Proceedings of the ASME 2013 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference
IDETC/CIE 2013
August 4-7, 2013, Portland, Oregon, USA

DETC2013-12695

ASSESSING LONG-TERM WIND CONDITIONS BY COMBINING DIFFERENT MEASURE-CORRELATE-PREDICT ALGORITHMS

Jie Zhang*
National Renewable Energy Laboratory
Golden, Colorado 80401
Email: Jie.Zhang@nrel.gov

Souma Chowdhury†
Syracuse University
Syracuse, New York, 13244
Email: sochowdh@syr.edu

Achille Messac‡
Syracuse University
Syracuse, New York, 13244
Email: messac@syr.edu

Bri-Mathias Hodge
National Renewable Energy Laboratory
Golden, Colorado 80401
Email: Bri.Mathias.Hodge@nrel.gov

ABSTRACT

This paper significantly advanced the hybrid measure-correlate-predict (MCP) methodology, enabling it to account for the variations of both wind speed and direction. The advanced hybrid MCP method used the recorded data of multiple reference stations to estimate the long-term wind condition at the target wind plant site with greater accuracy than possible with data from a single reference station. The wind data was divided into different sectors according to the wind direction, and the MCP strategy was implemented for each wind sector separately. The applicability of the proposed hybrid strategy was investigated using four different MCP methods: (i) linear regression; (ii) variance ratio; (iii) artificial neural networks; and (iv) support vector regression. To implement the advanced hybrid MCP methodology, we used the hourly averaged wind data recorded at six stations in North Dakota between the years 2008 and 2010. The station Pillsbury was selected as the target plant site. The recorded data at the other five stations (Dazey, Galesbury, Hillsboro, Mayville, and Prosper) was used as reference station data. The best hybrid MCP strategy from different MCP algorithms and reference stations was investigated and selected from the 1,024 combinations. The accuracy of the hybrid MCP method was found to be highly sensitive to the combination of individual MCP algorithms and reference stations used. It was also observed that the best combination of MCP algorithms was strongly influenced by the length of the correlation period.

Keywords: Measure-correlate-predict (MCP), power generation, resource assessment, wind distribution, wind energy

INTRODUCTION

Wind resource assessment is the process of estimating the power potential of a wind plant site and has been playing an important role in a wind energy project. In general, wind resource assessment includes (i) onsite wind conditions measurement; (ii) correlations between onsite meteorological towers to fill in missing data; (iii) correlations between long-term weather stations and short-term onsite meteorological towers; (iv) analysis of the wind shear and its variations; (v) modeling of the distribution of

*Postdoctoral Researcher, Transmission and Grid Integration Group, ASME member, Corresponding author.
†Research Assistant Professor, Multidisciplinary Design Optimization Laboratory (MDOL), ASME member.
‡Distinguished Professor and Department Chair, Department of Mechanical and Aerospace Engineering, ASME Fellow.

wind conditions; and (vi) prediction of the available energy at the site. MCP algorithms are used to predict the long-term wind resource at target sites using the short-term (one- or two-year) onsite data, and the co-occurring data at nearby meteorological stations (that also have long-term data). The accuracy of long-term predictions obtained using MCP methods is subject to (i) the availability of a nearby meteorological station; (ii) the uncertainty associated with a specific correlation methodology [1]; and (iii) the likely dependence of this correlation on physical features such as the topography, the distance between the monitoring stations, and the type of the local climate regime [2].

A wide variety of MCP techniques have been reported in the literature, such as: (1) linear regression [3, 4]; (2) variance ratio [4, 5]; (3) Weibull scale [5]; (4) artificial neural networks (ANNs) [2, 3, 6]; (5) support vector regression (SVR) [7, 8]; (6) Mortimer [2]; and (7) wind index MCP [9]. MCP methods were first used to estimate the long-term annual mean wind speed [10, 11]. Linear regression [12] was presented to characterize the relationship between the reference and target site wind speeds. Rogers et al. [13] compared four MCP algorithms: (i) a linear regression model; (ii) a model using distributions of ratios of wind speeds at the two sites; (iii) a vector regression method; and (iv) a method based on the ratio of the standard deviations of the two data sets. Perea et al. [4] proposed and evaluated three MCP methods based on concurrent wind speed time series for two sites: (i) linear regression derived from bivariate normal joint distribution; (ii) Weibull regression; and (iii) approaches based on conditional probability density functions.

Given the unavoidable practical constraints, the overall reliability of the predicted long-term wind distribution remains highly sensitive to the one-year distribution of recorded on-site data. Quantifying and modeling the uncertainty in the MCP methods would provide more credibility to wind resource assessment and wind plant performance estimation. Kwon [14] and Lackner et al. [15] presented different frameworks to analyze uncertainty in MCP-based wind resource assessment. The wind resource-based uncertainty models proposed by Messac et al. [16] can be applied also to the long-term data recorded at meteorological stations when MCP methods are used.

Research Objectives and Motivation

The hybrid MCP method recently developed by Zhang et al. [17] combines the component MCP algorithms by characterizing the distance and elevation difference between reference stations and the target wind plant site. The overall objective of this paper is to significantly advance the original hybrid MCP methodology by:

1. Considering both the wind speed and direction as the components of the hybrid MCP methodology; and
2. Investigating the best combination of different MCP methods and reference stations.

The advancements to the hybrid MCP method is presented in the next section. The application of the advanced hybrid MCP method and the corresponding results and discussion are presented in Section III. Section IV presents the concluding remarks of this research.

ADVANCING THE HYBRID MCP METHOD

A brief overview of the original hybrid MCP methodology is first presented, followed by the description of the advancements introduced in this paper.

Overview of the Original Hybrid MCP Method

The hybrid MCP method developed by Zhang et al. [17] correlates the wind data at the targeted wind plant site with that at multiple reference stations. The strategy accounts for the local climate and the topography information. In the original hybrid MCP method, all component MCP estimations between the targeted wind plant site and each reference station use a single MCP method (e.g., linear regression, variance ratio, Weibull scale, or neural networks).

The weight of each reference station in the hybrid strategy is determined based on: (i) the distance and (ii) the elevation differences between the target wind plant site and each reference station. The hypothesis here is that the weight of a reference station is larger when the reference station is closer (shorter distance and smaller elevation difference) to the target wind plant site. The weight of each reference station, w_i, is determined by

$$w_i = \frac{1}{2(n_{ref}-1)}\left(\frac{\sum_{j=1,j\neq i}^{n_{ref}}\Delta d_j}{\sum_{j=1}^{n_{ref}}\Delta d_j} + \frac{\sum_{j=1,j\neq i}^{n_{ref}}\Delta h_j}{\sum_{j=1}^{n_{ref}}\Delta h_j}\right) \qquad (1)$$

where n_{ref} is the number of reference stations; and Δd_j and Δh_j represent the distance and the elevation difference between the target plant site and j^{th} reference station, respectively.

In the following subsections, we briefly discuss how this paper advances the key components of the original hybrid MCP method. These advanced features provide helpful flexibility to the hybrid MCP method, and extends its applicability to designing full-scale commercial wind plants.

Modeling the Impact of Wind Direction on the Hybrid MCP Performance

Each wind data point was allocated to a bin according to the wind direction sector measurement at the target wind plant site. In this paper, we investigated four cases by choosing to bin into different number of sectors: (i) 4 sectors; (ii) 8 sectors; (iii) 16 sectors; and (iv) 32 sectors. At the multiple reference stations, the concurrent wind speed and direction measurement was allocated to the corresponding bin. Within each sector, the long-term wind speed was predicted by applying the hybrid MCP strategy based on the concurrent short-term wind speed data within that sector. A wind rose is a graphical tool used by meteorologists to provide a succinct illustration of how wind speed and wind direction are distributed at a location. Figure 1 shows a wind rose diagram with 16 direction sectors.

By putting the wind speed data in each sector together, we obtained the set of long-term wind data at the target wind plant site. The quality of the predicted long-term wind data was evaluated using the performance metrics described in the following subsection.

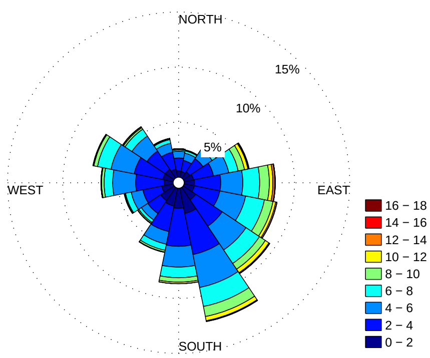


Figure 1. WIND ROSE DIAGRAM WITH 16 SECTORS

Performance Metrics for Evaluating the MCP Method

Three sets of performance metrics were proposed to evaluate the performance of the MCP methods: (i) statistical metrics; (ii) wind distribution metrics; and (iii) wind plant performance: power generation and capacity factor metrics.

Statistical Metrics Mean long-term wind speed is often used to characterize the potential of a wind plant site. It is an important measure of wind power potential. Four metrics are evaluated based on the estimated wind speeds by MCP methods and the reference wind speeds: (i) the ratio of mean wind speeds [13]; (ii) the ratio of wind speed variances; (iii) root mean squared error (RMSE) [18]; and (iv) maximum absolute error (MAE).

The ratio of mean wind speeds, R_μ, is expressed as

$$R_\mu = \frac{1/n_t \sum_{i=1}^{n_t} \tilde{v}(t^k)}{1/n_t \sum_{i=1}^{n_t} v(t^k)} \tag{2}$$

where $v(t^k)$ represents the measured hourly averaged wind speed at time t^k at the targeted wind plant site, $\tilde{v}(t^k)$ is the corresponding estimated wind speed value, and n_t is the total number of paired data points used in the analysis.

The ratio of wind speed variances, R_{σ^2}, is expressed as

$$R_{\sigma^2} = \frac{1/n_t \sum_{i=1}^{n_t} \left(\tilde{v}(t^k) - \tilde{\mu}\right)^2}{1/n_t \sum_{i=1}^{n_t} \left(v(t^k) - \mu\right)^2} \tag{3}$$

where $\tilde{\mu}$ and μ represent the mean of the estimated and measured wind speeds of all test paired data points.

The RMSE is given by

$$RMSE = \sqrt{\frac{1}{n_t} \sum_{k=1}^{n_t} \left(v(t^k) - \tilde{v}(t^k)\right)^2} \tag{4}$$

The MAE is expressed as

$$MAE = \max_k \left| v(t^k) - \tilde{v}(t^k) \right| \tag{5}$$

Wind Distribution Metrics Wind speed distributions are necessary to quantify the available energy (power density) at a site and to design optimal wind plant configurations. The Multivariate and Multimodal Wind Distribution (MMWD) model [19, 20] can capture the joint variation of wind speed, wind direction, and air density, and also allows representation of multimodally distributed data.

The MMWD model was developed based on the *Kernel Density Estimation* (KDE) [19, 20]. For a $d-variate$ random sample $U_1, U_2, \cdots, U_n$ drawn from a density f, the multivariate KDE is defined as

$$\hat{f}(u;H) = n^{-1} \sum_{i=1}^{n} K_H\left(u - U_i\right) \tag{6}$$

where $u = (u_1, u_2, \cdots, u_d)^T$ and $U_i = (U_{i1}, U_{i2}, \cdots, U_{id})^T$, $i = 1, 2, \cdots, n$. Here, $K(u)$ is the kernel that is a symmetric probability density function; H is the bandwidth matrix, which is symmetric and positive-definite; and $K_H(u) = |H|^{-1/2} K(H^{-1/2} u)$. The choice of K is not crucial to the accuracy of kernel density estimators [21]. In this paper, $K(u) = (2\pi)^{-d/2} exp\left(-\frac{1}{2} u^T u\right)$ is considered, the standard normal throughout. In contrast, the choice of H is crucial in determining the performance of $\hat{f}$ [22]. In the MMWD model, an optimality criterion, the *asymptotic mean integrated squared error* [22], is used to select the bandwidth matrix. The details of the MMWD model can be found in Ref. [19].

Wind Plant Performance Metrics This power generation model was adopted from Chowdhury et al. [23, 24]. The power generated by a wind plant is an intricate function of the configuration and location of the individual wind turbines. The flow pattern inside a wind plant is complex, primarily because of the wake effects and the highly turbulent flow. The power generated by a wind plant (P_{plant}) consisting of N wind turbines is evaluated as a sum of the powers generated by the individual turbines, which is expressed as [23]

$$P_{plant} = \sum_{j=1}^{N} P_j \tag{7}$$

Accordingly, the wind plant efficiency can be expressed as

$$\eta_{plant} = \frac{P_{plant}}{\sum_{j=1}^{N} P_{0j}} \tag{8}$$

where P_{0j} is the power that $turbine - j$ would generate if operating as a stand-alone entity, for the given incoming wind velocity. Detailed formulation of the power generation model can be found in the papers [23, 24].

The power generated by a wind plant with nine turbines was evaluated in this paper. Two types of wind turbines were selected: (i) the *GE 1.5-MW-XLE* [25] and (ii) the *GE 2.5-MW-XL* [26].

CASE STUDY: ASSESSING THE WIND RESOURCE POTENTIAL AT A WIND PLANT SITE

To implement the advanced hybrid MCP methodology, we used the hourly averaged wind data recorded at six stations in North Dakota between 2008 and 2010. The station Pillsbury was selected as the target wind plant site. The recorded data at the other five stations (Dazey, Galesbury, Hillsboro, Mayville, and Prosper) was used as reference station data. The station locations were shown in Fig. 2; the six stations inside the outer rectangle were used in this paper.

The wind data is obtained from the North Dakota Agricultural Weather Network (NDAWN) [27]. Table 1 shows the geographical coordinates and elevation of each station. The measurement information is listed as follows.

1. Wind speed is measured at 3 meters above the soil surface with an anemometer.
2. Wind direction is the direction from which wind is blowing (degrees clockwise from north) measured at 3 meters above the soil surface ($N = 0°; NE = 45°; E = 90°; SE = 135°; S = 180°; SW = 225°; W = 270°; NW = 315°; etc.$) with a wind vane. The value is the average of all measured wind directions for a 24-hour period from midnight to midnight.

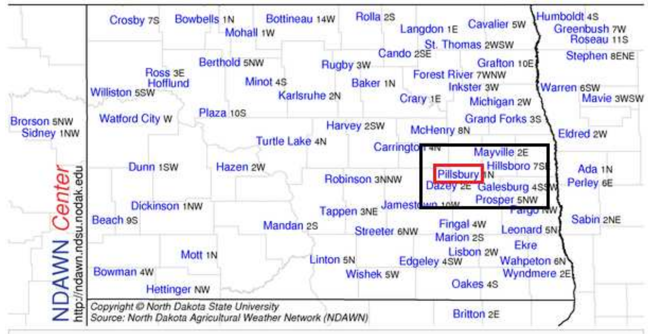


Figure 2. NDAWN STATION LOCATIONS [27]

Table 1. DETAILS OF NDAWN STATIONS [27]

Station	Latitude	Longitude	Elevation (m)
Dazey	47.183	-98.138	439
Galesburg	47.210	-97.431	331
Hillsboro	47.353	-96.922	270
Mayville	47.498	-97.262	290
Pillsbury	47.225	-97.791	392
Prosper	47.002	-97.115	284

Component MCP Methods

In this research, four MCP methods were investigated: (i) linear regression; (ii) variance ratio; (iii) ANNs; and (iv) SVR. It is helpful to note that other MCP methods can also be used in conjunction with the hybrid strategy, because the weights determination strategy is independent with the MCP method.

The Linear Regression Method Linear regression is a common method to characterize the relationship between the reference and target site wind speeds. The prediction equation is given as

$$\hat{y} = ax + b \tag{9}$$

where $\hat{y}$ is the predicted wind speed at the target site; x is the observed wind speed at the reference site; and a and b are the estimated intercept and slope of the linear relationship.

The Variance Ratio Method When using linear regression, the predicted mean wind speed at the target site will be close in value to the measured mean during the training interval. However, the predicted variance at the target site will be less than the measured variance. This can result in biased predictions of wind speed distributions.

The variance ratio method was proposed in response to the above limitations of linear regression. It involves forcing the variance of the predicted wind speed at the target site to be equal to the measured variance at the target site. The prediction equation is express as

$$\hat{y} = \mu_y - \frac{\sigma_y}{\sigma_x}\mu_x + \frac{\sigma_y}{\sigma_x}x \tag{10}$$

where μ_x, μ_y, σ_x and σ_y are the means and standard deviations of the two concurrent data sets.

The ANNs Method ANNs have been used to correlate and predict wind conditions because of their ability to recognize patterns in noisy or otherwise complex data. A neural network generally contains an input layer, one or more hidden layers, and an output layer. An ANN is developed by defining the following parameters:

1. The interconnection pattern between different layers of neurons;
2. The learning process for updating the weights of the interconnections; and
3. The activation function that converts a neuron's weighted input to its output activation.

The SVR Method SVR has gained popularity, both within the statistical learning community [28, 29] and within the engineering optimization community [30–32]. The SVR approach provides a unique way to construct smooth, nonlinear, regression approximations by formulating the surrogate model construction problem as a quadratic programming problem. The SVR approach can be expressed as [33]

$$\tilde{f}(x) = \langle w, \Phi(x) \rangle + b \tag{11}$$

where $\langle \cdot, \cdot \rangle$ denotes the dot product; w is a set of coefficients to be determined; and $\Phi(x)$ is a map from the input space to the feature space. To solve the coefficients, we can allow a predefined maximum tolerated error ε (with respect to the actual function value) at each data point, given by [33]

$$|\tilde{f}(x_i) - f(x_i)| \leq \varepsilon \tag{12}$$

where $f(x)$ is the actual function to be approximated. The flatness of the approximated function can be characterized by w. By including slack variables ξ to the constraint and a cost function, the coefficient w can be obtained by solving a quadratic programming problem given by [33]

$$\begin{aligned} \min_{w,\xi,\xi^*} \quad & \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n_p}(\xi_i + xi_i^*) \\ \text{subject to} \quad & f(x_i) - \tilde{f}(x_i) \leq \varepsilon + \xi_i \\ & f(x_i) - \tilde{f}(x_i) \leq \varepsilon + \xi_i^* \\ & \xi_i, \xi_i^* \geq 0 \end{aligned} \tag{13}$$

where n_p is the number of sample points. The parameter $C > 0$ is user-specified and represents the trade-off between flatness and the amount up to which errors larger than ε are tolerated. The above formulation is the primal form of the quadratic programming problem. In most cases, the dual form with fewer constraints is easier to solve, and is widely used to define the final form of the approximation. It can be shown that the dual form is convex and therefore has a unique minimum. Typical allowed mapping functions are radial basis functions, such as the gaussian function.

Selection of Parameters

The MATLAB Neural Network Toolbox [34] was used in this paper. The Levenberg-Marquardt algorithm was selected for neural network training. Eighty percent data points were randomly selected as training points; and 20 percent points were used to validate the network. The MSE metric was used to evaluate the performance of the developed neural network. For the SVR method, we used an efficient SVM package, LIBSVM (A Library for Support Vector Machines), developed by Chang and Lin [35].

Results and Discussion

Two scenarios were analyzed in the case study. In the first, the hybrid MCP strategy was implemented without binning wind data points to different sectors. The objective of the first scenario was to investigate the performance of the hybrid MCP method with mixing combinations of MCP algorithms and reference stations. In the second scenario, we evaluated the hybrid MCP performance, including consideration of both wind speed and direction. However, this paper discussed wind direction primarily related to the prediction of wind speed. The prediction of the long-term wind direction at the target wind plant site was not within the scope of this paper.

Scenario I: Hybrid MCP Methods with Mixing Combinations A preliminary comparison of the hybrid MCP method (using multiple reference stations) with the individual MCP method (using one reference station) was investigated in the paper by Zhang et al. [17]. However, the hybrid MCP methods discussed in Ref. [17] used a single MCP technique for all five reference stations, which might not be optimal. Each reference station has the flexibility to use any of the available MCP techniques. In this case study, the same five stations were selected as reference stations. In addition, each station can be combined into the hybrid MCP method with one of the four following MCP algorithms: (i) linear regression; (ii) variance ratio; (iii) neural network; and (iv) SVR. Therefore, a total of 1,024 (which is equal to 4^5) combinations were investigated to formulate the hybrid MCP strategy.

Figs. 3-5 illustrate the three sets of performance metrics. Each line in the figures represents one specific combination of stations and MCP algorithms. It was observed that the performance of the hybrid MCP technique varied significantly. For instance, the average value of RMSE during the length of correlation period (Fig. 3(c)) varied 4.91% during the 1,024 hybrid MCP models.

The smaller the RMSE value, the more accurate the estimated wind pattern. Based on the RMSE values, we found the best combination of MCP algorithms and reference stations, shown in Table 2. In the table, “Ratio,” “Linear,” “ANN,” and “SVR” represent variance ratio, linear regression, artificial neural networks, and support vector regression, respectively. Four combinations were observed based on the length of the correlation period. It was shown from Table 2 that: (i) the variance ratio algorithm was chosen at the Dazey, Mayville, and Prosper reference stations for all four correlation periods; (ii) the SVR algorithm was selected at the Galesbury reference station when the correlation period was between 3,500-11,500 hours; and (iii) at the Hillsboro reference station, different MCP algorithms were selected based on the length of the correlation period.

The two-parameter Weibull distribution was the most widely accepted distribution for wind speed. The shape parameter (k) and the scale parameter (c) determine the probability distribution. In this research, the ratios of $\tilde{k}$ (and $\tilde{c}$) for the predicted wind speeds to k (and c) for the observed targeted wind speeds, R_k and R_c, were evaluated. The values of R_k and R_c are given by

$$R_k = \frac{\tilde{k}}{k}, \quad and \quad R_c = \frac{\tilde{c}}{c} \tag{14}$$

Figure 4 shows the normalized Weibull k and c parameters for the total 1,024 combinations. The closer the value of the ratio is to one, the more accurate the estimated long-term wind condition. The average values of R_k and R_c varied 1.94% and 15.82%, respectively, throughout the 1,024 different combinations. The above observation indicates that the scale parameter (c) is more sensitive to the MCP strategies than the shape parameter (k).

The power generation of the nine-turbine wind plant is shown in Fig. 5. Figs. 5(a) and 5(b) show the 1,024 wind power generations of wind plants with 1.5-MW and 2.5-MW turbines, respectively. It was observed that: (i) the average value of the power generation with GE 1.5MW-XLE turbines during the length of correlation period (Fig. 5(a)) varied 5.57% during the 1,024 hybrid MCP models; and (ii) the average value of the power generation with GE 2.5MW-XL turbines during the length of correlation period (Fig. 5(b)) varied 4.81% during the 1,024 hybrid MCP models.

Scenario II: Hybrid MCP Methods Considering Wind Speed and Direction Each wind data point was allocated to a bin according to the wind direction sector measurement at the target plant site (Pillsbury station). In the case study, four cases were investigated: (i) 4 sectors; (ii) 8 sectors; (iii) 16 sectors; and (iv) 32 sectors. For the five reference stations (Dazey, Galesbury, Hillsboro, Mayville, and Prosper), the concurrent wind speed and direction measurement was allocated to the corresponding bin. Within each sector, the long-term wind speed was predicted by applying the hybrid MCP strategy based on concurrent short-term wind speed data within that sector. In Scenario II, the hybrid MCP strategy used a single MCP technique for all five reference stations.

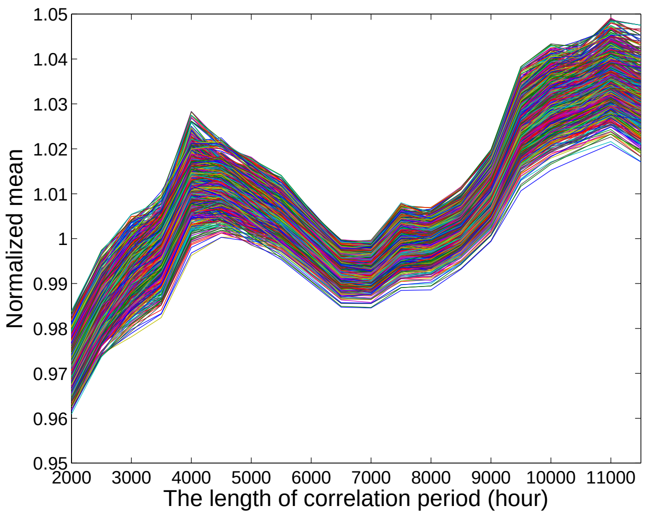
Figure 6 shows the ratio of mean wind speeds with the four different direction sectors. The closer the value of the ratio is to one, the more accurate the estimated wind pattern. It was observed that: (i) the hybrid MCP methods performed relatively better than the the individual MCP algorithms; (ii) the hybrid SVR algorithm performed relatively worse than the other three hybrid MCP methods; (iii) the hybrid linear regression, the hybrid variance ratio, and the hybrid neural network methods performed best when the correlation period was between 5,500-8,500 hours (approx. 8 to 12 months).

Figure 7 shows the wind speed distributions with the four different direction sectors. The closer the predicted distribution curve is to the actual distribution curve (the black line in the figure), the more accurate the estimated wind pattern. It was observed that: (i) for 8, 16, and 32 direction sectors, the hybrid variance ratio method (solid blue line) agreed more with the actual distribution curve (solid black line) than other MCP methods; (ii) for 4 direction sectors, the individual variance ratio method (dashed blue line) agreed more with the actual distribution curve.

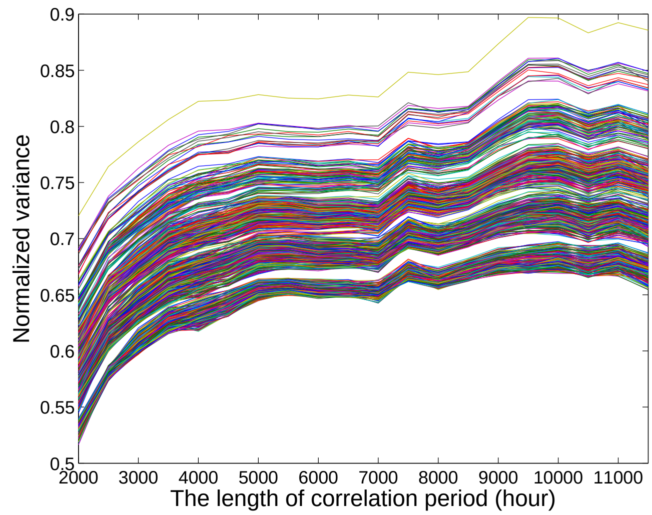
Figure 8 shows the wind power generation of the wind plant with nine GE 1.5MW-XLE turbines. The closer the predicted power generation curve to the actual power generation curve (the black line in the figure), the more accurate the estimated wind pattern. The predicted power generation of the wind plant was estimated using the long-term wind data predicted by the hybrid MCP method; the actual wind plant power generation was estimated using the measured long-term wind data. We observed that: (i) the hybrid variance ratio method performed best when the correlation period was between 2,000-3,500 hours (approxi-

Station	2,000-3,500 (hours)	3,500-5,500 (hours)	5,500-8,500 (hours)	8,500-11,500 (hours)
Dazey	Ratio	Ratio	Ratio	Ratio
Galesburg	Ratio	SVR	SVR	SVR
Hillsboro	Linear	Linear	ANN	SVR
Mayville	Ratio	Ratio	Ratio	Ratio
Prosper	Ratio	Ratio	Ratio	Ratio

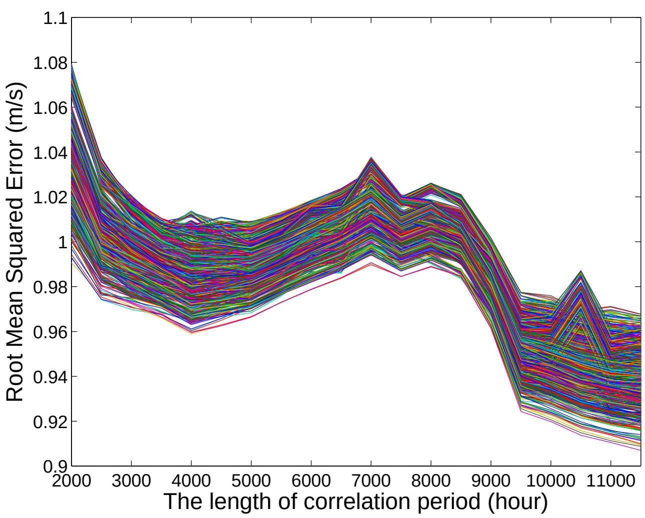
Table 2. BEST COMBINATION OF MCP ALGORITHMS AND REFERENCE STATIONS



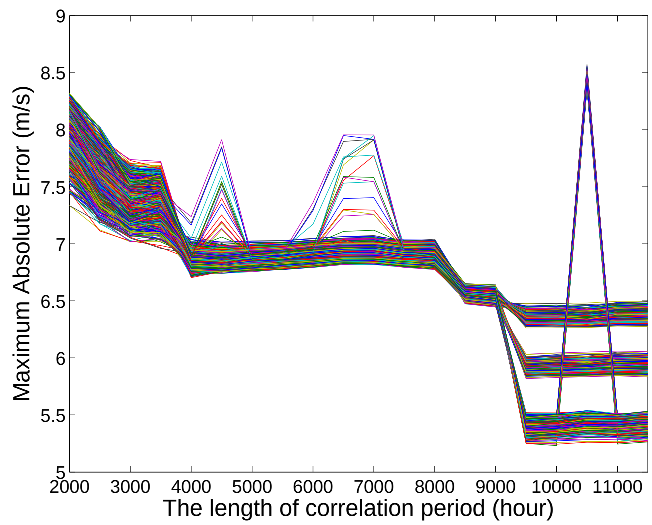
(a) Mean ratio



(b) Variance ratio

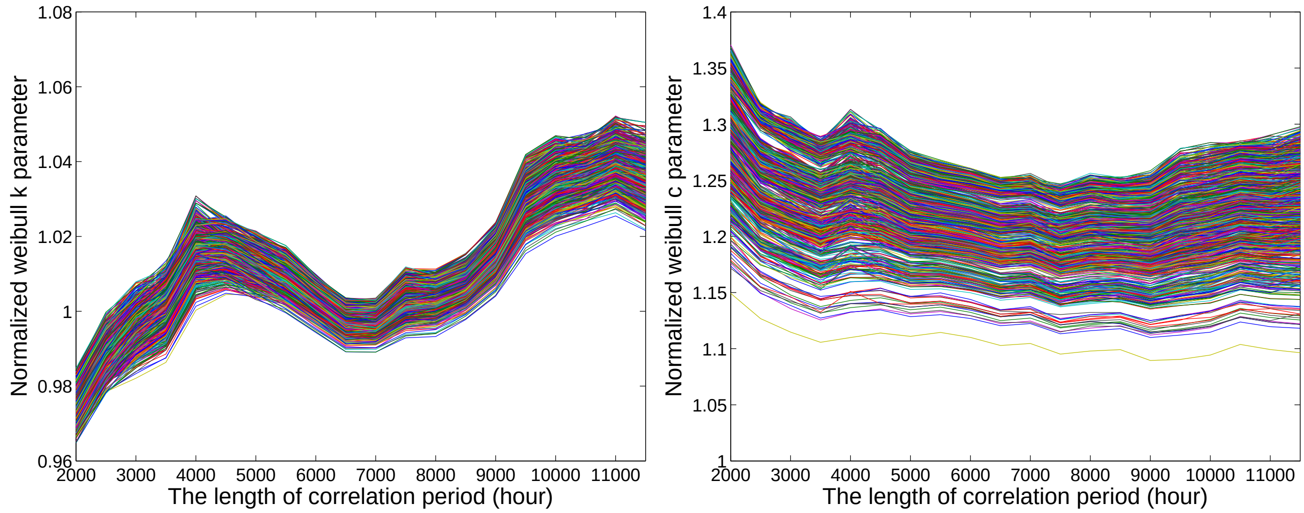


(c) RMSE



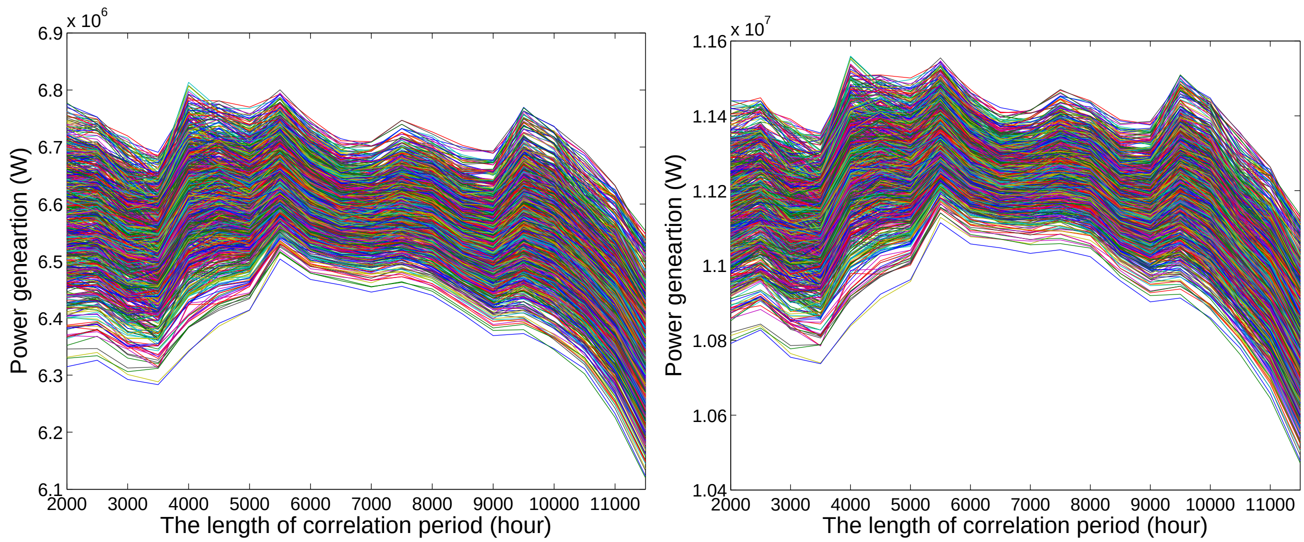
(d) MAE

Figure 3. Statistical performance metrics to evaluate the hybrid MCP with 1,024 combinations



(a) Normalized Weibull k parameter (b) Normalized Weibull c parameter

Figure 4. WIND DISTRIBUTION METRICS TO EVALUATE THE HYBRID MCP WITH 1,024 COMBINATIONS



(a) Power generation with *GE 1.5-MW-XLE* turbines (b) Power generation with *GE 2.5-MW-XL* turbines

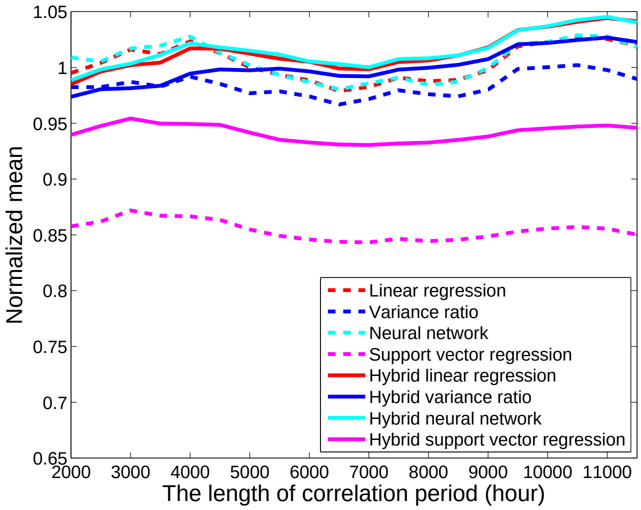
Figure 5. POWER GENERATION METRICS TO EVALUATE THE HYBRID MCP WITH 1,024 COMBINATIONS

mately 2.5 to 5 months) for all four direction sectors; (ii) the linear regression and the neural network methods had relatively better power generation estimations when the correlation period was between 6,000-9,000 hours (approximately 8.5 to 12.5 months); (iii) the power generation was generally overestimated by the neural network, hybrid neural network, linear regression, hybrid linear regression, and hybrid variance ratio methods; and (iv) the power generation was generally under-estimated by the SVR, hybrid SVR, and variance ratio methods.

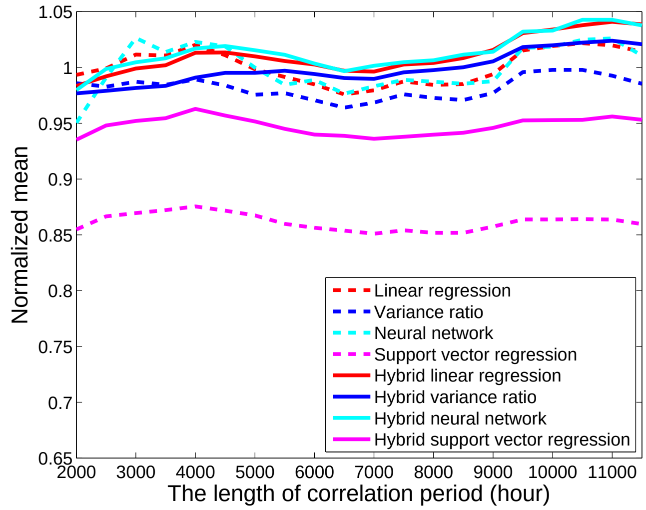
CONCLUSION

This paper developed an advanced hybrid MCP methodology that accounts for the variations of both wind speed and direction. The advanced hybrid MCP method uses the recorded data of multiple reference stations to estimate the long-term wind condition at a target plant site. Two scenarios were analyzed using the hybrid MCP methodology, and interesting results were observed and discussed.

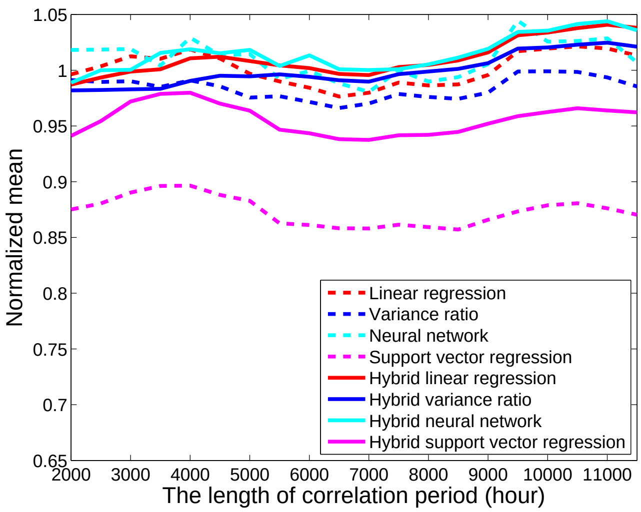
Because each reference station has the flexibility to use any of the available MCP techniques, the multiple reference weather



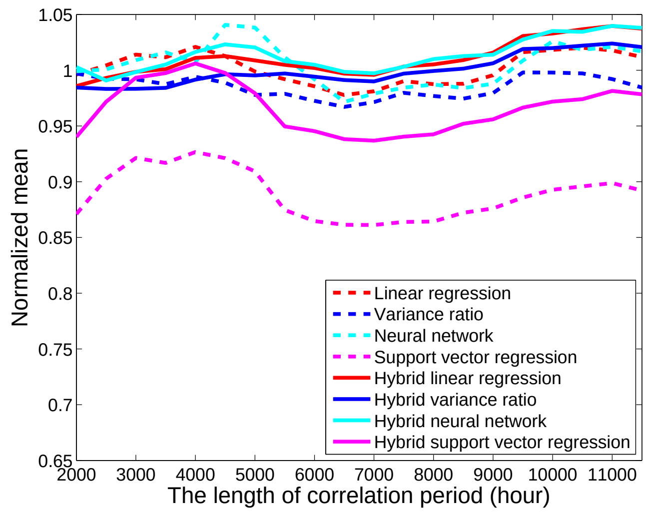
(a) 4 direction sectors



(b) 8 direction sectors



(c) 16 direction sectors



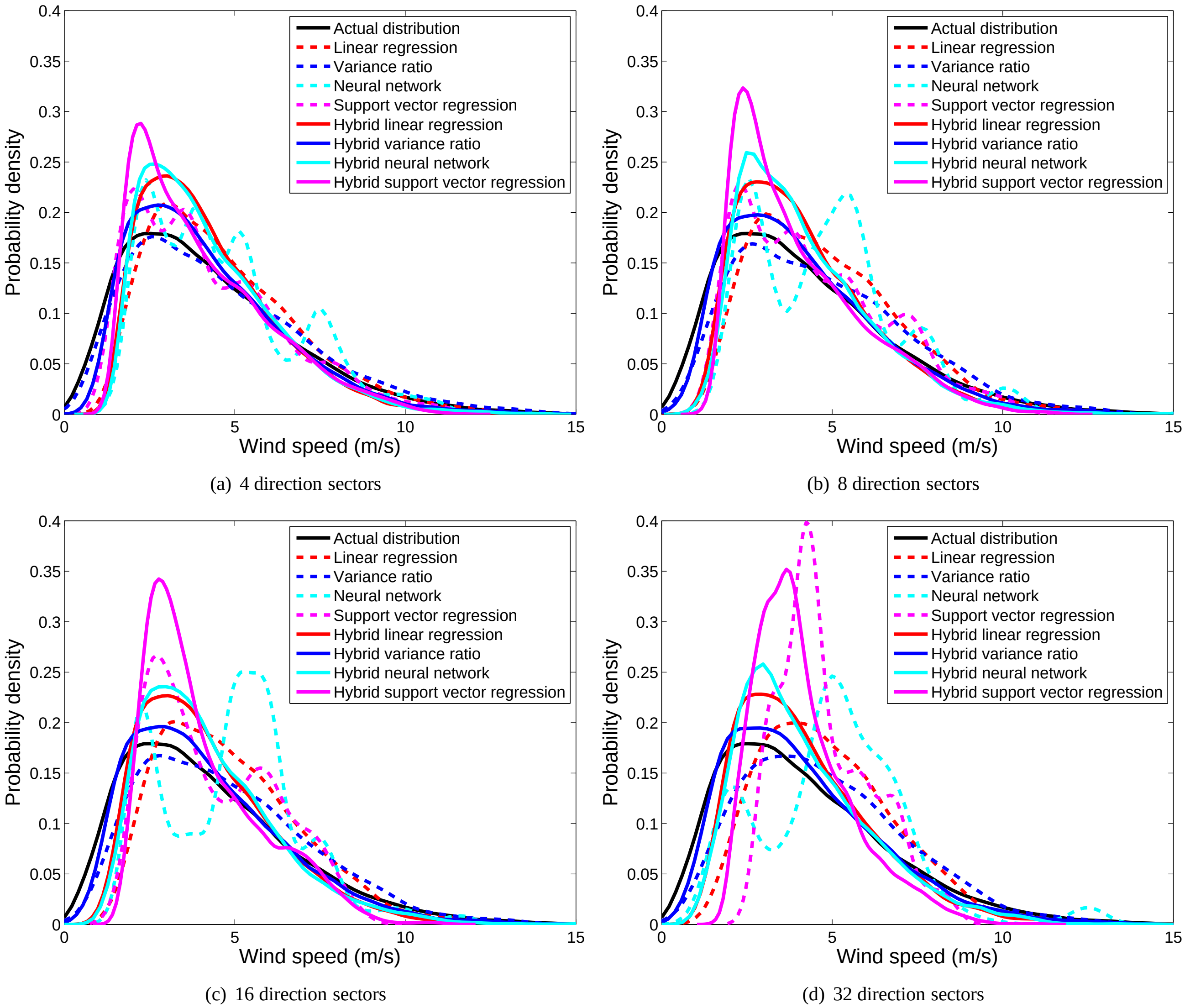
(d) 32 direction sectors

Figure 6. THE RATIO OF MEAN WIND SPEEDS WITH DIFFERENT DIRECTION SECTORS

stations were combined into the hybrid MCP strategy with the best suitable MCP algorithm for each reference station. In the first scenario, each reference weather station used one of the following MCP algorithms: (i) linear regression; (ii) variance ratio; (iii) neural network; and (iv) support vector regression. Therefore, a total of 1,024 (which is equal to 4^5) combinations were investigated to formulate the hybrid MCP strategy. The best hybrid MCP strategy of MCP algorithms and reference station combination was determined and analyzed. We found that the accuracy of the hybrid MCP method was highly sensitive to the combination of individual MCP algorithms and reference stations. We also found that the best hybrid MCP strategy varied based on the length of the correlation period.

In the second scenario, both wind speed and direction were considered in the application of the hybrid MCP strategy. We found that the hybrid MCP methodology performed best when the correlation period was between 5,500-8,500 hours (approximately 8 to 12 months) based on the ratio of mean wind speeds. For the nine-turbine wind plant, the power generation was generally overestimated by the neural network, hybrid neural network, linear regression, hybrid linear regression, and hybrid variance ratio methods; it was under estimated by the support vector regression, hybrid support vector regression, and variance ratio methods.

Quantifying and modeling the uncertainty in the MCP methods would provide more credibility to wind resource assessment



(a) 4 direction sectors

(b) 8 direction sectors

(c) 16 direction sectors

(d) 32 direction sectors

Figure 7. WIND DISTRIBUTION METRICS WITH DIFFERENT DIRECTION SECTORS

and wind plant performance estimation. Modeling the propagation of uncertainty through the MCP process would allow quantification of the expected uncertainty in on-site wind conditions and wind plant power generation. In addition, an investigation of how the uncertainties in the annual distribution of wind conditions interact with the uncertainties inherent in the MCP correlation methodology is also necessary. This investigation is an important topic for future research.

ACKNOWLEDGMENT

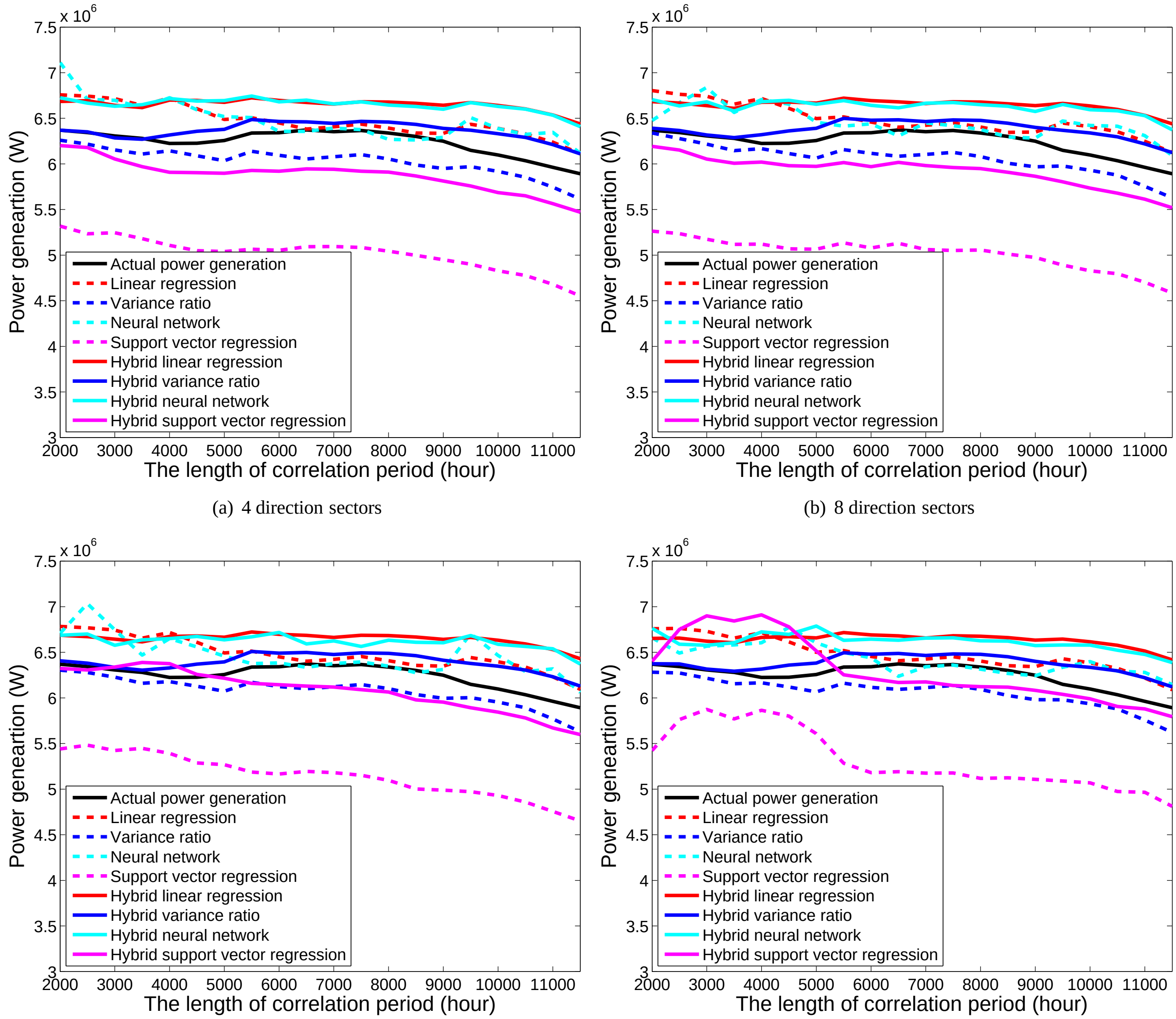
Support from the National Science Foundation Awards CMMI-1100948 and CMMI-0946765 is gratefully acknowledged. Any opinions, findings, conclusions, or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the NSF. This work was also supported by the U.S. Department of Energy under Contract No. DE-AC36-08-GO28308 with NREL.

REFERENCES

[1] EWEA, 2011. Wind energy - the facts. http://www.wind-energy-the-facts.org.

[2] Sheppard, C. J. R., 2009. "Analysis of the measure-correlate-predict methodology for wind resource assessment". Master's thesis, Environmental Resources Engineering, Humboldt State University, Arcata, CA.

[3] Velázquez, S., Carta, J. A., and Matías, J. M., 2011. "Com-



(a) 4 direction sectors

(b) 8 direction sectors

(c) 16 direction sectors

(d) 32 direction sectors

Figure 8. POWER GENERATION METRICS TO EVALUATE THE HYBRID MCP WITH DIFFERENT DIRECTION SECTORS (*GE 1.5MW-XLE*)

parison between anns and linear mcp algorithms in the long-term estimation of the cost per kwh produced by a wind turbine at a candidate site: A case study in the canary islands". *Applied Energy*, ***88***(11), pp. 3689–3881.

[4] Perea, A. R., Amezcua, J., and Probst, O., 2011. "Validation of three new measure-correlate-predict models for the long-term prospection of the wind resource". *Journal of Renewable and Sustainable Energy*, ***3***(2), p. 023105.

[5] Carta, J. A., and Velázquez, S., 2011. "A new probabilistic method to estimate the long-term wind speed characteristics at a potential wind energy conversion site". *Energy*, ***36***(5), pp. 2671–2685.

[6] Mohandes, M. A., Rehman, S., and Halawani, T. O., 1998. "A neural networks approach for wind speed prediction". *Renewable Energy*, ***13***(3), pp. 345–354.

[7] Mohandes, M. A., Halawani, T. O., Rehman, S., and Hussain, A. A., 2004. "Support vector machines for wind speed prediction". *Renewable Energy*, ***29***(6), pp. 939–947.

[8] Zhao, P., Xia, J., Dai, Y., and He, J., 2010. "Wind speed prediction using support vector regression". In The 5th IEEE Conference on Industrial Electronics and Applications (ICIEA), IEEE.

[9] Thøgersen, M. L., Motta, M., Sørensen, T., and Nielsen, P., 2007. "Measure-correlate-predict methods: Case studies

and software implementation". In European Wind Energy Conference & Exhibition, EWEA.
[10] Koeppl, G. W., 1982. *Putnam's Power from The Wind*, 2 ed. Van Nostrand Reinhold Company, New York.
[11] Justus, C. G., Mani, K., and Mikhail, A. S., 1979. "Interannual and month-to-month variations of wind speed". *Journal of Applied Meteorology,* ***18***(7), pp. 913–920.
[12] Derrick, A., 1992. "Development of the measure-correlate-predict strategy for site assessment". In Proceedings of the 14th British Wind Energy Association, BWEA.
[13] Rogers, A. L., Rogers, J. W., and Manwell, J. F., 2005. "Comparison of the performance of four measure-correlate-predict algorithms". *Journal of Wind Engineering and Industrial Aerodynamics,* ***93***(3), pp. 243–264.
[14] Kwon, S., 2010. "Uncertainty analysis of wind energy potential assessment". *Applied Energy,* ***87***(3), pp. 856–865.
[15] Lackner, M. A., Rogers, A. L., and Manwell, J. F., 2008. "Uncertainty analysis in mcp-based wind resource assessment and energy production estimation". *Journal of Solar Energy Engineering,* ***130***(3), p. 031006.
[16] Messac, A., Chowdhury, S., and Zhang, J., 2012. "Characterizing and mitigating the wind resource-based uncertainty in farm performance". *Journal of Turbulence,* ***13***(13), pp. 1–26.
[17] Zhang, J., Chowdhury, S., Messac, A., and Castillo, L., 2012. "A hybrid measure-correlate-predict method for wind resource assessment". In ASME 2012 6th International Conference on Energy Sustainability, ASME.
[18] Zhang, J., Chowdhury, S., and Messac, A., 2012. "An adaptive hybrid surrogate model". *Structural and Multidisciplinary Optimization,* ***46***(2), pp. 223–238.
[19] Zhang, J., Chowdhury, S., Messac, A., and Castillo, L., 2013. "A multivariate and multimodal wind distribution model". *Renewable Energy,* ***51***, pp. 436–447.
[20] Zhang, J., Chowdhury, S., Messac, A., and Castillo, L., 2011. "Multivariate and multimodal wind distribution model based on kernel density estimation". In ASME 2011 5th International Conference on Energy Sustainability, ASME.
[21] Epanechnikov, V. A., 1969. "Non-parametric estimation of a multivariate probability density". *Theory of Probability and its Applications,* ***14***, pp. 153–158.
[22] Duong, T., and Hazelton, M. L., 2003. "Plug-in bandwidth matrices for bivariate kernel density estimation". *Nonparametric Statistics,* ***15***(1), pp. 17–30.
[23] Chowdhury, S., Zhang, J., Messac, A., and Castillo, L., 2012. "Unrestricted wind farm layout optimization (uwflo): Investigating key factors influencing the maximum power generation". *Renewable Energy,* ***38***(1), pp. 16–30.
[24] Chowdhury, S., Zhang, J., Messac, A., and Castillo, L., 2013. "Optimizing the arrangement and the selection of turbines for wind farms subject to varying wind conditions". *Renewable Energy,* ***52***, pp. 273–282.
[25] GE, 2010. *GE Energy 1.5MW Wind Turbine Brochure*. General Electric, http://www.gepower.com/.
[26] GE, 2010. *GE Energy 2.5MW Wind Turbine Brochure*. General Electric, http://www.gepower.com/.
[27] NDSU, 2012. The north dakota agricultural weather network. http://ndawn.ndsu.nodak.edu.
[28] Bi, J., and Bennett, K. P., 2003. "A geometric approach to support vector regression". *Neurocomputing,* ***55***(1-2), pp. 79–108.
[29] Smola, A. J., and Scholkopf, B., 2004. "A tutorial on support vector regression". *Statistics and Computing,* ***14***(3), pp. 199–222.
[30] Clarke, S. M., Griebsch, J. H., and Simpson, T. W., 2005. "Analysis of support vector regression for approximation of complex engineering analyses". *Journal of Mechanical Design,* ***127***(6), pp. 1077–1087.
[31] Basudhar, A., and Missoum, S., 2008. "Adaptive explicit decision functions for probabilistic design and optimization using support vector machines". *Computers and Structures,* ***86***(19-20), pp. 1904–1917.
[32] Yun, Y., Yoon, M., and Nakayama, H., 2009. "Multi-objective optimization based on meta-modeling by using support vector regression". *Optimization and Engineering,* ***10***(2), pp. 167–181.
[33] Zhang, J., 2012. "Hybrid and uncertainty-based surrogate modeling with applications to wind energy". PhD thesis, Mechanical Engineering, Rensselaer Polytechnic Institute, Troy, NY, August.
[34] Beale, M. H., Hagan, M. T., and Demuh, H. B., 2011. Neural network toolbox: User's guide. Tech. rep., Mathworks.
[35] Chang, C. C., and Lin, C. J., 2011. "Libsvm: A library for support vector machines". *ACM Transactions on Intelligent Systems and Technology,* ***2***(3), pp. 27:1–27:27.