3) better atmospheric models. There is an expectation of an acceleration of improvement in remotely-sensed data mostly as a result of better instrumentation aboard geostationary and polar-orbiting satellites. Improved techniques of incorporating various types of data into regional and global scale models are being developed. Finally, the research community continues to develop and improve the physics-based atmospheric numerical models, benefiting particularly from the wide range of modeling groups in the government, university and private sectors. Underlying all of this is the relentless advance of computer technology, making more powerful computers available at lower costs, allowing the implementation of ever-more-sophisticated models directed at more specific problems such as wind energy. Research is also underway in the development of new forecasting techniques. It has already been shown that the ensemble technique of combining 10 or more individual forecast simulations can produce better forecasts than conventional single-simulation forecasts for beyond five days. With very active research in this area, it can be expected that the ensemble approach will be more widely used to improve the accuracy of shorter-term forecasts as well.

The challenges of day-ahead forecasting are conceptually similar to those associated with the medium range forecasting task. However, the manifestation of the issues is different because the time and space scales are different. The skill in day-ahead forecasting is mostly related to the prediction of regional scale and mesoscale atmospheric features. The use of conventional atmospheric data and physics-based atmospheric models have proven to be an effective tool for this application.

The current expectation is that the bulk of future improvements to this process will come from improvements in regional physics-based atmospheric models and improvements in the amount and quality of regional atmospheric data that is available to initialize these models. A new generation of high-resolution atmospheric models is currently under development. These models employ more advanced numerical techniques that permit a more accurate solution to the differential equations that are the basis for these models. They also utilize more advanced representations of atmospheric physics and employ more sophisticated data assimilation techniques. These new models include the Weather Research and Forecasting (WRF) model being developed by a consortium of government and academic agencies and institutions and the Operational Multi-scale Environment Model with Grid Adaptivity (OMEGA) being developed by a partnership of private companies.

The main hope for the improvement of the regional atmospheric data required to initialize the next generation of physics-based mesoscale atmospheric models is remote sensing. The expectation is that more sophisticated satellite-based sensors will be deployed over the next few years. These will provide more accurate and more detailed data describing the state of the regional atmosphere and that should translate into improved 1 to 2 day wind energy forecasts.

As noted earlier, a third component of the short-term forecast process is the MOS procedure. This is the link between the grid point data that comes from the physics-based atmospheric models and the quantity to be predicted. Most current MOS procedures use a fairly traditional multiple linear regression approach to create the MOS relationships. However, there may be some forecast accuracy benefit to the use of a more advanced statistical model (e.g. a neural network) for this purpose.

The development of an ability to more accurately simulate mesoscale features over a 24 to 36 hour period will help improve the quality of 1 to 2 day forecasts. A reasonable expectation is that in 10 years, the mesoscale features will be forecasted as well as they are now forecasted 6 to 12 hours in advance. Thus, the performance of 36-hour wind energy forecasts in the year 2015 is likely to be at the level now achieved for 6 to 12 hour forecasts. This translates into an MAE of about 10-15% of installed capacity and a skill score of about 40% over a climatology forecast for a typical midlatitude wind plant.

The skill of very short-term forecasts is mostly limited by the inability to (1) define the initial (at the time the forecast is made) structure of the atmosphere in the local area (0-200 km) around a wind plant (i.e., what is happening now?) and (2) extract the complex relationships between the measured data that serve as input to the forecast process (i.e. predictors) and the wind energy production (i.e., how is what is happening now related to what will happen in the future?).

It is intuitively obvious that the “what is happening now” part can be addressed by obtaining more atmospheric data from the local area surrounding the plant. The issue is what is the most cost-effective way of doing this? One suggestion that has been made numerous times in recent years has been to install "upwind" meteorological towers to provide information about atmospheric features that are approaching a wind plant. A paper presented at the WindPower 2003 Conference [3] demonstrated some forecast skill improvement for a wind plant on the Oregon-Washington border through the use of upstream-type meteorological tower data in the Columbia River Basin. Although there was some success in this instance, there are a number of issues with this approach including where to locate the towers and the cost of installation and maintenance. This approach may be cost effective in a narrow pass type of environment where the upstream direction is almost always in one of the two directions along the axis of the pass. However, it may be less cost effective in a more open setting in which the distribution of wind directions is more uniform. A possible way to achieve optimal cost effectiveness is to use a set of physics-based numerical model simulations of historical cases to identify the surrounding sites that are highly correlated with the variations in wind at the wind plant and to install measuring equipment at one or two of the most highly correlated sites.

An alternative approach is to use some type of remotely-sensed data. This could include surface-based remote sensing systems such as wind profilers, Doppler radars, lidars or sodars. These provide wind data over a limited atmospheric volume at a relatively high cost. It would most likely not be cost effective to install such equipment (only used for forecast applications) for a wind plant. However, it may be useful to use data from these devices if they are already operating in a region for other purposes and the data can be made available to the wind energy forecast process. Another possibility is data from satellite-based sensors. These sensors typically measure the amount of radiation coming from the atmosphere in multiple bands or channels that correspond to specific electromagnetic wave frequencies. These radiation measurements can be used to obtain estimates of temperature and moisture profiles of the atmosphere. They also can be used to provide some information about winds by tracking clouds.

The other part of the problem is to figure out how to develop better relationships between what is happening now and what will happen in the next few hours. One approach is to employ more advanced statistical models and to optimize their type and configuration for the wind energy-forecasting problem. Techniques, such as neural networks and fuzzy logic clustering, are possibilities. These may be able to find more subtle and complex relationships between the raw input data and the quantities to be predicted. However, these advanced statistical approaches are not magic. They typically carry a high computational cost. They also require large training samples to find subtle and complex relationships and to avoid "overfitting".

Another approach to mapping the relationship between the growing volume of input data and the variables to be forecasted is to use a high resolution physics-based model to assimilate local-area data and generate a very short-term three-dimensional simulation of the atmospheric conditions surrounding the wind plant. In this approach, the fundamental principles of physics (i.e. conservation of mass, momentum, etc.) provide the links between the measured data and the forecasted quantity. This approach has never been used to generate operational very short-term wind energy forecasts. This is mostly because of the high computational cost. However, the steadily declining cost of computations is now making this option economically viable.

Finally, an area that could potentially benefit very short-term forecasts as well as the longer time scales of wind energy forecasting is an improvement to the plant output model. The improvements are likely to come from more abundant and higher quality meteorological and energy generation data from more sophisticated wind plant data gathering and communications systems and more detailed statistical or physical plant output model formulations.

It is likely that the improvements in forecast models and the data coverage and quality in the local area environment of wind plants will yield meaningful improvements in the performance of 0 to 6 hour forecasts over the next 10 years. However, it is difficult to provide a quantitative estimate because the documented history of these very short-term wind energy forecasts is brief and the current state of the art in performance for this time scale has not been as firmly established as that for the 1 to 2 day forecasts. A reasonable expectation is that there will be a 15% to 25% reduction in the typical MAE values for 0 to 6 hour forecasts over the next 10 years. This level of MAE reduction would result in an increase in the persistence-based skill score from about 20% at the present time to the 30% to 40% range in the year 2013.

8. Summary

The current state-of-the-art of forecasting techniques exhibits considerable skill in both very short-term ("next hour") and short-term ("day ahead") forecasting. Very short-term (0-6 hrs) hourly forecasts typically outperform a persistence forecast by 10% to 30%. Short-term (1-2 days) hourly forecasts usually outperform persistence and climatology forecasts by 30% to 50% At present, medium range (3-10 days) forecasts of the hourly wind energy production typically do not outperform climatology forecasts and hence have limited usefulness. However, medium range forecasts of the *average* energy production over a day or half-day usually do outperform a climatological forecast out to 6 or 7 days and hence provide some value to the user who can effectively employ that type of information. Figure 8 provides an estimate of the typical range of mean absolute forecast errors (expressed as a percentage of the plant's installed capacity) as a function of the forecast time horizon (look-ahead period) for the 1 to 12 hour forecast period. The MAE of very short-term forecasts is typically in the range of 5% to 15% and the errors increase rapidly (about 1.5% of installed capacity per hour) with an increase in the forecast time horizon. After the short-term period, the error growth rate decreases to about 0.1% of installed capacity per forecast hour. This means that the mean absolute forecast errors remain in the 13% to 21% range for 1 to 2 days ahead and rise to the 20% to 25% range that is typical of a climatological forecast after about 3 days.

It should be noted that forecast performance can vary substantially (5% or more of installed capacity) as a function of location, season and weather regime. Much of this is related to the predictability of specific weather regimes. Some weather regimes are inherently more sensitive to small variations in the conditions at the time the forecast is made. This means that for some weather regimes slight differences in the current conditions can give rise to large differences in the future conditions. This makes forecasting more difficult since small differences in the initial state can produce large errors in the forecast. Forecast performance in these types of regimes is normally much worse than for regimes with less sensitivity.

Finally, there is an expectation that an improvement in the quality and quantity of global, regional and local area atmospheric data, the development and application of more sophisticated statistical and physics-based atmospheric models and data assimilation schemes for those models and the availability of greater and lower cost computing power will yield substantial improvement in forecast performance over the next 10 years. Although there is likely to be some improvement across all forecast time horizons, the most significant improvements are likely to be made for the start (3-5 days) of the medium range forecasting period and the start of the short-term forecast period (6-18 hours).

The forecasts for the 3 to 5 day period will most likely benefit from the upcoming increases in global data coverage from remote sensing systems and improved global scale physics-based models and data assimilation procedures. The trend in forecast performance over the last 20 to 30 years suggests that a 72-hour forecast, in the year 2015, will be as good as a 36-hour forecast now. This means that 72-hour forecasts should be able to record an MAE of about 13-18% of installed capacity with a skill score of about 30% over a climatology forecast. In addition, the typical level of skill for a 72-hour forecast now should be achieved for a 108-hour (4.5 day) forecast in 2015. This translates into an hourly forecast MAE of about 20 to 25% of installed capacity and a skill score of a few percent over climatology for a 4.5-day forecast.

The forecasts for the 6-18 hour period will likely benefit from increased data availability from the local-area environment of the wind plant due to advances in remote sensing technology and the availability of increased low-cost computational power which will permit very high resolution simulations by local-area physics-based numerical models to be executed in real-time. It is reasonable to expect that these new capabilities will yield a 25% to 35% reduction in the forecast error and a 50% or greater improvement in forecast skill relative to persistence for the 6 to 18 hour time range.

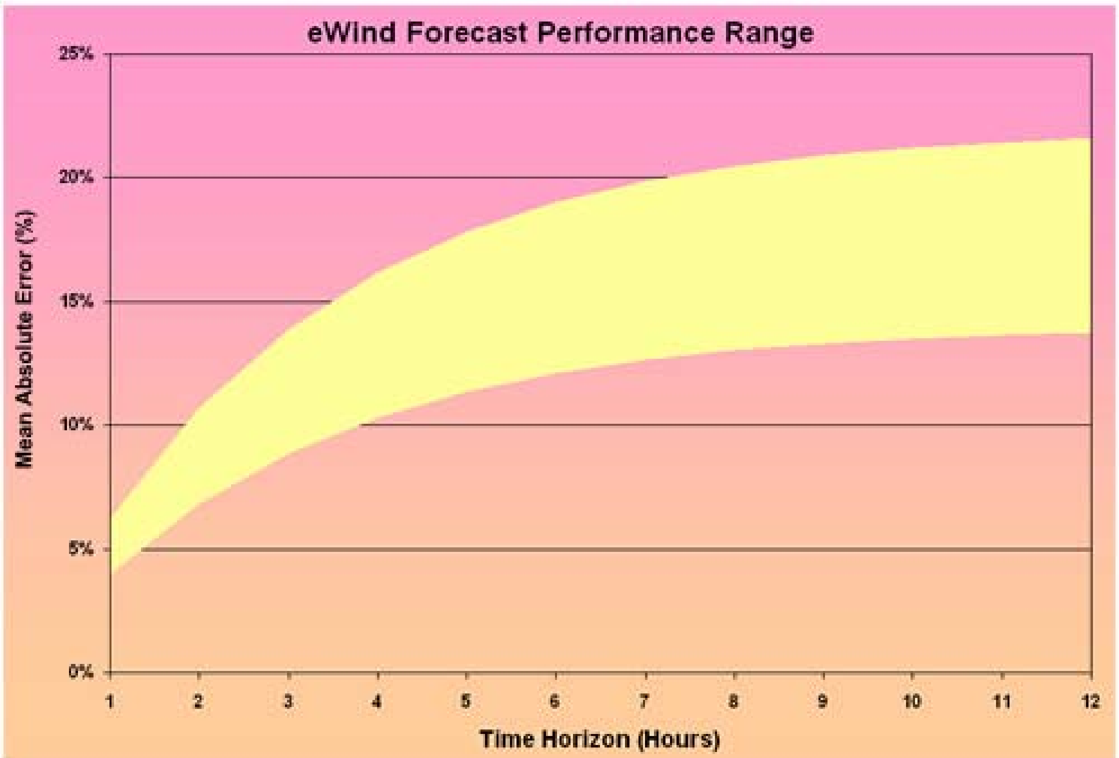

Figure 8. *Typical range of current wind energy forecast performance as a function of forecast time horizon. Forecast performance is expressed as a mean absolute error as a percentage of a wind plant's installed capacity*.

References

[1] 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, 1007339.

[2] Kalnay, E., S.J. Lord and R.D. McPherson, 1998: Maturity of Operational Numerical Weather Prediction: Medium Range, Bulletin of the American Meteorological Society, 79, 2753-2769.

[3] Westrick, K., 2003: Description and results from a comprehensive regional wind energy forecast system in the Pacific Northwestern U.S. *Proceedings of Windpower 2003 Conf.*, Austin, TX, American Wind Energy Assoc.