

# Multivariate Predictive Analytics of Wind Power Data for Robust Control of Energy Storage

Hamed Valizadeh Haghi, *Member, IEEE*, Saeed Lotfifard, *Member, IEEE*, and Zhihua Qu, *Fellow, IEEE*

**Abstract**—Short-term forecasting is frequently identified as an important tool for the effective management of wind generation. However, forecasting errors, inherent to the point forecasts, increase requirements for energy storage and can affect optimal system operation. Probabilistic forecasts can help tackle this issue by providing a proper characterization of forecasting errors in the optimization process. This paper proposes a multivariate model of forecasting data for wind generation. Predictive uncertainty intervals of wind power can be obtained by sampling from the proposed model. The main goal is to use empirical data models without linear or Gaussian approximations of the distributional or temporal variations. The predictive modeling is utilized within a case study of an energy storage system. A modified robust convex programming is used to maintain the practical robustness and feasibility of the solution based on the sampled scenarios from the model.

**Index Terms**—Data analytics, energy storage, forecasting, microgrid, prediction intervals (PIs), predictive ensembles, robust optimization, smart grid, wind power.

## I. INTRODUCTION

A BETTER integration of wind power forecasts in operational management tools can facilitate introduction of wind power in microgrids and smart grids [1]. A modern wind power forecasting tool may use numerical weather prediction (NWP) models, supervisory control and data acquisition (SCADA) data and other information about the characteristics of wind power plants [2], [3]. Some of the models that are referred to as the model-driven approaches, can perform very well in long-term forecasts [4], [5]. However, these approaches need large meteorological data, detailed terrain characteristics, and configuration of wind turbines in order to build an accurate model at the first place which makes them impractical for short-term applications.

Manuscript received September 20, 2015; revised February 05, 2016; accepted March 31, 2016. Date of publication May 17, 2016; date of current version August 04, 2016. This work is supported in part by the U.S. National Science Foundation under Grant ECCS-1308928, by the U.S. Department of Energy under Award DE-EE0006340 and Award DE-EE0007327, by the U.S. Department of Transportation under Grant DTRT13-G-UTC51, by Leidos under Contract P010161530, and by Texas Instruments awards. Paper no. TII-15-1440.R1.

H. Valizadeh Haghi and Z. Qu are with the Department of Electrical Engineering and Computer Sciences, University of Central Florida, Orlando, FL 32816 USA (e-mail: hvh@ucf.edu; qu@ucf.edu).

S. Lotfifard is with the School of Electrical Engineering and Computer Science, Washington State University, Pullman, WA 99164 USA (e-mail: s.lotfifard@wsu.edu).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TII.2016.2569531

Another category of forecasting tools for short-term wind power is based on data-driven or artificial intelligence approaches [6]–[9]. Model training in these approaches uses historical data of wind speed and possibly other NWP-related data such as wind direction. Well-known methods include persistence method [8], auto regressive-type models [10], and the methods that use stochastic processes [11], artificial neural networks (ANNs) [12], [13], and fuzzy systems [14]. The methods in this category can be used for different turbines, diverse geographical areas, and different wind locations. The generated forecast is the average wind power expected to be available from the wind farm during the considered look-ahead time. This is also referred to as the point forecast that provides only a single value for the considered time index.

Extensive research has been reported on point forecasting methods and their performance which use different types of regression models as mentioned above [1], [4]. These approaches use deterministic models and therefore cannot guarantee a robust estimation of future wind power uncertainty. In order to tackle forecast uncertainty, some propose a combination of model-driven and data-driven solutions [15].

In order to complement above-mentioned point forecasts which only provide projected mean values of the wind power (conditional on the previous observations), probabilistic forecasts, on the other hand, offer full distributions of wind power for the considered look-ahead times. In power system operations, it is desirable to predict exact probabilities associated with different wind power levels for future observations. This is particularly useful when a certain degree of robustness is sought for the reliable operation of storage against uncertainty of point forecasts.

Probabilistic forecasting may refer to a wide range of methods that study different models for characterization or online estimation of uncertainty. The uncertainty is reflected either in the point forecasts or in the wind power variations itself. The existing literature may also express the forecasting uncertainty and prediction intervals (PIs) in the form of a set of quantiles or intervals, mean, variance, or probability density functions (pdfs) [16]–[22]. Among these representations, pdfs are the most generic format for calculations yet the most commonly used representation in final decision making is by quantiles [23]. Interval evaluations of power system variables benefit from wind power PIs as recently reported for power flow analysis and locational marginal pricing [24], [25].

A set of quantiles represents the empirical cumulative distribution function (cdf) that can be calculated from a full pdf by using kernel smoothing [26]–[28]. This pdf estimates the true underlying forecasting error distribution of wind power and can

be provided by either a parametric or a nonparametric representation. Nonparametric or empirical representation is the most natural and accurate way of modeling the PIs; because it can reflect how the real wind and error data change and can be calculated from historical forecasting data [22]. The disadvantage of these quantile methods is that modeling requires building specific training sets; for example, in order to calculate with respect to a quantile curve, training and modeling for each quantile is inevitable. This increases the computational burden and chances of inaccuracies due to outliers such as quantile crossing.

The temporal dependence structure of the forecasting error series (or PIs) has been modeled by the covariance matrix [20]. Temporal dependence is studied as the correlation between values of the data at different points in time or across time. Effects of the dependence between adjacent wind farms, i.e., spatial dependence, has been studied [29]. Also, conditional PIs expressed by fuzzy classification has been proposed to account for the effects of wind power levels on the forecasting error [18].

On the other hand, existing probabilistic forecasts that are developed for consecutive look-ahead times, do not consider cross correlations between wind power and error time series. Error time series does not exhibit any particular data structure as they are intrinsic to the forecasting method. In fact, the cross correlations and dependences between the forecasting error series over the considered time horizon are valuable characteristics, particularly, for power system problems that have time-dependent memory in applications such as unit commitment or energy storage management [1].

This paper proposes a multivariate time series modeling of wind power PIs and uncertainty ensembles in order to capture the following nonlinear nonparametric characteristics all together.

- 1) Temporal dependence of forecasting error time series that is the dependence between forecasting error distributions at each time index.
- 2) Temporal dependence of wind power time series.
- 3) Interdependence or cross correlation between wind power time series and the corresponding error series.
- 4) Empirical distributions of forecasting error and wind power data at each time index as well as the considered time horizon. The calculated final distributions, ensembles, or PIs would be conditional on previous observations.

The proposed modeling approach is nonparametric in the sense that it does not rely on data belonging to any particular distribution. Also, it is nonlinear as it quantifies nonlinear dependence. The proposed algorithm uses the copula-based modeling. Once the model is fitted to the historical data, the conditional forecast error ensembles at each look-ahead time or PIs can be calculated very fast. The modeling is handled offline while the conditional ensembles or PIs are calculated in real time. This can be viewed as an adaptive probabilistic forecasting for different look-ahead time periods.

Optimization of power systems and energy storage calculations relies on the accuracy and reliability of probabilistic forecasts which are the main focus of this paper. However, it is also important how to solve for an optimal strategy. Robust optimization in contrast to stochastic optimization is selected in this paper. The motivation is twofold. First, the model of uncertainty in robust optimization is set-based. This is an appropriate

notion of parameter uncertainty for the presented application as the conditional forecast ensembles or PIs are calculated in a set by sampling from the proposed model. Second, computational tractability should be maintained for different objective functions and constraints. Here, the operator constructs a solution that is feasible for all realizations of the forecast in a given set.

An extensive review of different approaches to robust optimization is presented by [36]. A worst-case solution using the lower and upper bounds of the forecasts might not be desirable for most of the problems, since it considers all possible uncertainty scenarios, including those that are particularly unlikely to happen. Solutions to optimization problems can have significant sensitivity to future changes in the parameters of the problem. This can render a computed solution suboptimal or infeasible. In order to deal with this issue, robust programming paradigm covers a range of methods [30] that see a decision environment as characterized by a knowledge of the data and that it belongs to a given uncertainty set, feasibility of the solution meaning that it satisfies all viable realizations of the constraints, and computational tractability of the robust formulation counterpart to the deterministic problem.

Several optimization methods can be used to obtain robust optimal solutions based on these characteristics [30]–[35]. Construction of an uncertainty set from data models (step 1 above) can significantly affect the final tradeoff solution. Some of the existing methods are purely deterministic based on a set of parameters consistent with data, whereas other methods allow adaptive adjustments [36]–[38]. As a recently emerged approach [39], uncertain parameters (affected by the forecasting error) are modeled as empirical random scenarios defined on an uncertainty set; then, the decision-making program optimizes the system with observations of the uncertainty within that set. Successful implementation of this approach in robust control theory is first reported by Calafiore and Campi [39] and recently extended to nonconvex nonlinear cases [40]. The main idea is based on a probabilistic version of the well-known robust convex programming (RCP) [32]. The approach adopted in this paper has the following features.

- 1) A precalculated number of samples from the uncertainty model is used to modify the corresponding standard problem. The new robust problem is referred to as the aggregated robust program (ARP).
- 2) There is a probabilistic guarantee of robust stability and constraint satisfaction (as will be shown in Section III).
- 3) Only constraints in the ARP require convexity with respect to the design variable, while generic nonlinear functions with respect to the uncertain variables are allowed. This is not a strict requirement but only ensures feasibility over the unseen constraints [39].
- 4) The ARP is solved in a single run thus avoiding computational burden compared to the traditional scenario-based stochastic programming techniques.
- 5) Any deterministic optimization method can be used to solve the ARP.

The main objective of the ARP's robustness is to provide an explicit bound on the probability of possible violation of the constraints. This approach can effectively apply the proposed uncertainty model in this paper and is outlined in Section III.

Then, an alternative energy storage-based microgrid is studied in order to illustrate the effectiveness of the proposed forecasting error characterization. The main contributions of this paper that can help implement robust operations of wind power capacity are as follows.

- 1) Develop a probabilistic model of the forecasting error data that can utilize empirical distributions, as well as nonlinear temporal dependence of wind power.
- 2) Incorporate the cross correlations between wind power data and error data over the consecutive forecasting time horizons. The output from the forecasting model is conditional ensembles or PIs of wind generation for different look-ahead times.
- 3) Use a robust optimization of energy storage, as an alternative to the Monte Carlo-based algorithms, that can use the output from the proposed probabilistic forecast.

In the following sections, the proposed model for calculating the conditional PIs of wind power is described. Then, the RCP and ARP are explained in order to illustrate how to use the proposed probabilistic forecasts for optimizing a wind-based power system. The case study provides a solution to the management of the available energy storage that should be robust against the characteristic uncertainty of the wind power forecasts.

## II. MULTIVARIATE APPROACH TO CALCULATING CONDITIONAL PIS

PIs can be used in predictive inference procedures in order to deal with the uncertainty of point forecasts [41]. Future observations of wind power, according to the estimated probability, will fall in PIs, given what has already been observed. The main objective is to provide an empirical distribution of error for a current forecast value; hence, the inherent uncertainty of point forecasts can be predicted.

Most methods to construct PIs involve running a point forecast program (e.g., the regression neural network or the ANN) on a set of historical data [42]. A wind power forecast is obtained at each time instant within prespecified periods as a sliding window covering a certain number of hours. The forecast program predicts the wind power  $\tau$  hours later based on the past hourly wind power values as far as  $R$  number of hours. This procedure should be carried out using a comprehensive wind power dataset that is the training matrix of the predictor. Hence, a complete set of forecasts is obtained by training the neural network using  $Q$  vectors of length  $R$  as input for which the targets of interest were the recorded wind power  $\tau$  hours later.

For example, Fig. 1 illustrates the input and target for  $Q = 3$ ,  $\tau = 5$  and  $R = 4$ . Then, the forecasts are subtracted from the recorded target values in each period. This provides an error series along with the historical data. From this step forward, the probabilistic characterization of the point forecast uncertainty can be formulated as follows.

If wind power time series is of length  $\Gamma$ , the error time series would be of length  $(\Gamma - \tau - R)$ , according to Fig. 2. By rearranging the error series and the corresponding observations in the consecutive periods  $(0, T)$ ,  $(T, 2T)$ , ...,  $((\zeta - 1)T, \zeta T)$ , the following matrix can be obtained:

$$M = [M_\omega \ M_\varepsilon] \quad (1)$$

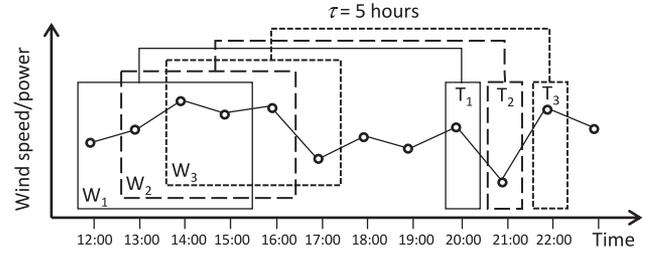


Fig. 1. Regression-based point forecasting for wind power:  $W_1$ ,  $W_2$ ,  $W_3$  are the inputs,  $T_1$ ,  $T_2$ ,  $T_3$  are the targets used for training. Here  $Q = 3$ ,  $R = 4$ , and  $\tau = 5$ .

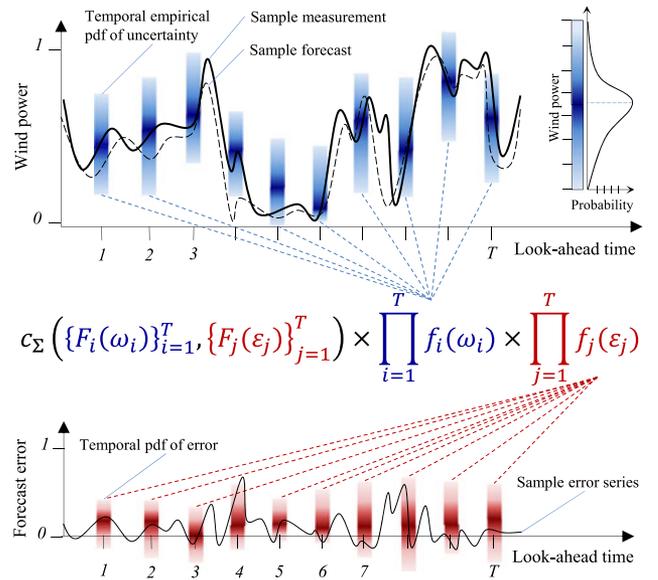


Fig. 2. Structure of a copula-based multivariate modeling for conditional characterization of forecast errors.

where

$$M_\omega = \begin{bmatrix} \omega_1^1 & \omega_1^2 & \cdots & \omega_1^T \\ \omega_2^1 & \omega_2^2 & \cdots & \omega_2^T \\ \vdots & \vdots & \ddots & \vdots \\ \omega_\zeta^1 & \omega_\zeta^2 & \cdots & \omega_\zeta^T \end{bmatrix} \quad (2)$$

$$M_\varepsilon = \begin{bmatrix} \varepsilon_1^1 & \varepsilon_1^2 & \cdots & \varepsilon_1^T \\ \varepsilon_2^1 & \varepsilon_2^2 & \cdots & \varepsilon_2^T \\ \vdots & \vdots & \ddots & \vdots \\ \varepsilon_\zeta^1 & \varepsilon_\zeta^2 & \cdots & \varepsilon_\zeta^T \end{bmatrix} \quad (3)$$

is the matrix of errors, where  $\omega_t^\zeta$  and  $\varepsilon_t^\zeta$  are the  $\zeta$ th observation of wind power and the corresponding forecast error at time  $t$ , respectively. Each column of  $M_\omega$  and  $M_\varepsilon$  can be used to calculate an empirical (nonparametric) pdf. It should be emphasized that the calculated pdfs are independent at this stage. Hence, it cannot be used as a reliable model unless the actual dependence structure and its parameters (the relationships between pdfs) are

added. Fig. 2 shows an outline of such a multivariate modeling using the concept of copulas. Color bars show the empirical marginal pdfs  $f_i(\omega_i)$  and  $f_i(\varepsilon_i)$  that are directly calculated from data in (1) and connected to each other using a copula function. This offers a unified multivariate model of historical forecasting data that includes exact nonlinear dependence structure required to estimate PIs.

A copula function can be simply defined as the joint distribution of two or more random vectors each transformed as uniform random variables. Advantages of using copulas for multivariate data are considerable [43]. First, by using copula modeling concept, the joint distribution can be decomposed into the dependency structure (copula function) and the marginal distributions. The marginal distribution of a subset of a collection of random variables is defined as the probability distribution of the variables contained in the subset. Hence, each variable can be described by a different distribution (e.g., Gaussian, Weibull, etc.). Furthermore, copula functions capture the complete dependence structure that is unlike the linear correlation coefficient which measures co-variations up to the second order.

Let  $G$  be an  $n$ -variate distribution function with marginal distributions  $F_1, \dots, F_n$ . There exists an  $n$ -dimensional copula  $C$  such that for all  $x$  in  $\mathcal{R}^n$

$$G(x_1, \dots, x_n) = C(F_1(x_1), \dots, F_n(x_n)). \quad (4)$$

Copula  $C$  is unique for all continuous marginal distributions. On the other hand, if  $C$  is an  $n$ -dimensional copula and  $F_i$  are cdfs, then the function  $G$  according to (4) is an  $n$ -variate joint distribution of random variable  $x$ .

Two main types of copulas exist: 1) elliptical; and 2) Archimedean. Full description of these types and their properties is out of the scope of this paper (information can be found in [43]). However, it should be noted that the Archimedean copulas cannot be easily implemented in higher than two dimensions and the practical option here is elliptical copulas. The Gaussian copula is employed in this paper as it fits wind data very well [44].

In order to construct the model by fitting a copula, the following procedure should be follows.

- 1) Calculate cdfs of all variables (i.e., marginal cdf).
- 2) Select a copula function. Gaussian copula is preferred here as it provides fast calculations over more than two variables.
- 3) Perform maximum likelihood estimation for calculating parameters of the selected copula considering cdf vectors in step 1. This gives the most fitting positive semidefinite matrix of the rank correlation [43].
- 4) The unified model of the data according to the Skalars theorem [43] can be calculated at this stage using (4).
- 5) Use inverse cdfs to get samples from the model.

### A. Model-Based Conditional Forecast Error

Assuming the finite collection of samples according to (1), the finite dimensional distribution in the cdf form can be written as

$$P(\Psi \leq \omega, E \leq \varepsilon | \omega, \varepsilon)$$

$$\begin{aligned} &= C_{\Sigma} \left( \{F_i(\omega_i)\}_{i=1}^T, \{F_j(\varepsilon_j)\}_{j=1}^T | \omega, \varepsilon \right) \\ &= C_{\Sigma} \left( \{u_i\}_{i=1}^T, \{u_j\}_{j=T+1}^{2T} | \omega, \varepsilon \right) \end{aligned} \quad (5)$$

where  $C_{\Sigma}$  is the copula function,  $F_i(\omega_i)$  and  $F_j(\varepsilon_j)$  are the marginal cdf of the wind power and error series, respectively for the corresponding  $\Psi$  and  $E$  random variables, and  $P$  is the probability.  $\{u_i\}$  is the set of transformed distributions according to [43]. The sequence of  $\{u_i\}$  is the same as the sequence of variables in  $M$  according to (1). The joint density of the data in (1), can be obtained by differentiating (5) with respect to  $(\omega, \varepsilon)$

$$\begin{aligned} P(\omega, \varepsilon | \Psi, E) &= c_{\Sigma} \left( \{F_i(\omega_i)\}_{i=1}^T, \{F_j(\varepsilon_j)\}_{j=1}^T | \omega, \varepsilon \right) \\ &\quad \times \prod_{i=1}^T f_i(\omega_i) \prod_{j=1}^T f_j(\varepsilon_j) \end{aligned} \quad (6)$$

where  $f(\cdot)$  are pdfs and  $c_{\Sigma}(\cdot)$  is the copula density function that by assuming  $n$ th cross-partial derivative of  $C_{\Sigma}$ , can be derived from (here, it is assumed that  $n = 2T$  and the joint density is equal to the product of the marginal densities by the copula density)

$$\begin{aligned} f \left( \{u_i\}_{i=1}^T, \{u_j\}_{j=T+1}^{2T} \right) &= \frac{\partial^{2T}}{\partial u_1 \partial u_2 \dots \partial u_{2T}} \\ &\quad \times F \left( \{u_i\}_{i=1}^T, \{u_j\}_{j=T+1}^{2T} \right) = \prod_{i=1}^T f_i(u_i) \prod_{j=T+1}^{2T} f_j(u_j) \\ &\quad \times \frac{\partial^{2T}}{\partial u_1 \partial u_2 \dots \partial u_{2T}} C_{\Sigma} \left( \{F_i(u_i)\}_{i=1}^T, \{F_j(u_j)\}_{j=T+1}^{2T} \right) \\ &= \prod_{i=1}^T f_i(\omega_i) \prod_{j=1}^T f_j(\varepsilon_j) c_{\Sigma} \left( \{F_i(\omega_i)\}_{i=1}^T, \{F_j(\varepsilon_j)\}_{j=1}^T \right). \end{aligned} \quad (7)$$

Once the multivariate model is calculated, the Monte Carlo method can be used to calculate the conditional PIs. This method can be carried out by constructing a comprehensive dataset. This dataset can then be used for calculating and updating the conditional PIs. The algorithm is as follows.

- 1) Construct a comprehensive look-up dataset that includes a large number of samples (e.g., 500 000) drawn from the multivariate model of forecasting data in (6). As the distributions are nonparametric, a kernel smoothing technique is recommended. The minimum number of samples to ensure a successful implementation can be determined by experiment or according to the Appendix.
- 2) Calculate conditional pdf of interest simply by searching for the realized matching past values of wind power and forecast errors over the forecasting interval  $T$  using the look-up dataset in step 1. This simple procedure is justified based on the following: Assume  $X_i$ 's are the random variables representing PIs

$$f(x_1, \dots, x_n) = f_1(x_1) \times f_2(x_2 | x_1)$$

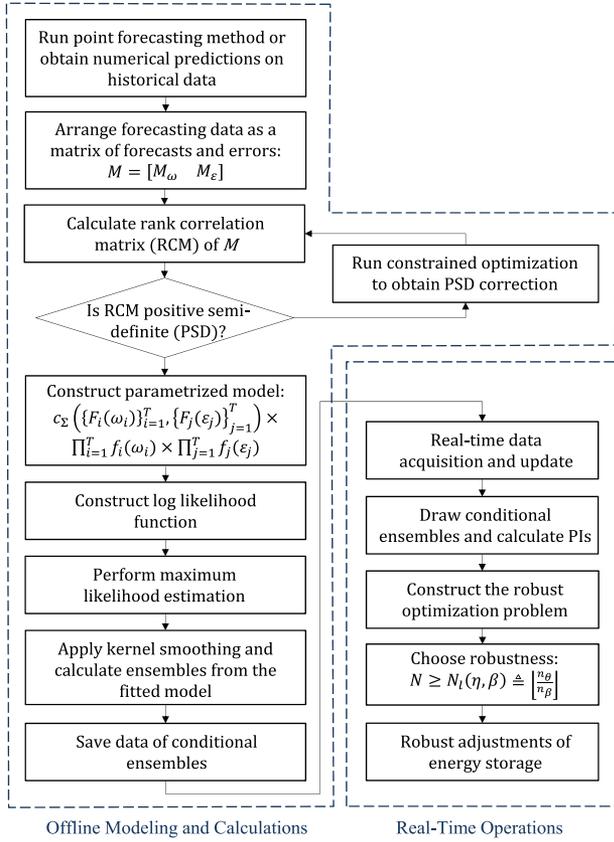


Fig. 3. Stages of the proposed method: offline modeling and characterization along with real-time calculation and robust optimization.

$$\begin{aligned} & \times f_3(x_3 | x_1, x_2) \times \\ & \times \cdots \times f_n(x_n^* | x_1, \dots, x_{n-1}) \end{aligned} \quad (8)$$

where  $f_n(x_n^* | x_1, \dots, x_{n-1})$ , as the only parameter to be determined, represents conditional error distribution of the PI of interest given that  $X_1 = x_1, \dots, X_{n-1} = x_{n-1}$ . It should be mentioned that a small tolerance/threshold should be considered when searching for the matching power and error values. This is due to the fact that the number of samples in step 1 is finite.

A detailed flowchart of the proposed method is shown in Fig. 3. The flowchart illustrates how the ensembles (equivalently, scenarios or data samples) can be generated for representing possible realizations of uncertain forecasts under study. These ensembles are then used as inputs to the optimization problem and should exploit critical characteristics leading to a robust decision policy. Three robust optimization stages are described in Section III.

It should be noted that the modeling and most of the required analytics can be handled offline. Hence, real-time or online operations would be more efficient. This is in addition to the increase in computational efficiency because of using a single-run robust optimization method which is described in Section III. The proposed forecasting and optimization approach, in general, can be viewed as an adaptive algorithm

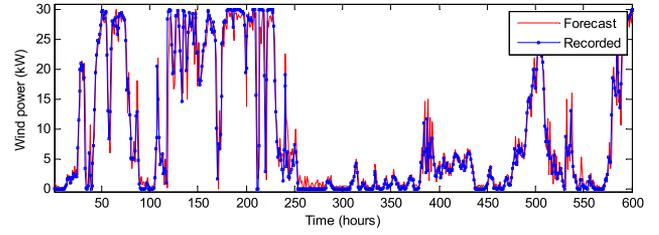


Fig. 4. Example of the point forecasting data from NREL [48]. Wind data along with the error data are used in constructing the proposed model.

that changes its behavior based on resources available and the history of data recently received.

### B. Application to Empirical Data, Correlation Analysis, and Verification

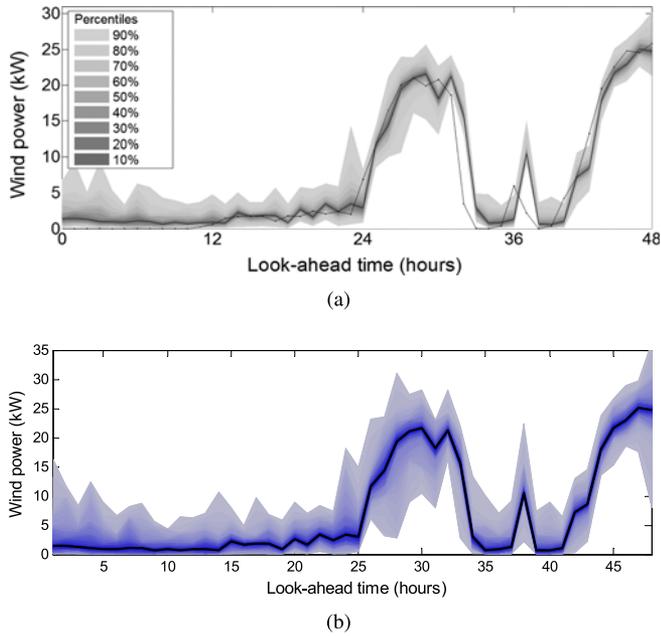
The suggested algorithm is implemented using the data recorded in California. Fig. 4 shows a typical window of recorded wind power along with the 4 h-ahead forecasts selected from the National Renewable Energy Library's (NREL) western wind dataset for year 2005, site no. 4492 [48]. The variable hour-ahead point forecast data is obtained by using a high-order multilayer neural network [49]. It should be emphasized that the proposed approach can be applied regardless of the type of the method used to obtain the point forecasts. Both forecasts and recorded data have an hourly time resolution. Forecasting time horizon can be varied from 1 to 24 h ahead based on the selected point forecast method and is updated every hour.

Parameters of the model using the aforementioned dataset are  $t = [1, 24]$ ,  $T = 24$  h, and  $\zeta = 365$ , as the modeling matrices  $M$  is of order  $365 \times 48$ , according to (1) and (2). Fig. 5(a) and (b) illustrates 48-h of 4 h-ahead forecasts using the proposed algorithm. Fig. 5(a) is a fan chart that shows ranges for likely values of forecasted data alongside a line showing the most likely value or a central estimate. As the predicted PIs are asymmetric (since the proposed model uses empirical densities), the fan chart is centered at the more likely forecast, i.e., the mode.

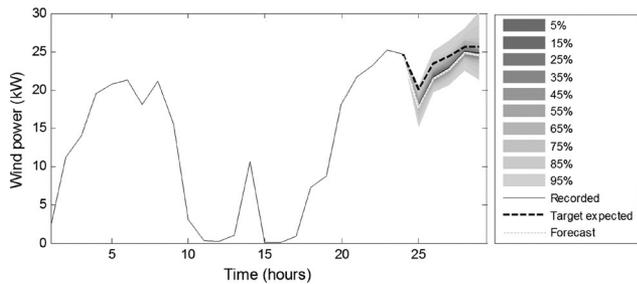
At each look-ahead time (e.g., every 4 h in Fig. 4), these prediction ranges or PIs spread out as forecasts become increasingly uncertain with time until the next look-ahead time. At this point, the algorithm is updated with the most recent outcomes and the copula-based model is simulated again in order to calculate conditional PIs according to the previous section. This is shown in Fig. 6, only for a single 4 h-ahead forecasting.

Performance of the proposed approach is compared to four other forecasting methods including the persistence method, the autoregressive integrated moving average (ARIMA), exponential smoothing method (ESM), and the Gaussian model using the same training and verification data for benchmarking. The Gaussian PIs are developed using samples from a normal distribution, whose parameters are calculated with the same procedure as the proposed method. The Gaussian PIs are also referred to as predictive confidence intervals [41].

The persistence method and ESM are known to be difficult to outperform for short-term forecasting, hence they are used



**Fig. 5.** Example of the conditional PIs calculated by the proposed algorithm (a) using a fan chart showing out-of-sample forecasts and (b) by showing all percentiles at a single-increment resolution with 99% percent confidence.

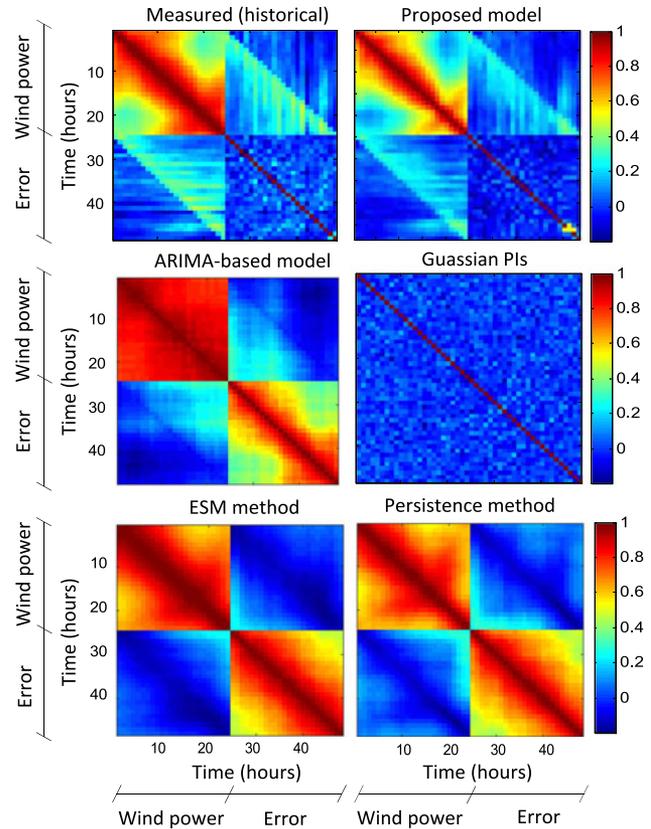


**Fig. 6.** Calculated conditional PIs using the proposed copula-based multivariate model based on newly updated measurements of wind power.

as benchmark [18]. The forecast error for the persistence point forecasting method is assumed to be normally distributed with the mean and variance given by the last observations. The ESM’s error is assumed to be normal as well with its mean conditional on smoothed values of past observations and its variance conditional on smoothed value of past squared residuals [45]. A high-order ARIMA model is also considered. The ARIMA model is adopted from [46] which provides a more advanced benchmark to better demonstrate the effectiveness of the proposed model.

Fig. 7 shows the actual dependence structure of the wind power together with the forecast error, compared to the simulated data using five models. As shown in Fig. 7, the proposed algorithm captures the actual structure of the temporal dependence. The spatial dependence between wind power levels and forecasting error is also well captured. It is interesting to note the following characteristics.

- 1) The temporal dependence structure of wind power time series is nonlinear. The specific multivariate modeling



**Fig. 7.** Multivariate temporal dependence structure of wind power PIs alongside forecast error series for the whole year. Color map shows Pearson’s correlation coefficient.

(the Gaussian copula that is assigned to  $C_{\Sigma}$  in ((5)) used to capture nonlinear temporal dependence outperforms autoregressive and Markov models [44].

- 2) There is a one-sided asymmetric dependence structure between wind power levels and the corresponding forecast error; that is, the fact that the forecast error at each time is correlated with the historical data of wind power levels used to make the forecast. Representing this kind of asymmetric dependence can only be made possible by using copulas.
- 3) As an inherent essential characteristic of the point forecasting algorithm, the forecast error has no specific structure of temporal dependence. This is misrepresented by the ESM, ARIMA, and persistence methods.

Fig. 8 illustrates application of the modeling strategy proposed by Fig. 2 from which the PIs in Fig. 5 are calculated. The PIs depicted in Fig. 8 shows that the proposed approach successfully uses empirical distributions. In other words, this is the first probabilistic forecast of wind power that provides nonparametric empirical distributions as PIs. This is particularly useful for more accurate representation of uncertainty in near future. It should be reemphasized that the error data from the point forecast engine is considered within the model, according to (5)–(7) which include cross correlations.

Four scores are used to quantify the outcomes of the proposed and benchmark methods. Two deterministic scores include

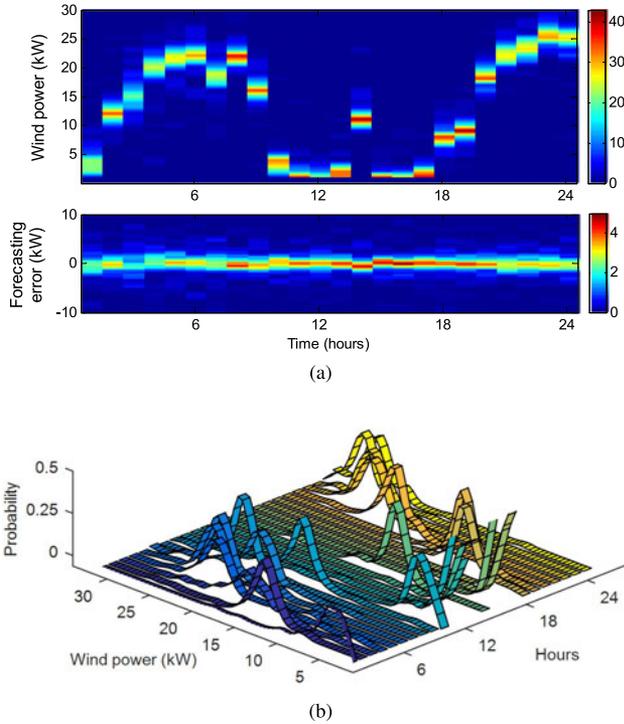


Fig. 8. (a) Complete modeled distributions of wind power PIs and forecast error for the corresponding data in Fig. 5, using second 24-h time interval. (b) The same probabilistic information of (a) in 3-D. This illustration shows the direct application of the proposed modeling concept in Fig. 2 to real data.

root mean square error,  $RMSE = \sqrt{1/T \sum_{t=1}^T (x_t - \hat{x}_t)^2}$ , and mean absolute error,  $MAE = 1/T \sum_{t=1}^T |x_t - \hat{x}_t|$ , where  $\hat{x}_t$  is the predicted value of  $x_t$ .

The continuous rank probability score (CRPS) and log score are utilized to evaluate the performance of the probabilistic forecasts [47]. The CRPS has become one of the popular and reliable tools for probabilistic forecast evaluations, particularly for ensemble forecasts [46]. For the  $h$ -step ahead probabilistic forecast pdf  $f_{t+h|t}$ , let  $F_{t+h|t}$  be the corresponding cdf. Then, the CRPS is defined as

$$CRPS = \int_0^1 [F_{t+h|t}(x) - \mathbf{1}(x - x_{t+h})]^2 dx \quad (9)$$

where  $\mathbf{1}(x - x_{t+h})$  is the indicator function which is equal to one when  $(x - x_{t+h})$  is positive.

The log score is defined as the mean negative log of the forecast pdf evaluated at the corresponding observation, log score =  $1/T \sum_{t=1}^T -\log(f_{t+h|t})$ .

Forecast performance scores of all methods are listed in Table I, calculated in percent of rated power. These are average values assuming 95% confidence. Table II lists probabilistic scores by calendar months. Both probabilistic and deterministic performance scores show noticeable improvement over the benchmark and state-of-the-art methods.

TABLE I  
FORECASTING PERFORMANCE OF THE PROBABILISTIC METHODS (MEAN OF ENSEMBLE SKILL SCORES AS % OF RATED POWER AND 95% CONFIDENCE INTERVAL)

	RMSE	MAE	CRPS	Log Score
Proposed Ensembles	9.2374	6.0801	6.4001	7.0168
ARIMA-Based PIs	24.8038	18.2355	14.0399	12.0141
Gaussian PIs	15.7000	12.9189	11.9236	13.1523
Persistence	27.1923	17.3575	n/a	n/a
ESM	16.1590	10.3211	n/a	n/a

TABLE II  
PROBABILISTIC FORECAST SKILL SCORES AS % OF RATED POWER AND 95% CONFIDENCE INTERVAL FOR SELECTED SITE AND CALENDAR MONTHS

Month	Score	Proposed	ARIMA	Gaussian
January	CRPS	8.7566	12.1764	10.8639
	Log Score	9.014	10.6834	14.6758
February	CRPS	7.1053	15.7392	12.6834
	Log Score	10.4975	12.9157	12.2844
March	CRPS	6.6007	16.3714	12.1445
	Log Score	9.23658	13.4943	13.211
April	CRPS	5.0046	20.3158	11.9222
	Log Score	7.9884	14.3135	11.3132
May	CRPS	4.6078	15.1173	15.7494
	Log Score	8.1993	11.5927	14.1255
June	CRPS	6.7277	11.7928	11.5692
	Log Score	7.0768	10.8676	13.5737
July	CRPS	4.7458	13.4213	13.9934
	Log Score	6.9217	11.1687	16.2326
August	CRPS	5.6498	11.9618	12.9955
	Log Score	7.0642	10.6933	13.5204
September	CRPS	6.2188	13.2992	10.666
	Log Score	7.4220	12.3324	14.1512
October	CRPS	6.2910	13.5862	9.7724
	Log Score	7.4757	11.3956	10.6387
November	CRPS	6.7152	13.2846	9.9455
	Log Score	7.8243	13.3804	10.9508
December	CRPS	8.382	11.4137	10.7784
	Log Score	7.4815	11.3325	13.15129

### III. ROBUST OPTIMIZATION UNDER FORECAST UNCERTAINTY

This section investigates practical effectiveness of the proposed uncertainty model in Section II for robust optimization of energy storage which corresponds to the last three stages in Fig. 3. A case study involving optimal operation of an energy storage-based microgrid is presented. Then, the robust optimization problem is explained in order to compensate for the uncertainty of wind forecasts.

#### A. Case Study: Optimal Operation of Energy Storage in a Microgrid

Fig. 9 shows the architecture of a theoretical microgrid that operates using several distributed micro-sources and a hydrogen-based energy storage system in an islanded situation. Consumption of hydrogen should be controlled in order to provide a higher reliability and to minimize load interruptions during islanded operation.

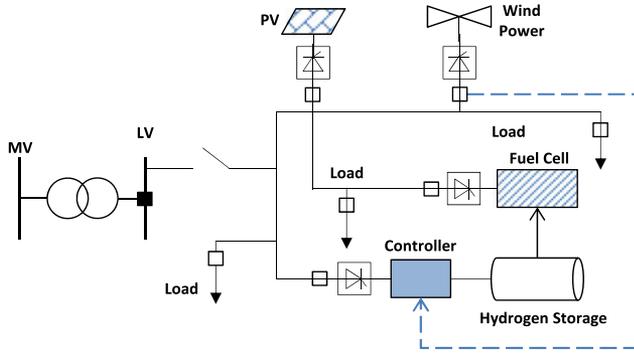


Fig. 9. Diagram of the hybrid system under study. Both the generation side and load side are distributed and based on the short-term wind power forecasts, the only operational control strategy is assumed for the energy storage.

By minimizing the following objective function, the optimal adjustment of energy storage ( $\Delta H$ ) at each look-ahead time horizon can be calculated.

$$\min \sum_n MC_n + LC \quad (10)$$

$$\text{in terms of } \Delta H = H_{t+\tau} - H_t \quad (11)$$

$$\text{subject to } 0 \leq H_t \leq \text{Max} \quad (12)$$

$$P_{\text{wind}} + P_{PV} + P_{FC} = (L - LC)/\eta_C \quad (13)$$

$$\text{Sim}(\Delta H) = 1 \quad (14)$$

where

$MC_n$	Maintenance cost of component $n$ ;
$LC$	load loss due to curtailment;
$\Delta H$	hydrogen storage adjustment at each look-ahead time;
$H_t$	stored hydrogen at time $t$ ;
$H_{t+\tau}$	stored hydrogen at the end of the forecasting horizon;
$P_{\text{wind}}$	forecasted wind power used for supplying load;
$P_{PV}$	preassigned available PV power;
$P_{FC}$	available fuel cell power from $\Delta H$ according to (15);
$L$	total load;
$\eta_C$	converters efficiency.

Hydrogen is utilized through a proton exchange membrane fuel cell. The equivalent heating value of hydrogen is  $3.4 \text{ kWh/m}^3$  and its density is about  $0.09 \text{ kg/m}^3$ . Therefore, the amount of energy yield per kilogram of hydrogen is  $37.8 \text{ kWh/kg}$  and the electricity produced by the fuel cell can be calculated as  $P_{FC} [\text{kWh}] = \Delta H [\text{kg}] \times \eta_{FC} \times 37.8$ , where  $\eta_{FC}$  is the efficiency of the fuel cell. Hydrogen is produced from the electrolysis of water that happens within the controller in Fig. 9. The amount of hydrogen storage is proportional to the available electricity from the wind generation as follows:

$$\Delta H [\text{kg}] = \frac{P_{w2h} [\text{kWh}]}{41.97 [\text{kWh/kg}]} \quad (15)$$

where  $P_{w2h}$  is the predicted wind power used for producing hydrogen and the constants are according to [51]. The  $\text{Sim}(\Delta H)$  function performs the system simulation in order to verify that the system operation fulfills the uninterrupted power supply requirement of the sensitive loads during the simulation time span. If verified, the output would be 1, otherwise 0. Detailed definition of the  $\text{Sim}(\Delta H)$  is as follows. Assume

$$P_G = (P_{\text{wind}} + P_{w2h}) + P_{PV} \quad (16)$$

and

$$P_L = (L - LC)/\eta_C \quad (17)$$

are the total generation and load, then, the rules for calculating the output is as follows.

- 1) If  $P_G = P_L$ , then the storage capacity remains unchanged.
- 2) If  $P_G > P_L$ , then the power surplus is used to produce hydrogen and supply the storage tank according to (15). If the tank reaches its maximum limit, the remainder of the power will not be used.
- 3) If  $P_G < P_L$ , then the power deficit required to supply the load is supplied by the fuel cell, according to (17).

The above steps are carried out for all times, eventually indicating one of the following outcomes: 1) the successful operation of the system which returns 1 in the output, or 2) the storage tank is discharged below the lowest permissible limit. If the latter occurs, the system operation is considered as a failure for which the simulation might not guarantee reliable power supply of the sensitive loads. It is assumed that 20% of the load consists of sensitive loads and cannot be curtailed. Hence, the output would be 0 if the chances of curtailing sensitive loads are higher than 2%. It should be noted that for regular loads the cost of curtailment is added to the objective function as the parameter  $LC$ .

### B. Calculation of Robust Solution Considering Forecast Uncertainty of Wind Power

It is assumed that the only uncertainty in the introduced microgrid results from the wind powers variable behaviour and forecasting errors. The controller/electrolyzer for the hydrogen storage receives short-term forecasts of wind power as input and determines optimal operational adjustments. Robust optimization is required to tackle effects of forecast uncertainty.

Robust optimization [39] deals with problems subject to a family of constraints that are parameterized by uncertainty terms. The ARP, as outlined in the Introduction, contributes to determining an optimal solution that is feasible for all possible constraints in the parameterized family. Meanwhile, it is important to note that this guarantees performance not in a deterministic sense, i.e., satisfaction of all possible uncertain outcomes, but instead in a probabilistic sense, i.e., satisfaction of a major set of the uncertainty possibilities or outcomes. In mathematical terms, an ARP may be formalized as

$$\min_{\theta \in \mathbb{R}^p} c^T \theta \quad \text{subject to : } \theta \in \bigcap_{s \in \{1, \dots, N\}} \Theta_s \quad (18)$$

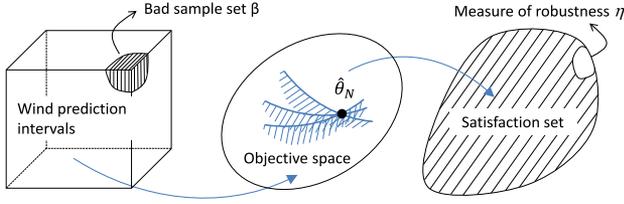


Fig. 10. Interpretation of the ARP's scenario approach to the RCP (as modified from [39]).

where  $\theta$  is the optimization variable,  $s$  is the uncertainty index represented by  $N$  samples, drawn according to the probability distribution of the uncertainty, and  $\{\Theta_s\}_{s=1,\dots,N}$  is a finite collection of convex sets in  $\mathbb{R}^n$ . The objective to be minimized can be nonlinear or in any other form. It should be noted that  $\theta$  can be assumed separately from  $s$ , as it is the case for the presented case study. Hence, the constraints are a function of the optimization variable ( $\theta \doteq \Delta H$ ) and the uncertainty variable ( $p_s$ ):  $f(\theta, p_s) \leq 0$ , where  $p_s = \{P_{\text{wind}} + P_{w2h}\}_s$  for the  $s$ th sample and  $s = 1, \dots, N$  represents samples calculated from the proposed PIs.

Furthermore, as in an application of the particle swarm optimization for solving the deterministic problem, the objective function can be of any form and even nonlinear nonconvex functions can be handled. Hence, the ARP is based on sampling at random a finite number of constraints in the family  $\bigcap_{s \in \{1, \dots, N\}} f(\theta, p_s) \leq 0$  and solving the corresponding deterministic problem.

The constraints, according to the problem of (10)–(14), are selected based on the probabilistic PIs or ensembles, hence, the resulting optimal solution  $\hat{\theta}_N$  is a random variable that depends on the random samples. Therefore,  $\hat{\theta}_N$  would be a  $\eta$ -level solution for a given random sampling

$$N \geq N_{\text{linear}}(\eta, \beta) \doteq \left\lceil \frac{n_\theta}{\eta\beta} \right\rceil \quad (19)$$

where  $n_\theta$  is the number of variables and the parameter  $\beta$  bounds the probability that  $\hat{\theta}_N$  is not a  $\eta$ -level solution. Accordingly,  $\beta$  is the confidence related to the solution algorithm.

Therefore, with probability no smaller than  $1 - \beta$ ,  $\hat{\theta}_N$  is  $\eta$ -level robustly feasible. Both parameters have a similar effect on the algorithm by (19). For example, %95 confidence and %95 robustness can be achieved by  $\beta = 0.05$  and  $\eta = 0.05$ , respectively. The inequality (19) provides the minimum number of sampled constraints that are needed in order to attain the desired probabilistic levels of robustness in the solution. The function  $N_{\text{linear}}$  gives therefore a bound on the generalization rate of the scenario approach which relates to the ability of the scenario solution of being feasible (with high probability) also with respect to constraints that were not explicitly taken into account in the solution of the ARP (unseen scenarios).  $N_{\text{linear}}$  is a linear function with respect to  $\beta^{-1}$ . Fig. 10 gives a visual interpretation of this theorem.

Optimization of all modeling outcomes drawn from the PIs provides a sampling of the optimal solution space. The final

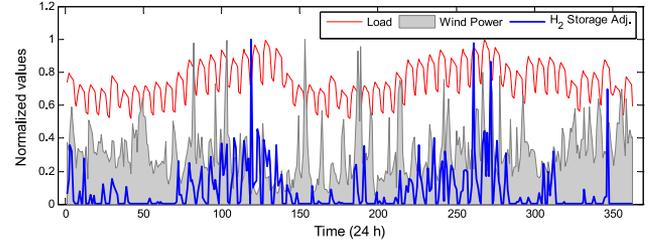


Fig. 11. Daily robust optimal operation of hydrogen storage based on multivariate conditional PIs of wind power forecasts.

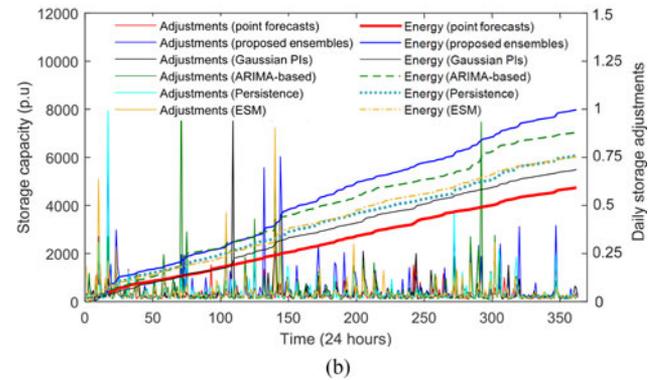
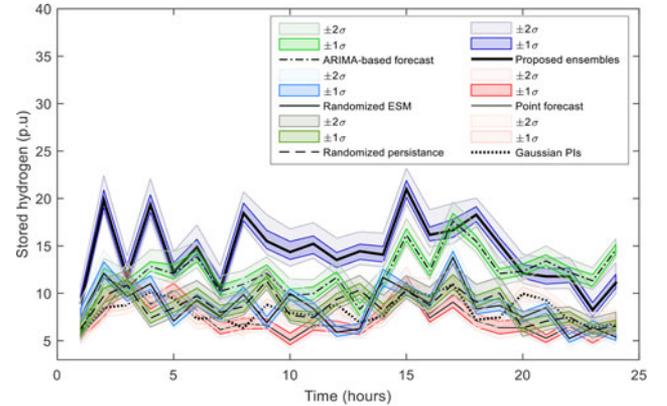


Fig. 12. Comparison of (a) distribution of the amounts of available hydrogen at each hour for the operation of the microgrid over one year and (b) the energy available from hydrogen over one year versus daily adjustments.

admissible and implementable optimal outcomes can then be extracted using the ARP. Wind forecast data of Section II is used and the uncertainty is according to the proposed and benchmark methods.

Fig. 11 shows normalized optimal adjustments of energy storage along with the load and wind power variations. The wind power forecasts provide 24 PIs for each day throughout the year. Each PI is represented by conditional scenarios of forecast uncertainty calculated from 500 000 samples of the multivariate copula-based model (see the Appendix). The ARP used to solve the problem of (10)–(14) uses a minimum of  $N = 200$  scenario constraints based on (19).

The main practical importance of using conditional PIs can be illustrated by Fig. 12. Fig. 12(a) shows that by using conditional PIs, the availability of energy from the energy storage system is often higher compared to other methods and deterministic point forecasts. The main reason is that the temporal dependence of forecast data is preserved by conditional PIs; hence, optimal operation of the energy storage conforms to the actual available energy from wind generation. A similar outcome is shown in Fig. 12(b) for the available energy from the hydrogen storage over a year of daily operation. It is interesting that by using the proposed robust optimization, more energy can be extracted from the system which makes the microgrid more reliable.

On the other hand, by randomly sampling the uncertainty of wind power forecasts, given that the uncertainty model is accurate and appropriate, the original infinite constraint set has been substituted with a finite set of  $N$  constraints. The resulting operations of storage [e.g., in Fig. 11 or Fig. 12(b)] fail to satisfy only a small portion of the original infinite constraints; however, by satisfying (19), it is explicitly guaranteed that the solution is practical and sufficiently robust. In the presented case study, the storage operations with probability no smaller than 0.9, are robustly feasible with a risk factor of 0.1.

#### IV. CONCLUSION

A stochastic model is proposed to characterize forecasting error of wind power. Both conditional and temporal interdependence between forecast error data and wind power levels can be captured by the model. By obtaining numerous short-term forecasting data over a long-term period (e.g., 6 h-ahead forecasts over a year), the accuracy of the proposed algorithm can be verified. The copula-based modeling is then used to calculate conditional empirical PIs or scenarios of future wind power uncertainty to assign the point forecasts with likely outcomes. Scenarios of wind power forecasts are then applied to optimize the required energy storage adjustments using a new method of robust programming. The presented method provides an efficient and explicit bound on the number of scenarios essential to obtain a solution that guarantees an *a-priori* specified robustness. Hence, there is a probabilistic guarantee of constraint satisfaction for decisions made on the modeled PIs. This is demonstrated by the results of a microgrid operation strategy involving hydrogen storage systems.

#### APPENDIX

##### ESTIMATING MINIMUM SIZE OF SAMPLE POOL

This appendix provides a simple formula for determining the minimum sample size for an efficient simulation. However, a more reasonable way is to experiment with different sample sizes in order to obtain sufficient output from the algorithm in Section II-A. Suppose that the sample pool required by the algorithm estimates  $\gamma = E[g(\mathbf{X})]$ , where  $\mathbf{X}$  is the random vector, and  $f(x)$  is the modeled density of  $\mathbf{X}$ . Then, simulation by sampling gives the estimate

$$\hat{\gamma} = \frac{1}{N_s} \sum_{j=1}^{N_s} g(\mathbf{X}_j) \quad (20)$$

with standard error  $\hat{\delta} = \delta\sqrt{N_s}$  where  $N_s$  is the sample size and  $\delta^2$  is the sample variance. The quality of estimation  $Q$  can be represented by half of the confidence interval ( $\alpha$ ) for  $\gamma$

$$Q = z_{\alpha/2} \sqrt{\frac{\text{Var}(g(\mathbf{X}))}{N_s}} \quad (21)$$

where  $z$  is the standard normal distribution. For better accuracy of the estimate,  $Q$  should be made small. Minimum sample size for achieving accuracy of  $Q$  is

$$N_s \geq \frac{z_{\alpha/2}^2}{Q} \times \text{Var}(g(\mathbf{X})). \quad (22)$$

For example,  $z_{\alpha/2} = 2.58$  for 99% confidence. If the variance is 10, assuming simulation of 48 variables, then a simple choice of  $N_s \geq 123\,000$  roughly provides quality of at least  $Q = 1\%$ .

#### REFERENCES

- [1] A. Botterud *et al.*, "Wind power trading under uncertainty in LMP markets," *IEEE Trans. Power Syst.*, vol. 27, no. 2, pp. 894–903, May 2012.
- [2] C. S. Ioakimidis, L. J. Oliveira, and K. N. Genikomsakis, "Wind power forecasting in a residential location as part of the energy box management decision tool," *IEEE Trans. Ind. Informat.*, vol. 10, no. 4, pp. 2103–2111, Nov. 2014.
- [3] N. Chen, Z. Qian, I. T. Nabney, and X. Meng, "Wind power forecasts using Gaussian processes and numerical weather prediction," *IEEE Trans. Power Syst.*, vol. 29, no. 2, pp. 656–665, Mar. 2014.
- [4] A. Tascikaraoglu and M. Uzunoglu, "A review of combined approaches for prediction of short-term wind speed and power," *Renew. Sustain. Energy Rev.*, vol. 34, no. 3, pp. 243–254, 2014.
- [5] M. B. Ozkan and P. Karagoz, "A novel wind power forecast model: Statistical hybrid wind power forecast technique (SHWIP)," *IEEE Trans. Ind. Informat.*, vol. 11, no. 2, pp. 375–387, Apr. 2015.
- [6] M. Khalid and A. Savkin, "Closure to discussion on "a method for short-term wind power prediction with multiple observation points", *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 1898–1899, May 2013.
- [7] J. Tastu, P. Pinson, P.-J. Trombe, and H. Madsen, "Probabilistic forecasts of wind power generation accounting for geographically dispersed information," *IEEE Trans. Smart Grid*, vol. 5, no. 1, pp. 480–489, Jan. 2014.
- [8] C. Wan, Z. Xu, P. Pinson, Z. Y. Dong, and K. P. Wong, "Probabilistic forecasting of wind power generation using extreme learning machine," *IEEE Trans. Power Syst.*, vol. 29, no. 3, pp. 1033–1044, May 2014.
- [9] P.-J. Trombe, P. Pinson, and H. Madsen, "A general probabilistic forecasting framework for offshore wind power fluctuations," *Energies*, vol. 5, no. 3, pp. 621–657, 2012.
- [10] P. Chen, T. Pedersen, B. Bak-Jensen, and Z. Chen, "ARIMA-based time series model of stochastic wind power generation," *IEEE Trans. Power Syst.*, vol. 25, no. 2, pp. 667–676, May 2011.
- [11] N. Chen, Z. Qian, I. Nabney, and X. Meng, "Wind power forecasts using gaussian processes and numerical weather prediction," *IEEE Trans. Power Syst.*, vol. 29, no. 2, pp. 656–665, Mar. 2014.
- [12] H. Quan, D. Srinivasan, and A. Khosravi, "Short-term load and wind power forecasting using neural network-based prediction intervals," *IEEE Trans. Neural Network Learn. Syst.*, vol. 25, no. 2, pp. 303–315, Feb. 2014.
- [13] S. Buhari and I. Cadirci, "Multi-stage wind-electric power forecast by using a combination of advanced statistical methods," *IEEE Trans. Ind. Informat.*, vol. 11, no. 1, pp. 1231–1242, Oct. 2015.
- [14] G. Zhang, H.-X. Li, and M. Gan, "Design a wind speed prediction model using probabilistic fuzzy system," *IEEE Trans. Ind. Informat.*, vol. 8, no. 4, pp. 819–827, Nov. 2012.
- [15] N. Amjadi, F. Keynia, and H. Zareipour, "Wind power prediction by a new forecast engine composed of modified hybrid neural network and enhanced particle swarm optimization," *IEEE Trans. Sustain. Energy*, vol. 2, no. 3, pp. 265–276, Jul. 2011.
- [16] Y. Ren, P. N. Suganthan, and N. Srikanth, "Ensemble methods for wind and solar power forecasting: state-of-the-art review," *Renew. Sustain. Energy Rev.*, vol. 50, pp. 82–91, 2015.

- [17] Y. Zhang, J. Wang, and X. Wang, "Review on probabilistic forecasting of wind power generation," *Renew. Sustain. Energy Rev.*, vol. 32, no. 1, pp. 255–270, 2014.
- [18] P. Pinson and G. Kariniotakis, "Conditional prediction intervals of wind power generation," *IEEE Trans. Power Syst.*, vol. 25, no. 4, pp. 1845–1856, Nov. 2010.
- [19] J. B. Bremnes, "Probabilistic wind power forecasts using local quantile regression," *Wind Energy*, vol. 7, no. 1, pp. 47–54, 2004.
- [20] L. Yang, M. He, J. Zhang, and V. Vittal, "Support-vector-machine-enhanced Markov model for short-term wind power forecast," *IEEE Trans. Sustain. Energy*, vol. 6, no. 3, pp. 791–799, Jul. 2015.
- [21] J. W. Taylor, P. E. McSharry, and R. Buizza, "Wind power density forecasting using ensemble predictions and time series models," *IEEE Trans. Energy Convers.*, vol. 24, no. 3, pp. 775–782, Sep. 2009.
- [22] C. Wan, Z. Xu, P. Pinson, Z. Y. Dong, and K. P. Wong, "Optimal prediction intervals of wind power generation," *IEEE Trans. Power Syst.*, vol. 29, no. 3, pp. 1166–1174, Nov. 2014.
- [23] H. Quan, D. Srinivasan, and A. Khosravi, "Incorporating wind power forecast uncertainties into stochastic unit commitment using neural network-based prediction intervals," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 26, no. 9, pp. 2123–2135, Sep. 2015.
- [24] X. Fang, Y. Wei, and F. Li, "Evaluation of LMP intervals considering wind uncertainty," *IEEE Trans. Power Syst.*, vol. 31, no. 3, pp. 2495–2496, May 2016.
- [25] T. Ding *et al.*, "Interval power flow analysis using linear relaxation and optimality-based bounds tightening (OBBT) methods," *IEEE Trans. Power Syst.*, vol. 30, no. 1, pp. 177–188, Jan. 2015.
- [26] R. J. Bessa, V. Miranda, A. Botterud, J. Wang, and E. M. Constantinescu, "Time adaptive conditional kernel density estimation for wind power forecasting," *IEEE Trans. Sustain. Energy*, vol. 3, no. 4, pp. 660–669, Oct. 2012.
- [27] R. J. Bessa, V. Miranda, A. Botterud, Z. Zhou, and J. Wang, "Time-adaptive quantile-copula for wind power probabilistic forecasting," *Renew. Energy*, vol. 40, no. 1, pp. 29–39, 2012.
- [28] H. V. Haghi, M. T. Bina, M. A. Golkar, and S. M. M. Tafreshi, "Using copulas for analysis of large datasets in renewable distributed generation: PV and wind power integration in Iran," *Renew. Energy*, vol. 35, no. 9, pp. 1991–2000, 2010.
- [29] N. Zhang, C. Kang, Q. Xia, and J. Liang, "Modeling conditional forecast error for wind power in generation scheduling," *IEEE Trans. Power Syst.*, vol. 29, no. 3, pp. 1316–1324, May 2013.
- [30] V. Gabrel, C. Murat, and A. Thiele, "Recent advances in robust optimization and robustness: An overview," LAMSADE, Université Paris Dauphine, Paris, France, Tech. Rep., 2012. [Online]. Available: [http://www.optimization-online.org/DB\\_HTML/2012/07/3537.html](http://www.optimization-online.org/DB_HTML/2012/07/3537.html)
- [31] G. Calafiore and F. Dabbene, *Probabilistic and Randomized Methods for Design Under Uncertainty*. New York, NY, USA: Springer, 2006.
- [32] A. Ben-Tal and A. Nemirovski, "Robust convex optimization," *Math. Oper. Res.*, vol. 23, no. 4, pp. 769–805, 1998.
- [33] L. El Ghaoui and H. Lebret, "Robust solutions to uncertain semi-definite programs," *SIAM J. Optimisation*, vol. 9, no. 1, pp. 33–52, 1998.
- [34] A. Prkopa, *Stochastic Programming*. Norwell, MA, USA: Kluwer, 1995.
- [35] S. Vajda, *Probabilistic Programming*. New York, NY, USA: Academic, 1972.
- [36] D. Bertsimas, D. B. Brown, and C. Caramanis, "Theory and applications of robust optimization," *SIAM Rev.*, vol. 53, no. 3, pp. 464–501, 2011.
- [37] D. Bertsimas and D. B. Brown, "Constructing uncertainty sets for robust linear optimization," *Oper. Res.*, vol. 57, no. 6, pp. 1483–1495, 2009.
- [38] W. Chen, M. Sim, J. Sun, and C. P. Teo, "From CVAR to uncertainty sets: Implications in joint chance-constrained optimization," *Oper. Res.*, vol. 58, no. 2, pp. 470–485, 2010.
- [39] G. Calafiore and M.C. Campi, "The scenario approach to robust control design," *IEEE Trans. Autom. Control*, vol. 51, no. 5, pp. 742–53, May 2006.
- [40] C. Calafiore and L. Fagiano, "Robust model predictive control via scenario optimization," *IEEE Trans. Autom. Control*, vol. 58, no. 1, pp. 219–224, Jan. 2013.
- [41] C. Chatfield, "Prediction intervals for time series forecasting," in *Principles of Forecasting* (A Handbook for Practitioners and Researchers), J. S. Armstrong, Ed. New York, NY, USA: Springer, 2001, pp. 475–494.
- [42] G. E. P. Box, G. M. Jenkins, and G. C. Reinsel, *Time Series Analysis. Forecasting and Control*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1994.
- [43] R. B. Nelsen, *An Introduction to Copulas*. New York, NY, USA: Springer, 2006.
- [44] H. V. Haghi, M. T. Bina, and M. A. Golkar, "Nonlinear modeling of temporal wind power variations," *IEEE Trans. Sustain. Energy*, vol. 4, no. 4, pp. 838–848, Oct. 2013.
- [45] P. Pinson, "Very short-term probabilistic forecasting of wind power with generalized logit-Normal distributions," *J. Roy. Stat. Soc., C*, vol. 61, no. 4, pp. 555–576, 2012.
- [46] A. Lau and P. McSharry, "Approaches for multi-step density forecasts with application to aggregated wind power," *Ann. Appl. Stat.*, vol. 4, no. 3, pp. 1311–1341, 2010.
- [47] T. Gneiting, F. Balabdaoui, and A. E. Raftery, "Probabilistic forecasts, calibration and sharpness," *J. Roy. Stat. Soc. B*, vol. 69, no. 2, pp. 243–268, 2007.
- [48] C. W. Potter, D. Lew, J. McCaa, S. Cheng, S. Eichelberger, and E. Gritmit, "Creating the dataset for the western wind and solar integration study (U.S.A.)," in *Proc. 7th Int. Workshop Large Scale Integr. Wind Power Trans. Netw. Offshore Wind Farms*, 2008. [Online]. Available: [http://www.nrel.gov/electricity/transmission/western\\_wind.html](http://www.nrel.gov/electricity/transmission/western_wind.html)
- [49] G. N. Kariniotakis, G. S. Stavrakakis, and E. F. Nogaret, "Wind power forecasting using advanced neural networks models," *IEEE Trans. Energy Convers.*, vol. 11, no. 4, pp. 762–767, Dec. 1996.
- [50] G. Calafiore and M. C. Campi, "Uncertain convex programs: Randomized solutions and confidence levels," *Math. Program.*, vol. 102, no. 1, pp. 25–46, 2005.
- [51] S. M. Hakimi and S. M. Moghaddas-Tafreshi, "Optimal sizing of a stand-alone hybrid power system via particle swarm optimization for Kahnouj area in south-east of Iran," *Renew. Energy*, vol. 34, pp. 1855–1862, 2009.



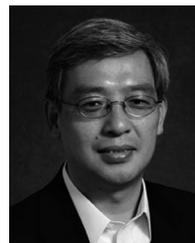
**Hamed Valizadeh Haghi** (M'16) is currently working toward the Ph.D. degree at the University of Central Florida, Orlando, FL, USA.

He is currently a Ph.D. Researcher at the University of Central Florida. His experience and expertise are in smart grid and data analytics. His research interests include predictive modeling and control of power systems, distributed and stochastic optimization, demand response, energy storage, and wind generation.



**Saeed Lotfifard** (S'08–M'11) received the Ph.D. degree in electrical engineering from Texas A&M University, College Station, TX, USA, in 2011.

He is currently an Assistant Professor with Washington State University, Pullman, WA, USA. His research interests include power systems protection and control.



**Zhihua Qu** (M'90–SM'93–F'09) received the Ph.D. degree in electrical engineering from the Georgia Institute of Technology, Atlanta, GA, USA, in 1990.

Since 1990, he has been with the University of Central Florida, Orlando, FL, USA, where he is currently the SAIC Endowed Professor with the College of Engineering and Computer Science, a Professor and the Chair of Electrical and Computer Engineering, and the Director of the Foundations for Engineering Education for Distributed Energy Resources (FEEDER) Center (one of the Department of Energy-funded national centers on distributed technologies and smart grid). His areas of expertise are nonlinear systems and control with applications to energy and power systems. In energy systems, his research covers such subjects as low speed power generation, dynamic stability of distributed power systems, anti-islanding control and protection, distributed generation and load sharing control, distributed VAR compensation, distributed optimization, and cooperative control.