Very Short-Term Wind Forecasting for Tasmanian Power Generation

Article *in* IEEE Transactions on Power Systems · June 2006

DOI: 10.1109/TPWRS.2006.873421 · Source: IEEE Xplore

Michael Negnevitsky
University of Tasmania

Very Short-Term Wind Forecasting for Tasmanian Power Generation

Cameron W. Potter, *Student Member, IEEE,* and Michael Negnevitsky, *Member, IEEE*

***Abstract*—This paper describes very short-term wind prediction for power generation, utilizing a case study from Tasmania, Australia. Windpower presently is the fastest growing power generation sector in the world. However, windpower is intermittent. To be able to trade efficiently, make the best use of transmission line capability, and address concerns with system frequency in a re-regulated system, accurate very short-term forecasts are essential. The research introduces a novel approach—the application of an *adaptive neuro-fuzzy inference system* to forecasting a wind time series. Over the very short-term forecast interval, both windspeed and wind direction are important parameters. To be able to be gain the most from a forecast on this time scale, the turbines must be directed toward on oncoming wind. For this reason, this paper forecasts wind vectors, rather than windspeed or power output.**

***Index Terms*—Adaptive neuro-fuzzy inference systems (ANFIS), intelligent systems, very short-term forecasting, windpower.**

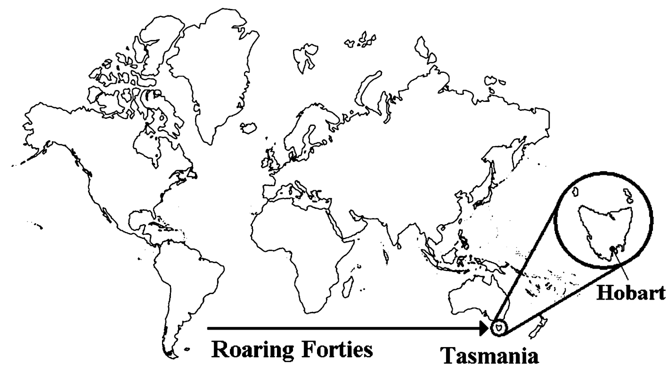

Fig. 1. Location of Tasmania on a world map.

I. Introduction

Tasmania is in a unique situation in the field of power generation; 98.5% of all power production is through renewable means. Specifically, the monopoly generator, Hydro Tasmania, utilizes hydro and wind, with a gas turbine power station as a back-up source. At present, Tasmania's hydro power dominates the power production with 54 dams, 27 hydro power stations, and numerous canals and waterways in seven catchment areas around Tasmania. This system generates almost 10 000 GWh of energy annually.

However, the growth area is in windpower. At the start of 2005, Tasmania had five operational windfarms. These included two on Flinders Island, one on King Island, and also Stages 1 and 2 of the Woolnorth Windfarm development. These farms total to a capacity of 67 MW, or a potential wind penetration of 5.7% (although the installations have not been operating long enough for accurate yearly assessment). Developments that are underway will take the total to 143 MW. Furthermore, there are proposals being investigated for a further 674 MW [1], [2]. This means there is a potential for almost 50% wind penetration.

A smaller independent system (that is also operated by Hydro Tasmania) on King Island is another interesting problem. This power system has no connections to a larger grid and has a maximum demand of approximately 3 MW. On this island, there is 2.5 MW of wind generation installed, resulting in a very high wind penetration.

The geographic location of Tasmania puts the west coast inline with the renowned "Roaring Forties" wind pattern that builds up over the Atlantic Ocean and the Indian Ocean, producing prevailing westerlies—see Fig. 1. The west coast of Tasmania is also sparsely populated, allowing for more cost-effective, onshore wind developments.

This abundance of wind makes the installation of large amounts of windpower viable. One of the major problems with using wind for power generation is that it is not able to be stored. One of the major problems with hydropower is that dry years can jeopardize power production. The combination of wind and hydropower represents new advantages. When the wind is producing power, extra river flow can be stored, when the wind is not producing power that extra water can be released. In fact, easy access to hydro dams also allows for consideration of pumped storage, possibly resulting in more effective utilization of power resources.

New challenges, as well as new opportunities, are provided by the construction of an underwater cable that will connect the Tasmanian power grid to Australia's National Electricity Market (NEM). Increasing Tasmania's windpower resources will provide further flexibility in the trading. Furthermore, as consumers become increasingly environmentally conscious, how the power is generated becomes more important. A recent survey [3] across Australia showed that 76% of respondents were prepared to pay an additional 5% on their energy bills for a 10% increase in renewable energy.

Now that Tasmania will soon be connected to the NEM, excess energy will be able to be exported. In fact, due to the relative ease of start-up and shut-down operations for a hydropower system, Hydro Tasmania will be able to conserve power resources at low price periods and expend power resources at high price periods. The utilization of Australia's Mandatory Renewable Energy Target (MRET) system will also supplement the ability of renewable energy to compete with fossil-fuel produced power. This system indicates that all distributors must effectively source an increasing percentage (up to 2010) of their power from renewable sources. This is facilitated through the

Manuscript received June 20, 2005; revised October 11, 2005. Paper no. TPWRS-00365-2005.

The authors are with the School of Engineering, University of Tasmania, Tasmania, Australia (e-mail: cameron.potter@utas.edu.au; michael.negnevitsky@utas.edu.au).

Digital Object Identifier 10.1109/TPWRS.2006.873421

purchase of MRET credits, which are traded in a separate market [4].

The NEM is a wholesale power pool in Australia to which the generators sell the power they produce, and the distributors buy the power they need to supply their clients. This form of power system was initially known as a “deregulated system”; however, it is now more commonly known as a “re-regulated system,” as some restrictions still apply. The generators make a bid to supply a certain amount of power at a chosen price. The overseeing body, called the National Electricity Market Management Company (NEMMCO), receives the generators’ bids, ranks them according to price, and accepts enough bids to satisfy the projected demand plus a safety margin. Usually this means that the last bid accepted is accepted as a partial amount. The highest accepted bid price is the price that each supplier receives for the power that they produce in that time period, regardless of their own bid price. Most generators and distributors have a majority of their power resources set in contracts. These contracts are unaffected by the spot price market—from the customer perspective. However, the generators can opt to buy power from the spot price market to fill their contracts if they either cannot meet the contract themselves or it would be more economical to do so. Hydro Tasmania will soon be changing from its present state-monopoly to joining this market bringing an increased need competitiveness. Accurate and reliable forecasts are essential to competitive renewable energy supply.

Another driver for very short-term wind forecasting (especially in Australia) is that with the rapid growth of windfarming, the old scheduling methods are becoming inefficient [5]. At present, the NEM simply schedules all windfarm production straight into the market. Until recent years, there has been no need for a bidding system as the wind generation has been insignificant in comparison with total system requirements. However, as generation from wind sources increases, so too does the influence of these sources on the market. For this reason, NEMMCO is reviewing the scheduling policy; one possible outcome will result in the utilities to have to bid their wind assets into the market, just like fossil-fuel or hydro production.

Accurate forecasts of power generation are also of importance to electricity transmission. As windfarms grow in capacity, the strain they place on the transmission grids also becomes more pronounced. This is then further compounded by the fact that many windfarms are being built in remote areas with sparse transmission grids. This means that operation near transmission limits will not be unusual for wind generation, and sometimes the transmission grid may not be able to transmit all the windfarm generation. In Tasmania, the west coast has the greatest wind potential and also has the majority of the hydroelectric catchments. Using accurate short-term forecasts, surplus of wind generation could potentially be directed into pumped hydroelectric storages.

II. Present Forecasting Techniques

A. General Overview

Forecasting is a vital part of business planning in today’s competitive environment. However, while there is significant interest in research regarding prediction for wind power generation, there is no reliable system for the very short-term time-series prediction [6].

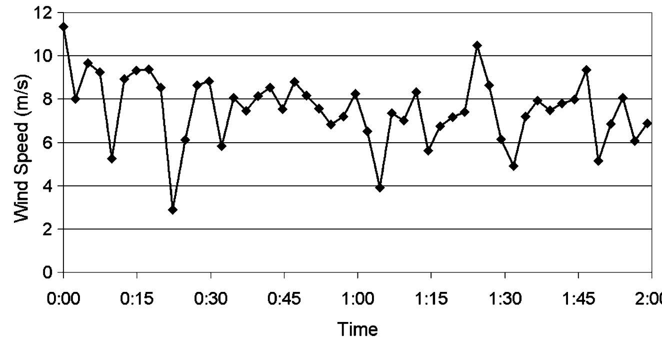

Fig. 2. Typical wind time-series plot for a wind site in Tasmania, Australia.

TABLE I
WIND VARIATION FOR A SITE IN TASMANIA, AUSTRALIA, 2002

Measured Variable	Mean	Standard Deviation
Wind Vector (m/s)	4.9	3.0
Change in Wind Vector (m/s) (over 2.5 minutes)	1.8	3.0

Windspeed represents a time series that can be defined as a set of observations of a parameter, or set of parameters, taken at a number of time intervals. These intervals are usually (although not always) of a regular length. If the time step between data points is not consistent, or there are data missing, this should be corrected to a regular time step if the data are to be used for forecasting. Real-world time series are diverse. Some time-series data change slowly and relatively smoothly. Monthly electricity demand may represent such a time series. Other time series can exhibit relatively chaotic behavior, making them difficult to predict. A windspeed time series, such as in Fig. 2, possesses these characteristics.

Wind prediction is complex due to the wind’s high degree of volatility and deviation. Time-length scales need definition. In this paper, *very short-term forecasting* is defined as look-ahead periods from a few minutes up to an hour, while *short-term forecasting* will indicate hours out to a few days ahead.

This difference between the two forecasting time periods is important when trying to create a prediction system. Three main classes of techniques have been identified for wind forecasting. These are numeric weather prediction (NWP) methods, statistical methods, and methods based upon artificial neural networks (ANNs).

The NWP methods are based on mathematical fluid mechanics models and have been found to dominate the meteorological literature, almost exclusively. NWP methods are the most accurate technique for short-term forecasting and beyond. However, statistical and ANN methods based on observations perform more accurately over the very short-term forecast range [7], [8].

One reason for NWP method’s failure on the very short-term time scale is that wind has a high variation over brief periods. As a result statistical, ANN and other similar methods perform better. An example of the variation in wind data is shown in Table I.

These values show that for the wind vector, the standard deviation is high when compared to the mean value, but more significantly, the standard deviation of the *change* in windspeed vector is even higher compared to its mean. This level of variation in a data set causes problems when trying to achieve accurate predictions.

This paper proposes a novel approach to very short-term wind prediction. However, first we will investigate what is presently available and *also* what is currently considered as industry present practice (as these two are not necessarily the same).

B. Weather Prediction Scales

Until recently, there were two main wind forecasting spheres of influence: planetary and synoptic. The planetary scale is for extremely large weather patterns. The time scale for this tends to be measured in months or even years. The synoptic scale is smaller but still focuses on large weather patterns that may influence multiple countries. The time scale for synoptic prediction tends to be in the order of days or months. This was acceptable for coarse weather prediction of patterns such as fronts, highs, and lows. However, as more precise forecasts were sought after, a new scale for forecasting was developed. This scale is mesoscale forecasting.

However, the mesoscale forecasts were also too coarse for small area and very short-term prediction. Consequently, other time periods have been defined for use in meteorological studies:

1) mesoscale alpha (2 days—6 h);
2) mesoscale beta (6 h—30 min);
3) mesoscale gamma (30 min or less).

C. Numerical Weather Prediction Models

NWP models were developed by meteorologists and are well researched. The focus of these models is on the synoptic or planetary scales. These models are very successful in this area. With the need for shorter term forecasts, NWP methods have been supplemented to increase the accuracy over the smaller time steps.

Short-term NWP models operate by solving conservation equations (mass, momentum, heat, water, etc.) numerically at given locations on a spatial grid that is three dimensional: latitude, longitude, and elevation. This is considered across the fourth dimension of time as well. The finer the spatial grid, the finer the ability to forecast. NWP models of synoptic scale (or greater) must be assumed hydrostatic (equilibrium of the vertical pressure gradient in the atmosphere) and therefore cannot model thermally driven winds such as sea breeze. Mesoscale models, however, can be nonhydrostatic and can predict smaller scale wind patterns such as land/sea breeze, venturi effect (mountain winds), etc. In fact, the shorter term predictions can be dominated by terrain and nonhydrostatic phenomena. Thus, mesoscale NWP models need to include accurate digital elevation models (DEMs) to represent the topography over which the weather prediction occurs.

Predictions from a NWP model have four limiting factors: data variability; DEM resolution; grid spacing; and computation time. A NWP would also suffer from inaccurate input data, as is true for any forecast model.

The variability of the data cannot be avoided and will always limit the accuracy of a NWP model. As forecasts become more precise, so too must the DEM. This becomes a limiting factor once the resolution gets small enough that forests and towns become significant factors as these features can change rapidly (causing need for remodeling the environment).

However, even if these two obstacles could be overcome, the spacing of the monitoring stations can be an issue. For any numerical model, five grids points are required to resolve a wave's structure. Thus, for a 5-km (tight) grid separation, the smallest weather pattern that can be accurately modeled would be 20 km across. This is insufficient for very short-term wind forecasting.

The last limiting factor is that NWP models are heavily mathematical and are usually run on super computers, even then taking over an hour to obtain a result. This by itself may limit the usefulness of NWP methods for online applications in power systems.

D. Persistence Models

The persistence technique is based upon the high correlation between the present windspeed and the windspeed in the immediate future. This method was developed by meteorologists as a comparison tool to supplement a NWP. Since the accuracy of very short-term prediction was historically deemed unimportant, persistence was sufficient. In fact, this simplified method proved to be more effective than a NWP model over very short-term prediction [9].

E. Statistical and Neural Network Methods

The statistical and neural network-based methods are mostly aimed at very short-term predictions [6], [7], [10], [11]. The statistical methods are auto-recursive. This means they use the difference between the predicted and actual windspeeds in the immediate past to tune the model parameters. The neural networks use past data taken over a longer time-frame to learn the relationship between the input data and output windspeeds. The accuracy of these methods degrade rapidly with increasing prediction lead time.

Most techniques are either based on past and present measurements or measurements provided by weather observation stations, principally upstream from the prevailing weather movement direction [7], [10] or at the wind site itself [11]. A significant difference between statistical models and neural networks versus NWP methods is that the former use the present windspeed at the measurement site(s) as their inputs. This is one of the reasons for the comparably accurate forecast speeds over the very short-term forecast time scale.

However, as yet, no very short-term wind forecasting system has gained widespread industry acceptance [6], and often the persistence technique is used instead. Consultation with industry experts indicated that the available statistical techniques were impractical to implement. Often they are not portable, needing an expert to create an individualized model for each windfarm. There is also a high reliance on upstream observation stations, which are not always available or practical to install. Due to their lack of portability, changes in conditions very often need changes in the model. A change in the model requires the

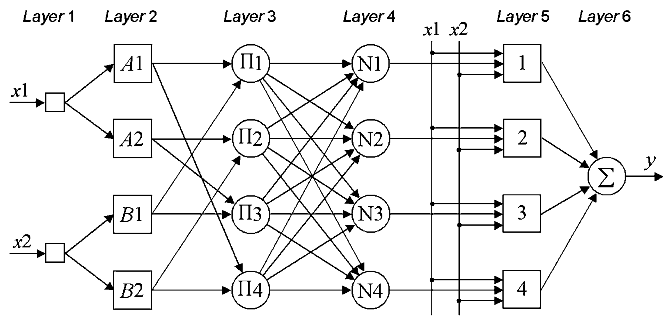

Fig. 3. ANFIS model representation.

Rule 1:
IF $x1$ is $A1$
AND $x2$ is $B1$
THEN $y = f_1 = k_{10} + k_{11}\,x1 + k_{12}\,x2$

Rule 2:
IF $x1$ is $A2$
AND $x2$ is $B2$
THEN $y = f_2 = k_{20} + k_{21}\,x1 + k_{22}\,x2$

Rule 3:
IF $x1$ is $A2$
AND $x2$ is $B1$
THEN $y = f_3 = k_{30} + k_{31}\,x1 + k_{32}\,x2$

Rule 4:
IF $x1$ is $A1$
AND $x2$ is $B2$
THEN $y = f_4 = k_{40} + k_{41}\,x1 + k_{42}\,x2$

Fig. 4. Example of ANFIS rules.

attention of an expert again for the retuning of the statistical models.

Prediction research is a growth area, and increasingly often, this research involves the use of artificial intelligence [12], [13]. In this paper, we investigate a hybrid approach—a combination of an ANN and fuzzy logic for very short-term wind prediction.

III. Adaptive Neuro-Fuzzy Inference System

Fuzzy systems and neural networks are natural complementary tools in building intelligent systems. While neural networks are low-level computational structures that perform well when dealing with raw data, fuzzy logic deals with reasoning on a higher level. However, fuzzy systems lack the ability to learn and cannot adjust themselves. The merger of a neural network with a fuzzy system into one integrated system therefore offers a promising approach to building very short-term wind prediction models.

A neuro-fuzzy system is, in fact, a neural network that is functionally equivalent to a fuzzy inference model. For example an *adaptive neuro-fuzzy inference system* (ANFIS) proposed by Roger Jang [14] is a six-layer feedforward neural network. Fig. 3 shows the ANFIS architecture. For simplicity, we assume that the ANFIS has two inputs—x1 and x2—and one output—y. For additional simplicity, each input is represented by only two fuzzy sets, although three or more are not uncommon. Extra membership functions will increase the accuracy of the results but will take longer to train. For this example, the two fuzzy sets are converted to the output by a first-order polynomial. The ANFIS in Fig. 3 implements four rules, shown in Fig. 4, where x1, x2 are input variables; A1 and A2 are fuzzy sets on the universe of discourse X1; B1 and B2 are fuzzy sets on the universe of discourse X2; and $\{k_{i0}, k_{i1}, k_{i2}\}$ is a set of parameters specified for rule i.

Layer 1 is the input layer. Neurons in this layer simply pass external crisp signals to Layer 2.

Layer 2 is the fuzzification layer. Neurons in this layer perform fuzzification. In Jang's model, fuzzification neurons normally use a bell activation function.

Layer 3 is the rule layer. Each neuron in this layer corresponds to a single fuzzy rule. A rule neuron receives inputs from the respective fuzzification neurons and calculates the firing strength of the rule it represents.

Layer 4 is the normalization layer. Each neuron in this layer receives inputs from all neurons in the rule layer and calculates the normalized firing strength of a given rule—the ratio of the firing strength of a given rule to the sum of firing strengths of all rules. It represents the contribution of a given rule to the final result.

Layer 5 is the defuzzification layer. Each neuron in this layer is connected to the respective normalization neuron and also receives the initial inputs, x1 and x2. A defuzzification neuron calculates the weighted consequent value of a given rule.

Layer 6 is represented by a single summation neuron. This neuron calculates the sum of outputs of all defuzzification neurons and produces the overall ANFIS output, y.

An ANFIS uses a hybrid learning algorithm that combines the least-squares estimator and the gradient descent method [14]. First, initial activation functions are assigned to each membership neuron. The function centers of the neurons connected to input x_i are set so that the domain of x_i is divided equally, and the widths and slopes are set to allow sufficient overlapping of the respective functions. In the ANFIS training algorithm, each training epoch is composed from a forward pass and a backward pass. In the forward pass, a training set of input patterns (an input vector) is presented to the ANFIS, neuron outputs are calculated on the layer-by-layer basis, and rule consequent parameters are identified.

IV. Case Study

A. Defining the Problem

Currently, there is no very short-term wind prediction package commercially available [6]. There is also little available literature for this time scale for power generation purposes. Some research focuses on a time scale of half an hour to two hours ahead [10]. There has been attempts using ANNs to forecast intervals as low as 10 min ahead [7]. However, this research does not consider one of the major intervals for wind power management—the need for 2-3-min predictions for governing power output due to gusts [15]. This paper addresses this time scale and provides 2.5-min-ahead forecasts of wind vectors.

B. Developing the Model

The ANFIS model design is flexible and capable of handling rapidly fluctuating data patterns. This meant that it covered the criterion necessary for very short-term wind prediction. This model has two major goals. The first one is to increase prediction accuracy (and at least outperform the simple-to-implement industry standard of persistence). The second one, which is no less important, is to make a system that can be easily installed at a variety of different sites.

To use the ANFIS model, the user first must define a number of variables:

TABLE II
COMPARISON OF SINGLE EPOCH TRAINING TIMES FOR TASMANIAN ELECTRICITY DEMAND DATA

Input and MF Setup	Training Time [s]	RMS Error
6 inputs and 3 MFs	54436	0.0296
6 inputs and 2 MFs	264	0.0296
4 inputs and 3 MFs	220	0.0345
4 inputs and 2 MFs	8	0.0362

1) size of the training set;
2) number of training epochs (iterations);
3) type of the fuzzy membership functions (MFs);
4) number of MFs associated with each input.

The size and diversity of the training set is important. If the set has insufficient variation to properly model the characteristics of the data, the training of the ANFIS will fail.

The standard MF for an ANFIS is a "bell function," such as the Gaussian distribution function [14].

The number of MFs depends upon the complexity of the system and the size of the training set. As a general rule, more MFs will better represent a complex system but will take longer train—especially for a large training set. In this case study, two MFs were used for each input.

Table II demonstrates the relationship between training times and the number of membership functions. It should be noted that this test was done on Tasmanian electricity demand data as this is far easier to predict and so smaller training sets could be used.

Table II also includes a variable number of inputs: six or four. More inputs to the ANFIS model will result in a wider variety of data in each pass. This will result in more selective training, producing better predictions. However, as with MF's, a greater number of inputs will have the consequence of a longer training time.

C. Choosing the Inputs

In order to predict a time series with an ANFIS model, it is important to select the right inputs for the system. Unfortunately, there are no rules available to expedite this process. Every data set has different deviations and rates of change. Thus, what might work for one data set will not necessarily be the best configuration for another.

If the system can be trained over a long period of time, the developer can afford to use larger data sets, more inputs, and more membership functions; however, the developer must also consider the need for multiple training and testing runs. A system will need to be trained at least several times to show that it produces reliable results and many times before that, in order to tune the parameters. Once the training set size is chosen, the number of membership functions and the number of inputs are chosen to allow for training time and the available computing resources.

The most common method of choosing inputs for very short-term time-series prediction is by considering a subset of the available data. Often these are known points from the same time series, chosen to best determine the desired prediction point. Equation (1) is a generalized example input to try to predict the point at $\mathrm{x(t+1s)}$. X is the input pattern, $x(t)$ is the value of the time series at the present time t and s is the chosen time step

$$X = [x(t-3s) \quad x(t-2s) \quad x(t-1s) \quad x(t)]. \qquad (1)$$

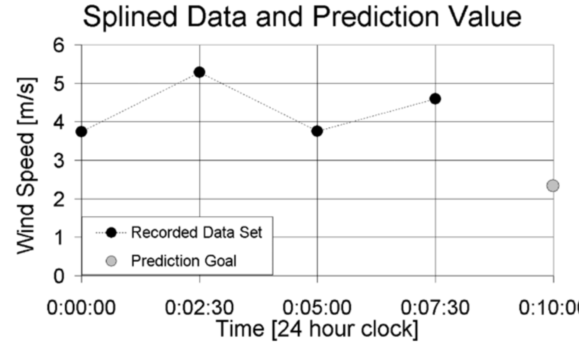

Fig. 5. Example data set, trying to predict the point at 0:10:00.

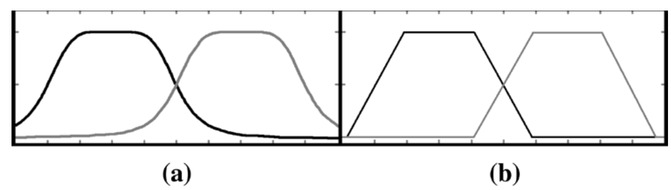

Fig. 6. Example MFs: (a) "Bell" MFs. (b) Trapezoidal MFs.

Fig. 5 provides a sample of wind data. The goal is to predict the point at 0:10:00. Using an input time step of 2.5 min, the equation would become

$$X = [x(0:00:00) \quad x(0:02:30) \quad x(0:05:00) \quad x(0:07:30)]. \qquad (2)$$

The example uses four inputs and a single output. The number of data points in the input pattern can be adjusted. In Section IV-F, the authors used six inputs to the ANFIS model and one output, the predicted value. The spacing of the input data was changed through the use of interpolation splines, but the forecast was always for a period of 2.5 min.

D. Choosing the Membership Functions

Initially, the ANFIS model was trained with a bell function, shown in Fig. 6(a). This is an example of a soft-edged MF as the transitions have no abrupt corners. Note that the boundaries of these curves are extremely wide. When tested on less variable data, such as electricity demand, this model gave accurate results (less than 1% mean absolute error over a half hour interval). However, the results were not satisfactory for highly stochastic (or "spikey") test data. Large events in the time series would result in large error spikes, such as in Fig. 7. In order to improve robustness, the MF type was altered. It was found that hard-edged MFs were more robust. A trapezoidal MF, shown in Fig. 6(b), has clearly defined boundaries, and removes the large overshoot (see Fig. 7). However, the robustness did come at the cost of some fine-tuning accuracy. Examining Fig. 7 shows that other than the large error spikes, the results using the bell MF were more accurate.

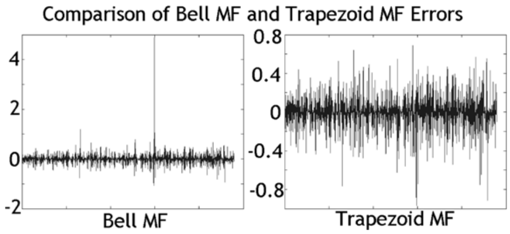

Fig. 7. Twenty-four hours of results obtained using different MFs on the same data set.

TABLE III
COMPARISON OF WIND TURBINE TOWERS FROM THE SAME WINDFARM

Wind Speed [m/s]		
Mean for Tower A	Mean for Tower B	Mean Absolute Difference
0.80	0.77	0.19

The final design resulted in using a bell MF and overcoming the overshoot/undershoot issue by putting a hard-limiter in the system to truncate predictions outside a defined range—selected by simple maximum/minimum statistical data analysis. The use of data analysis was chosen instead of expert input so that the procedure could be easily replicated at other wind sites without the need for external input.

E. Predicting Difference to Aid Transportability

Wind forecasting models are site specific. Each recording site produces unique data. A comparison of recordings from the same windfarm (over very short-term) showed that even sites in close proximity produce significantly different data. To test this, two adjacent wind turbine towers on the same windfarm were compared. The results are shown in Table III.

Although *quantitative* proof of time-series (data-based) prediction is difficult, and no system can be *guaranteed* as being portable to another site, there are steps that can be taken to improve portability. It was determined that while different sites may have varying wind patterns, most sites will behave similarly when considered on a very short-term basis. To further increase the portability, this idea of similar very short-term behavior is able to be exaggerated through using the difference between data points as the training data and as the output. The forecast difference will then be added to the present data point to determine the effective offset. A mathematical proof showing that difference data has a smaller range than the sampled data is included in the Appendix.

For the wind site used in this case study, some histograms were developed. Fig. 8 shows a wind vector histogram for Spring. It shows that westerlies are the predominant wind pattern but also has a large reading at zero. Fig. 9 shows the differenced wind vector's histogram. It resembles a Laplacian distribution, which shows a marked likelihood for little or no alteration. This resulted in much better predictions. Using histograms such as these, it is possible to create a probabilistic model to develop additional data should it be required for training at a later stage.

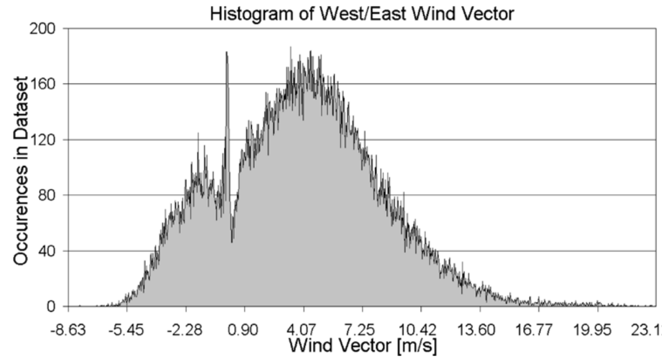

Fig. 8. Histogram of Spring wind data from a site in Tasmania, Australia.

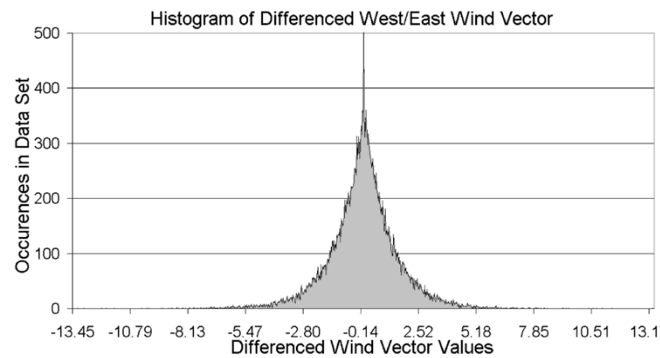

Fig. 9. Histogram of the differenced values of Spring wind data from a site in Tasmania, Australia.

F. Case Study Results

The case study used a wind site in Tasmania as the data set, providing a 21-month time series in steps of 2.5 min. This was to be the forecast period as well. After consultation with industry experts, the data was converted from speed and direction into $u - v$ vectors. For this case study only one vector was considered at one tower height. This allows the training of more systems (within finite computing potential) resulting in a better comparison of the various techniques.

In order to adequately test the performance of the proposed ANFIS system, multiple architectures were evaluated on the same data. A persistence model was also developed for comparison. Persistence is presently an industry benchmark for very short-term wind forecasting and so is the most indicative assessment.

The ANFIS model was developed in several different formats. This was to highlight the usefulness of intermediary splines through data for very short-term forecasting. The results are shown in Fig. 10. The chart also includes the results from the persistence model. A useful comparison is available through considering the persistence results and the ANFIS model with no spline. The ANFIS model shows some improvement, in the order of 5%.

The tests using splined data as the input to the ANFIS model resulted in an improvement in prediction. The system architecture used is identical to the system used without a spline except that the data used as inputs (and outputs) were interpolated using a spline of 2, 5, and 10 points. The differences between these results are small, but the difference between these three

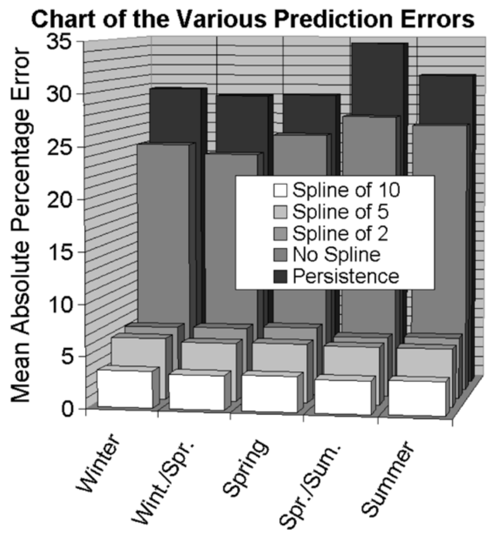

Fig. 10. Chart of prediction errors for various systems, tested over a period of eight months on a wind site in Tasmania, Australia.

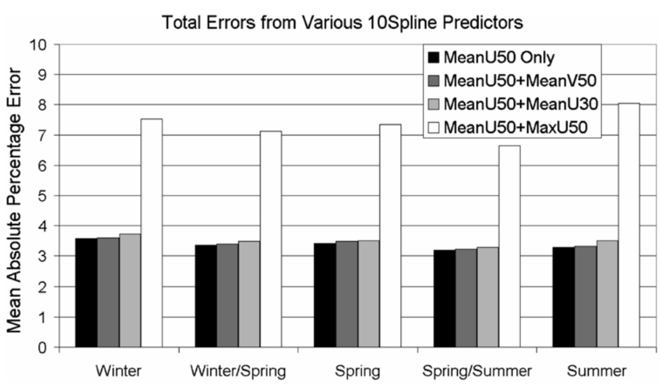

Fig. 11. Chart of errors comparing various instances of ANFIS models using data from a wind site in Tasmania, Australia.

results and the results using no spline or a persistence model is significant. It is thus concluded that the results were improved by using some degree of spline. It is important to demonstrate that the persistence model would not gain from such a method. Persistence by nature does not change between known points, so each intermediate prediction would be identical in value.

Further tests were developed to study the effectiveness of the model on a variety of training data. The results of these additional trials are summarized by Fig. 11. The goal was to check whether additional data from a nearby data source could be used to improve prediction. In three out of the four cases, the change was negligible. The exceptional case was a consequence of the addition of data (from the same site) but recording the maximum value over the interval instead of the mean value. This proved to increase the mean absolute percentage error.

V. Conclusion

This paper demonstrates the development of a wind prediction system for forecasts of wind vectors 2.5 min ahead. In the case study on a site in Tasmania, Australia, the implementation of an ANFIS model produced results with less than 4% mean absolute percentage error. To provide comparison, a persistence model was developed using the same data. This system produced a mean absolute percentage error of approximately 30%. The ANFIS model was developed such that additional expert input is not required (on the condition that the data provided are reliable and have an even time step).

Appendix

This Appendix provides two mathematical proofs that address the use of difference data used as forecasting inputs. The first proof shows that the range of difference data will never be greater than the data from which it was derived. The second proof shows that for faster sampling, the range of difference data must be smaller than the range of the sampled data. Together these proofs show that there is no risk of increasing data range using difference data rather than the sampled data, and for an increased sample rate, the range will in fact be reduced.

A smaller range results in fewer samples required to represent that time series. The need for fewer samples may allow for training set reduction, thus increasing the rate of training. However, more importantly, the ability to use a smaller training set makes better use of limited resources. If the available data set is smaller than desired, the system will be better trained through training with differenced data.

A. Proof 1: Initial Difference Data and Sample Data

We show here that the range of the difference between two adjacent data points can never exceed the range of the original data set.

Show that $|\Delta y_j| = |y_{j+1} - y_j| \leq (y_{\max} - y_{\min})$ by contradiction.

Let $y_{j+1} \geq y_j$.

Now, suppose that

$$\begin{aligned}(y_{j+1} - y_j) &> (y_{\max} - y_{\min})\\ \therefore\ y_{j+1} &> y_{\max} - y_{\min} + y_j.\end{aligned}$$

By definition

$$\begin{aligned}y_j &\geq y_{\min}\\ \Rightarrow y_{j+1} &> y_{\max} - y_{\min} + y_{\min}\\ \therefore\ y_{j+1} &> y_{\max}\end{aligned}$$

which is a contradiction; thus, $(y_{j+1} - y_j) \leq (y_{\max} - y_{\min})$.

Similarly, it can be shown that if $y_{j+1} \leq y_j$, then

$$-(y_{j+1} - y_j) \leq (y_{\max} - y_{\min}).$$

Thus, $|\Delta y_j| = |y_{j+1} - y_j| \leq (y_{\max} - y_{\min})$.

B. Proof 2: Increased Sampling Rate

We show here that increased sampling provides a result such that the range of the difference will be strictly less than the range of the sample.

Suppose that the observations y_j can be regarded as values of some function $f(t)$, so that $y_j = f(t_j)$. Then

$$|\Delta y_j| = |y_{j+1} - y_j| = |f(t_{j+1}) - f(t_j)| \text{ and } \Delta t_j = t_{j+1} - t_j.$$

Suppose now that the function $f(t)$ is Lipschitz continuous [16], so that the condition

$$|f(t_{j+1}) - f(t_j)| \leq L|t_{j+1} - t_j|$$

is obeyed. Here, L is a positive constant. Notice that this is a weaker form of continuity, which does not necessarily require f to be differentiable; if f does indeed have a derivative, then the constant L can be taken to be the maximum value of $|f'(t)|$ over the interval, by the Mean-Value Theorem [17]. It follows that

$$|\Delta y_j| \leq L\Delta t_j \leq L\Delta t_{\max}.$$

The right-hand side of this inequality is independent of j

$$\therefore |\Delta y|_{\max} \leq L\Delta t_{\max}.$$

As the number of sampling points increases, $\Delta t_{\max}$ decreases, and consequently, the range of the difference $|\Delta y|_{\max}$ also decreases. It follows that when the number of sampling points is sufficiently large, the range of the difference can be made smaller than the range of the sample.

Acknowledgment

The authors would like to thank Hydro Tasmania for their collaboration and provision of time-series data for the authors' research on power system prediction models. Specific thanks go to K. Jacka, R. Allen, M. Skoklevski, and P. Clark for their personal assistance in the research. The authors also would like to especially thank Prof. L. Forbes, School of Maths and Physics, University of Tasmania, for his aid in refining the proofs in the Appendix.

References

[1] Hydro Tasmania—hydro.com.au, in 2005: Hydro Tasmania.
[2] G. Chua, "The sleeping dragon," *ReFocus*, pp. 26–33, Nov./Dec. 2003.
[3] "Australian Reasearch Group 2003—National Renewable Energy: Quantitative Research, Report to the Wind Energy Association,", 2003.
[4] G. E. J. Tambling *et al.*. (2004) Mandatory Renewable Energy Target Review—A Review of the Operation of the Renewable Energy (Electricity) Act 2000. Ministry of Environment and Heritage. [Online]Available: http://www.mretreview.gov.au/.
[5] L. Owens, South Australia wind study 2005, in 2005: Electricity Industry Supply Planning Council.
[6] G. N. Bathurst, J. Weatherill, and G. Strbac, "Trading wind generation in short term energy markets," *IEEE Trans. Power Syst.*, vol. 17, no. 3, pp. 782–789, Aug. 2002.
[7] S. Li *et al.*, "Using neural networks to estimate wind turbine power generation," *IEEE Trans. Energy Convers.*, vol. 16, no. 3, pp. 276–282, Sep. 2001.
[8] M. C. Alexiadis, P. S. Dokopoulos, and H. C. Sahsamanoglou, "Wind speed and power forecasting based on spatial correlation models," *IEEE Trans. Energy Convers.*, vol. 14, no. 3, pp. 836–842, Sep. 1999.
[9] D. Milborrow, "Forecasting for scheduled delivery," *Windpower Monthly*, p. 37, Dec. 2003.
[10] I. G. Damousis *et al.*, "A fuzzy model for wind speed prediction and power generation in wind parks using spatial correlation," *IEEE Trans. Energy Convers.*, vol. 19, no. 2, pp. 352–361, Jun. 2004.
[11] T. S. Nielsen, H. Madsen, and J. Tofting, "WPPT, a tool for on-line wind power prediction," in *Proc. Int. Energy Agency Expert Meeting Wind Forecasting Techniques*, Boulder, CO, 2000.
[12] M. Negnevitsky, *Artificial Intelligence: A Guide to Intelligent Systems*, 2nd ed. Reading, MA: Addison-Wesley, 2005.
[13] C. Potter and M. Negnevitsky, "ANFIS application to Competition on Artificial Time Series (CATS)," *FUZZ-IEEE*, 2004.
[14] J.-S. R. Jang, "ANFIS: adaptive network-based fuzzy inference systems," *IEEE Trans. Syst., Man, Cybern.*, vol. 23, no. 3, pp. 665–685, May-Jun. 1993.
[15] J. P. Barton and D. G. Infield, "Energy storage and its use with intermittent renewable energy," *IEEE Trans. Energy Convers.*, vol. 19, no. 2, pp. 441–448, Jun. 2004.
[16] A. B. Dickinson, *Differential Equations—Theory and Use in Time and Motion*. Reading, MA: Addison-Wesley, 1972.
[17] J. Stewart, *Calculus*, 3rd ed. Pacific Grove, CA: Brooks/Cole, 1995.

Cameron W. Potter (S'04) received the B.E. (Hons) degree from the University of Tasmania, Tasmania, Australia, in 2002. He is currently working toward the Ph.D. degree.

His major research interests are forecasting techniques, intelligent systems, and power system operations.

Mr. Potter is a Student Member of the Institution of Engineers Australia.

Michael Negnevitsky (M'95) received the B.S.E.E. (Hons) and Ph.D. degrees from Byelorussian University of Technology, Minsk, Belarus, in 1978 and 1983, respectively.

Currently, he is a Professor in the School of Engineering, University of Tasmania. From 1984 to 1991, he was a Senior Research Fellow and Senior Lecturer in the Department of Electrical Engineering, Byelorussian University of Technology. After his arrival in Australia, he was with Monash University, Monash, Australia. His major interests are power system analysis, power quality, and intelligent systems applications in power systems.

Dr. Negnevitsky is a Chartered Professional Engineer, a Senior Member of the Institution of Engineers Australia, and a Member of CIGRE AP36 (Electromagnetic Compatibility), the Australian Technical Committee.