Long Term Yield Predictions Using a Fuzzy Logic Method

Carlos Ferreira, Joseph O'Hagan, and Elizabeth Traiger

AWEA WINDPOWER 2010
May 23 - 26, 2010

Corresponding Author: carlos.ferreira@genivar.com
GENIVAR WIND ENERGY, 103 - 2710 – 3rd Ave NE, Calgary, Alberta, Canada T2A 2L5

Long Term Yield Predictions Using a Fuzzy Logic Method

Carlos Ferreira, Joseph O'Hagan, and Elizabeth Traiger

Abstract - This work proposes a method using data from multiple reference stations as inputs to a fuzzy logic model to improve accuracy in predicting the wind speed distribution and yields at the site of interest. Wind yield predictions are investigated at a site surrounded by three Environment Canada reference stations with varying ranges. Yields at 60m computed using a generic power curve were calculated for the predicted and measured wind speeds at the target site. Overall yields were predicted within a maximum error of ~1%. The specifics of the proposed method involve the use of a fuzzy logic model where the inputs for the yield prediction model are selected using a sequential forward search (SFS). Possible input variables include wind speeds, wind directions, month, and time of day for each of the three reference stations, for a total of twelve possible inputs. These candidates are narrowed to four or five total inputs using a SFS. The resulting fuzzy logic model output is an average; however, the desired output is a probability distribution. Simulating data from a specified distribution that is applied to the output achieves a suitable pseudo probability distribution. Results from our simulations show an increased correlation between the wind speeds from our model as compared to individual site correlations. In addition, these simulations suggest that the model can capture some of the seasonal, diurnal, and terrain factors. The three most important learning objectives of this paper can be summarized as follows: 1. Correlation with a single reference site can be insufficient to capture the complex variations between the wind regimes 2. Combining data from multiple reference stations leads to accurate predictions. 3. Fuzzy logic methods can be used as a valid MCP (measure, predict, correlate) process for predicting yields. This method can be used to predict long term wind yields by using meaningful data from reference stations that are relatively far away from the desired site.

Long Term Yield Predictions Using a Fuzzy Logic Method

Introduction

With the current attractiveness of wind potential as a reliable energy source, a reliable method for wind prediction is increasingly needed to assess the availability of wind energy. Current practice commonly involves the use of correlated data from a single nearby reference site to predict long terms yields at a site of interest. As yields are influenced by meteorological parameters and terrain conditions, the accuracy of predictions depends greatly on the similarity of conditions at the location of the reference site versus the target site. In many cases, the correlation is poor at a single nearby site or a data set of a suitable length does not exist for forecasting. This causes the predictions to be generated using a site dissimilar from the target site returning unreliable yield estimates.

This work proposes a method using data from multiple reference stations to improve accuracy in predicting average wind speeds and yields at the site of interest: given conditions at our reference sites, a probability distribution of the wind speeds at our target site is established. By combining data selected from several sites based on specific similarity criteria, the results better capture the conditions at the target site. The statistical characteristics, persistence, availability, and diurnal variation of wind can be captured by using the data from several nearby reference stations as inputs in a fuzzy logic model. The resulting wind speeds are used as the averages in the probability distribution at the target site, as opposed to current methods resulting in point estimates. The probabilistic result gives greater insight in variation of the wind resource at the target site under varying climatic, seasonal and terrain conditions.

Data Collection

The target station used in the analysis is located in an inland region of relatively flat terrain at an elevation of 850 meters. Three meteorological reference stations maintained by Environment Canada within approximately 30 kilometers of the target site were selected. Observed temperature, 10 meter wind speed and 10 meter wind direction data recorded every hour over a five year period, December 2004 to December 2009, is available for all three reference stations. Data for the target site consists of ten-minute 60 meter wind speed and wind direction averages for the same five year period. Figure 1 shows the locations of the target site and reference stations relative to one another.

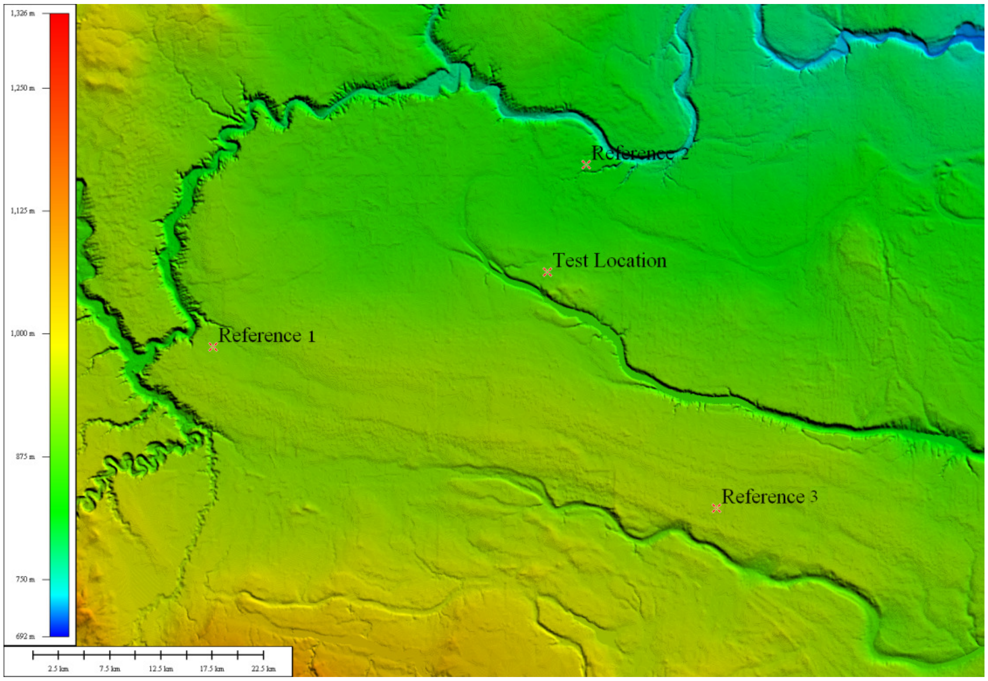

Figure 1: Map showing the location of meteorological stations. The distance between the Site and Reference 1 is 34 km, the Site and Reference 2 is 11 km, the Site and Reference 3 is 28 km.

Methodology

The data set is split into a one-year training period and four one-year verification sets. The training set is used to choose the inputs, train the fuzzy logic model and specify the probability distribution, while the verification data set is used for the validation of the model. Possible input variables include wind speeds, wind directions, month, and time of day for each of the three reference stations. One hour (+/-) lags/leads can also be incorporated into the reference data sets if desired – to account for the spatial temporal variations of meteorological conditions across the region. While additional variables aside from speeds and directions, such as temperature, may be added to the input selection simulations indicated no benefit in including such parameters for the present case. Using a sequential forward search the wind speed and directional candidates are then narrowed to the four inputs that most influence the output .

These selected inputs are fed into a primary fuzzy inference system to capture the behavior of the wind speed and direction. A secondary fuzzy inference system is used to capture the seasonal and temporal variation. Utilizing Sugeno-type fuzzy inference, the resulting model output is a type of weighted average. The resulting wind direction predictions are complete and require no further manipulation. For the wind speeds the desired outcome is a probability distribution rather than an average value. The differences between the modeled and actual wind speeds define this probability distribution of the wind speeds at the target site given the conditions at the reference sites. For practical implementation, the probability distribution is mimicked by generating six random numbers from this distribution such that for each wind speed at the target site, six predicted wind speeds are generated.

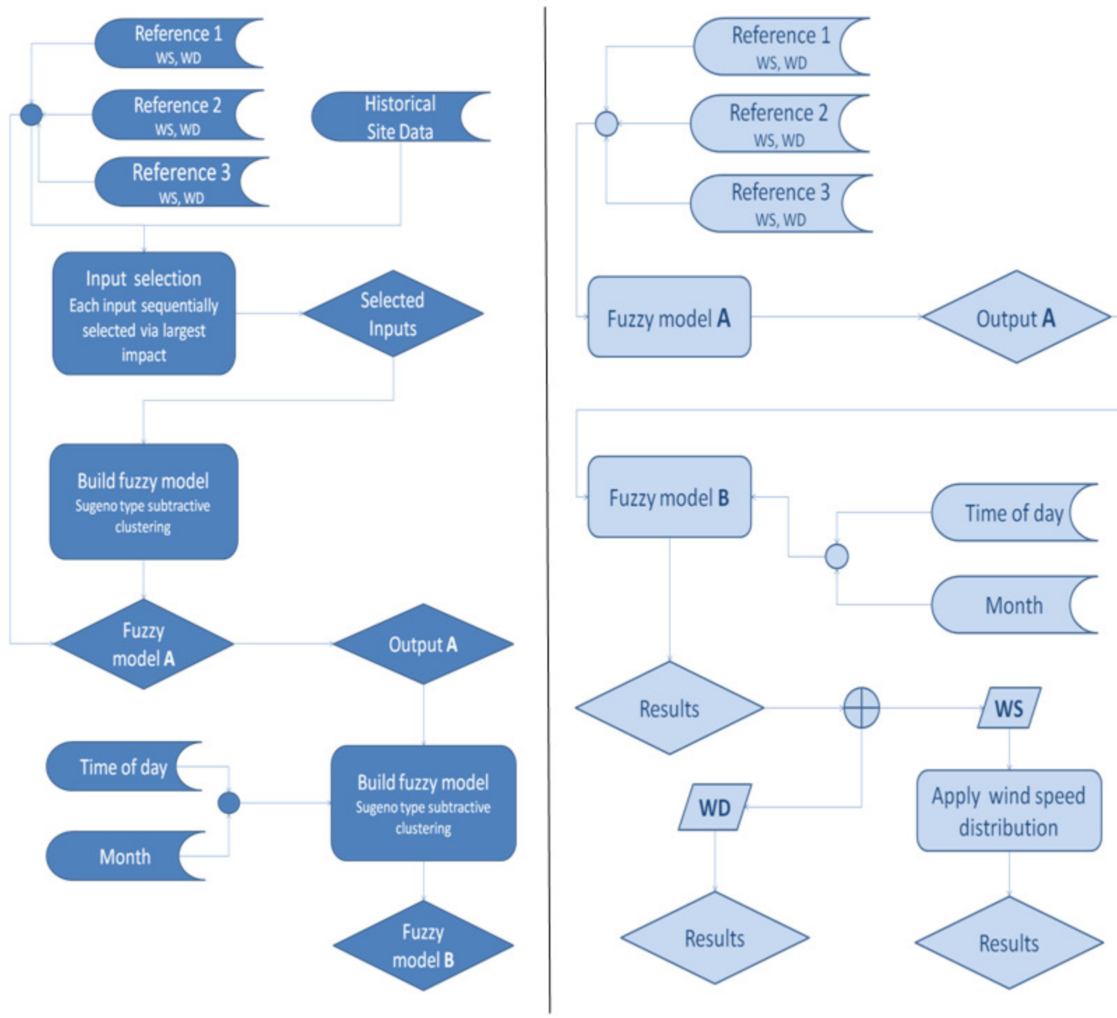

Figure 2: Summary of the model construction (left) and implementation (right).

Results and Discussion

The model was run with the first year of data, December 2004 - December 2005 or Year 1, used as the training set. The results were then used to predict the conditions at the target site for each subsequent year and then compared to the actual recorded yields. The results for Year 2 are presented in detail in Figure 3 through Figure 7, with the results of all years summarized in Table 1.

Figure 3 presents the wind speed distributions for the modeled and real (validation) data, as well as Weibull distribution fits. There is some slight discrepancy in the binned data between the model and real data sets. The real data distribution shows a double 'hump' feature while the model data does not. This characteristic is absent from the data used in the training set (Figure 4). This may explain why the model data generated for Year 2 does not exhibit this feature.

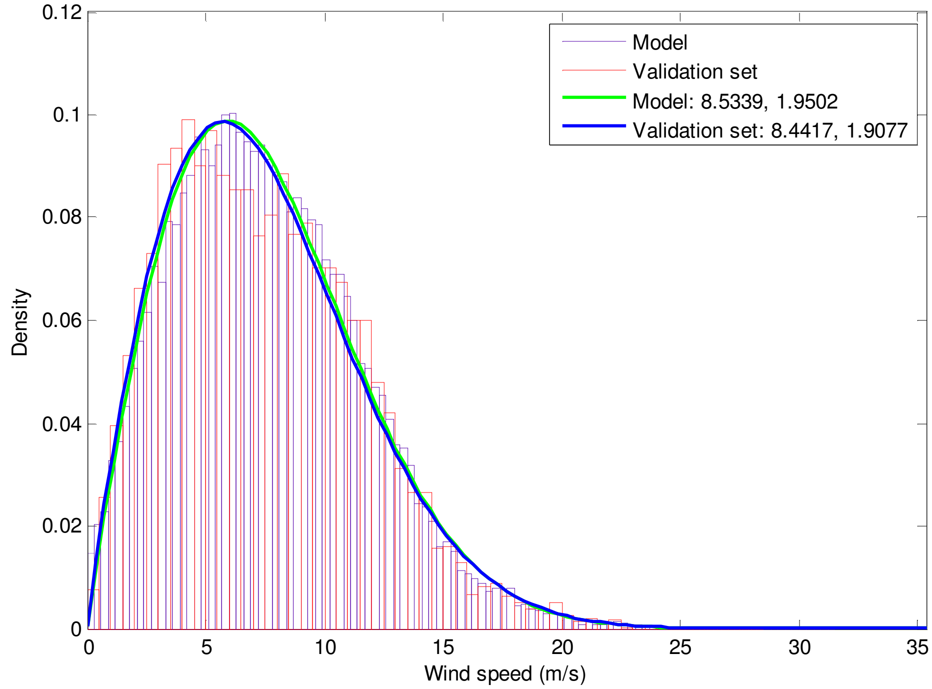

Figure 3: Weibull Distribution fit to wind speed data from the model and validation sets for Year 2. Note that the units presented here for *Data* are m/s and for *Density* is %.

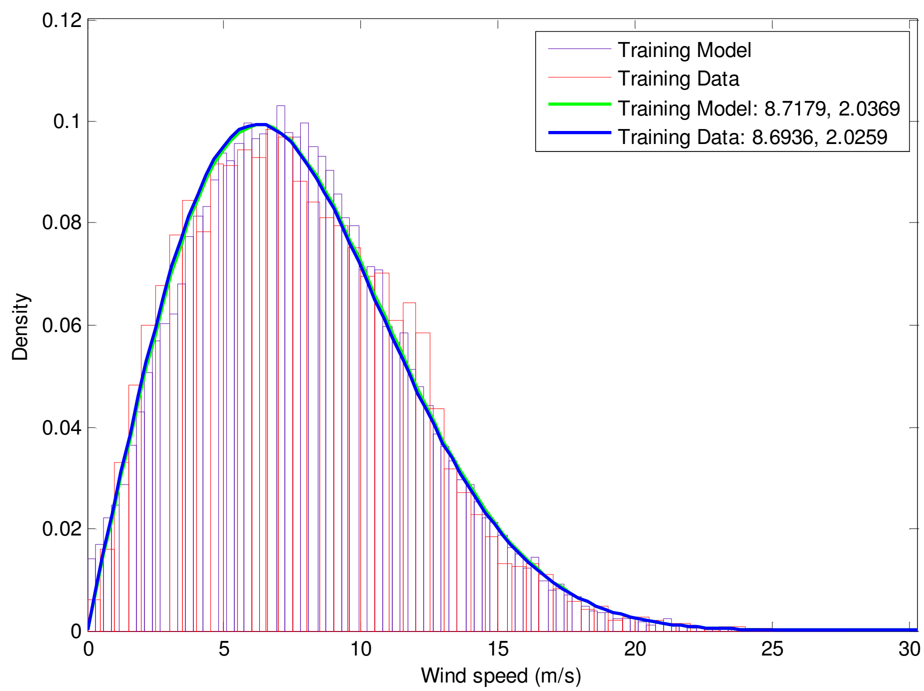

Figure 4: Weibull Distribution fit to wind speed data from the training model and training data. Note that the units presented here for *Data* are m/s and for *Density* is %.

The difference between the average wind speed obtained from the fuzzy system and actual wind speeds presented above is seen in Figure 5. The average error (0.036 m/s) between the simulated and real wind speeds for Year 2 is similar to the results from the training set (0.003 m/s).

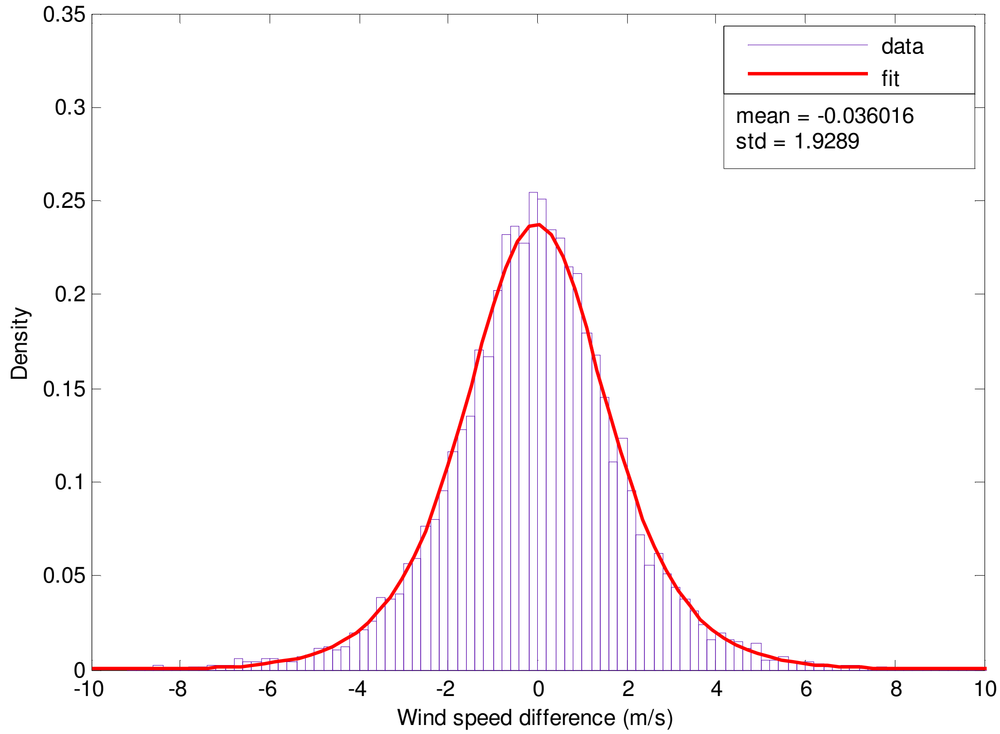

Figure 5: Distribution of MCP error. Note that the units presented here for *Data* are m/s and for *Density* is %.

The yields calculated on a monthly basis for Year 2 for both the real and modeled data sets are presented in Figure 6. These yields were calculated using a generic turbine performance curve and are the sums of the hourly yields. The largest relative error neared 10% with no marked seasonal trend. The results from the wind direction modeling are shown in Figure 7, overlaying both the actual and predicted wind direction distributions. It can be seen that the model agrees with the actual data.

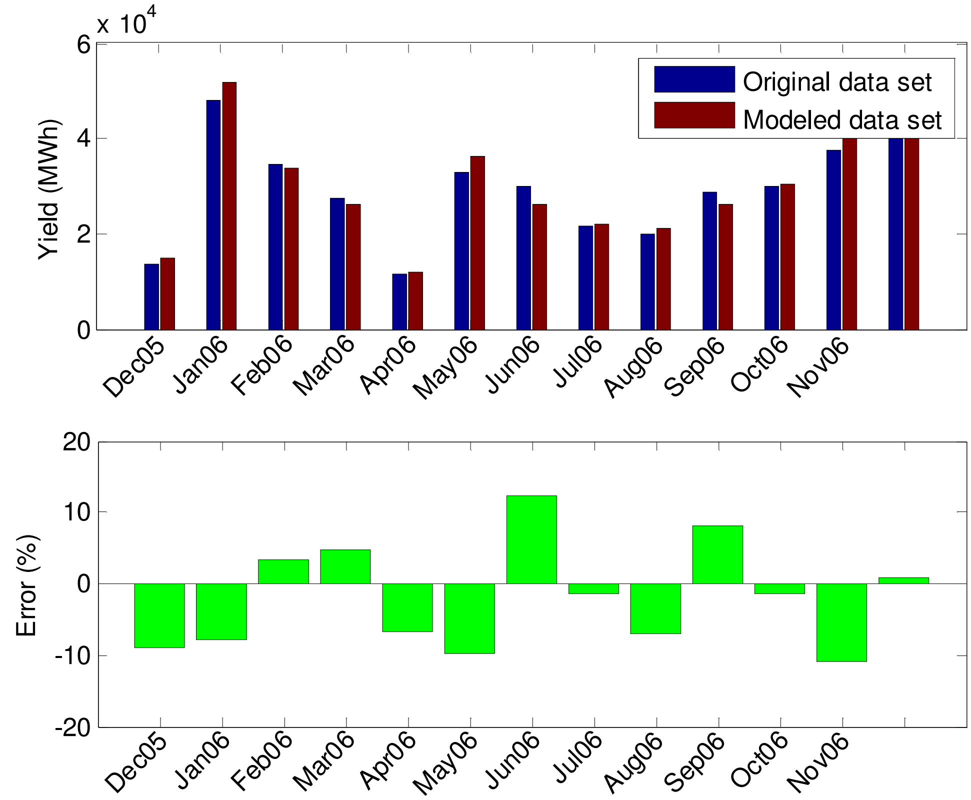

Figure 6: Predicted monthly power yields (MWh) at the target site using the model predictions versus the site data and resulting differences in yield (%).

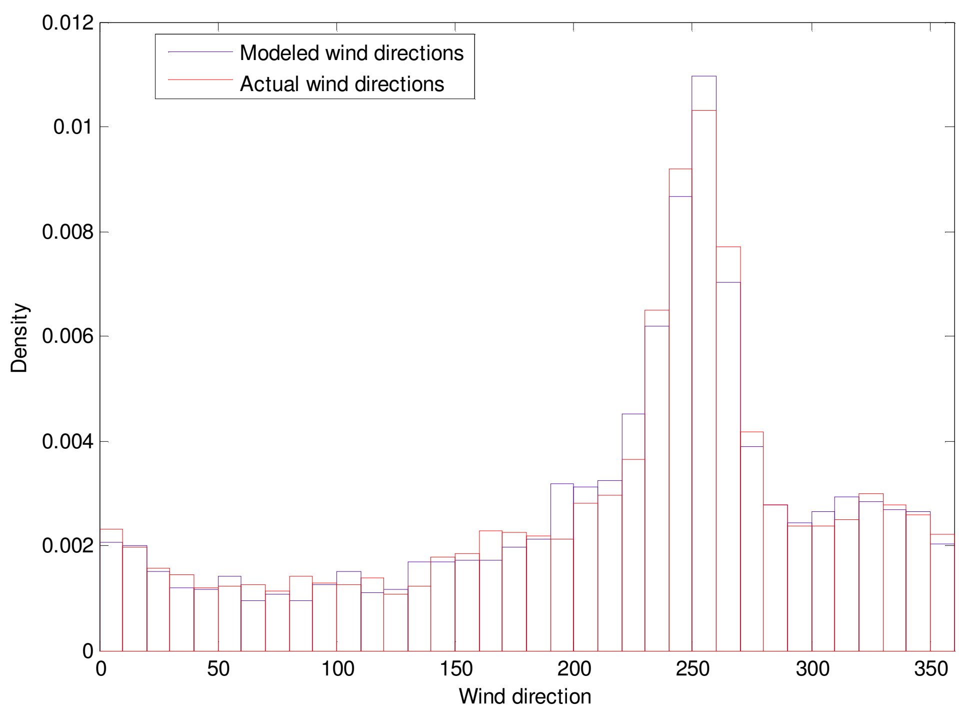

Figure 7: Wind direction distribution results.

Checks on the consistency of the method were performed. The model assumes the error mean to be zero and the error standard deviation to remain constant. From Table 1, we can see two clusters formed with respect to the wind speed error mean. To check the validity of the method, the model was run backwards with the last year of data, Year 5 used as the training set as shown in Table 2.

Table 1: Yield error and Weibull parameter estimates from model using Year 1 as training set.

	WS Error mean (m/s)	WS Error standard deviation (m/s)	Weibull parameters (validation set | model)		Overall yield error (%)
Train (12/23/04 - 12/23/05)	-0.00	1.77	8.69, 2.03	8.72, 2.04	0.62
Year 2 (12/23/05 - 12/23/06)	-0.04	1.93	8.44, 1.91	8.53, 1.95	1.46
Year 3 (12/23/06 - 12/23/07)	-0.08	1.90	8.94, 1.92	9.08, 1.92	2.69
Year 4 (12/23/07 - 12/23/08)	-0.13	1.80	8.29, 1.91	8.47, 1.95	4.41
Year 5 (12/23/08 - 12/23/09)	-0.12	1.88	8.50, 1.87	8.70, 1.89	4.47

Table 2: Yield error and Weibull parameter estimates from model using Year 5 as training set.

	WS Error mean (m/s)	WS Error standard deviation (m/s)	Weibull parameters (validation set | model)		Overall yield error (%)
Train (12/23/08 - 12/23/09)	-0.00	1.74	8.50, 1.87	8.54, 1.88	0.57
Year 4 (12/23/07 - 12/23/08)	0.01	1.78	8.29, 1.91	8.31, 1.92	0.93
Year 3 (12/23/06 - 12/23/07)	0.05	1.89	8.94, 1.92	8.92, 1.95	-1.02
Year 2 (12/23/05 - 12/23/06)	0.14	1.93	8.44, 1.91	8.31, 1.91	-3.40
Year 1 (12/23/04 - 12/23/05)	0.15	1.90	8.69, 2.02	8.54, 2.00	-3.91

The error is consistent between results obtained by the forward and backward training indicating the method to be robust. The change in mean error through the years using Year 1 as the training set mirrors that of the change in error mean using Year 5 as the training set. The differences in magnitude must be an attribute of the data itself. Hence, a site effect accounts for the pattern found in our mean error results. The error groups show Year 1 and Year 2 as one cluster, Year 4 and Year 5 as the other cluster, and Year 3 stands alone as a transition year. An investigation of the target site history indicates that maintenance comprising of wind sensor replacements occurred halfway into Year 3. The same anemometer was used during Year 1 and Year 2. The new anemometer was deployed for Year 4 and Year 5. Both anemometers were non-calibrated sensors. This equipment change most likely accounts for this observed discrepancy.

By incorporating wind direction variables as inputs to the fuzzy inference system, the average predicted wind speed at our target site are more accurately estimated based on the wind directions and terrain conditions on site. Figure 8 shows the predicted average speed of the wind at the target site with the wind speed at all three reference sites set to 5 m/s. For the same wind speed conditions at the reference sites, the wind speed at the target site depends strongly on where the wind is coming from. Thus, for a more accurate prediction, not only are the magnitudes of wind speed at reference sites used, but the direction is also utilized to specify the average conditions at the target site.

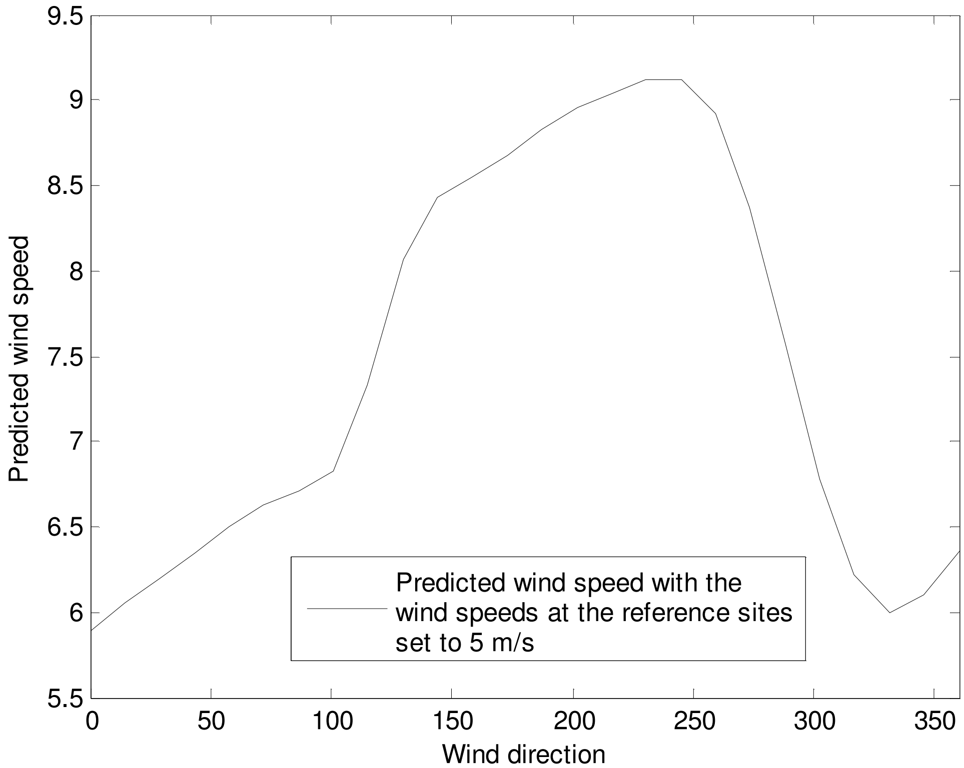

Figure 8: Predicted wind speed defined by direction at the target site given winds speeds of 5 m/s at all three reference stations.

Seasonal and diurnal characteristics of the wind on the target site are also captured by this method. While the primary fuzzy inference system deals with the wind conditions at the reference sites, the secondary fuzzy inference system uses the month and time of day to further refine the prediction of the wind speeds at the target site. For the same wind conditions at the reference sites the output of the secondary fuzzy inference system varies significantly depending on the season and the time of day. For a constant wind speed of 9 m/s obtained from the primary fuzzy inference system, the differences in predicted daily wind speeds for the months of January and July can be seen in Figure 9. The wind at the target site is more consistent in speed and of a higher magnitude during the winter months as seen in the January plot. The July plot shows the summer characteristics of more variable wind throughout the day with a dip through the daylight hours and an overall lower average speed. Both the January and July curves in Figure 9 are consistent with the daily shear profile exponents observed at the target site.

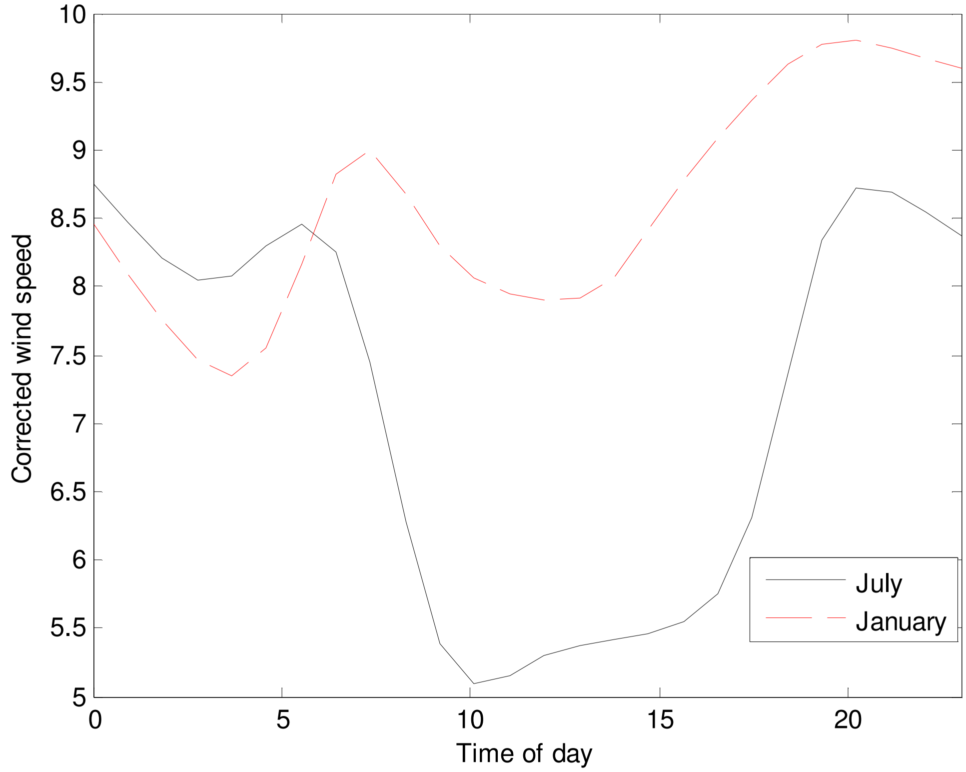

Figure 9: Predicted daily average wind speeds (m/s) at the target site for the months of January (dashed) and June (solid) given a constant wind speed of 9 m/s obtained from the primary fuzzy inference system.

Conclusions

Using data from multiple reference stations as inputs to a fuzzy logic model is a valid MCP method for use in predicting average wind speeds and yields at the site of interest. Results from our simulations show a strong similarity between the actual and predicted wind regimes. The results demonstrate the ability to predict overall yields within a maximum error of approximately 1%. The method also captures some of the seasonal, diurnal, and terrain features which are key factors for the prediction of wind speeds.