Ellipsoidal Prediction Regions for Multivariate Uncertainty Characterization

Faranak Golestaneh, *Student Member, IEEE,* Pierre Pinson, *Senior Member, IEEE,*
Rasoul Azizipanah-Abarghooee *Student Member, IEEE* and Hoay Beng Gooi, *Senior Member, IEEE*

***Abstract*—While substantial advances are observed in probabilistic forecasting for power system operation and electricity market applications, most approaches are still developed in a univariate framework. This prevents from informing about the interdependence structure among locations, lead times and variables of interest. Such dependencies are key in a large share of operational problems involving renewable power generation and electricity prices for instance. The few methods that account for dependencies translate to sampling scenarios based on given marginals and dependence structures. However, for classes of decision-making problems based on robust, interval chance-constrained optimization, necessary inputs take the form of multivariate prediction regions rather than scenarios. The current literature is at very primitive stage of characterizing multivariate prediction regions to be employed in these classes of optimization problems. To address this issue, we introduce a new class of multivariate forecasts which form as multivariate ellipsoids for non-Gaussian variables. We propose a data-driven systematic framework to readily generate and evaluate ellipsoidal prediction regions, with predefined probability guarantees and minimum conservativeness. A skill score is proposed for quantitative assessment of the quality of prediction ellipsoids. A set of experiments is used to illustrate the discrimination ability of the proposed scoring rule for potential misspecification of ellipsoidal prediction regions. Application results based on three datasets with wind, PV power and electricity prices, allow us to assess the skill of the resulting ellipsoidal prediction regions, in terms of calibration, sharpness and overall skill.**

***Index Terms*—Probabilistic forecasting, uncertainty sets, ellipsoids, robust optimization, chance-constrained optimization**

I. Introduction

THE RAPID deployment and integration of renewable energy generation capacities have increased the level of variability and uncertainty in power systems, possibly also magnified by new electricity consumption patterns. This comes in a context of deregulation of energy markets, eventually resulting in a more complex environment for decision-makers. This calls for the development of a number of forecasting methodologies providing the suitable input to a wealth of decision-making problems in power system operation and electricity market participation, under uncertainty and with a view on risk management [1].

The primary goal of this paper is to propose a new class of multivariate probabilistic forecasts to characterize the uncertainty in those random variables in power systems which are correlated in time and space. The proposed framework predicts the correlated uncertainty in a form which is tailored for the classes of decision-making problems based on robust, interval chance-constrained optimization. The output of the proposed framework can serve as a valuable and informative input for those operational problems such as optimal power flow which need an estimation of dependent and uncertain variables.

The most common type of forecasts appeared in the literature and practice is deterministic and it takes the form of single values for each variable of interest, location and lead time individually and independently. Although these are easier to interpret and to use as input to decision-making problems, they are always subject to errors [2]. Costs induced by such errors often motivate to provide a full picture of potential forecast errors and to accommodate uncertainty estimates in associated optimization problems. Probabilistic forecasting then comprises the appropriate framework to generate that information [3]. Probabilistic forecasts are, however, most often produced in a univariate framework, i.e., still providing uncertainty information for every variable, lead time and location, individually. They are only suboptimal inputs to decision-making when temporal, spatial and/or inter-variable dependencies are to be considered. Besides, due to the inertia in meteorological systems and their impact of renewable power generation, load and electricity markets, such dependencies are expected to be present.

The common practice in the literature for modeling spatial/temporal dependencies is to generate space/time scenarios as random draws of multivariate prediction distributions [4]. This is motivated by the fact that a number of optimization methods e.g. stochastic programming are to use scenarios and scenario trees as inputs. However, chance-constrained [5], robust [6], [7], interval [8] optimization require the definition of multivariate probabilistic forecast regions rather than scenarios. A multivariate probabilistic forecast region defines a region where the realization of a multivariate random variable is expected to lie, with a certain probability. Only few proposals on generation of multivariate prediction regions can be found in the literature [9]–[11], referred to as adjusted intervals and Chebyshev-based intervals. The idea is to use already generated sets of scenarios and to deduce prediction regions as a minimum volume that covers given proportions of these scenarios. As reported in [9], the prediction intervals with low nominal coverage rates are too wide and conservative, while difference in the size of the regions for varying nominal coverage rates is low.

For many of the decision-making problems formed based on robust optimization, multivariate uncertainty sets (another term for multivariate prediction regions when used as inputs to optimization) are assumed to have ellipsoidal geometry [12], [13]. For example in [13], two types of cuts are proposed for minmax regret problems with ellipsoidal uncertainty sets, where ellipsoidal uncertainty sets are considered as more flexible and realistic uncertainty sets compared to finite or hyper-boxes. Also in [14], ellipsoidal uncertainty sets are introduced as a relevant uncertainty representation in robust unit-commitment. Ellipsoidal approximations have been of interest

in reachability analysis for hybrid systems. The ellipsoidal approximations are performed to calculate the reach sets for linear control systems with time-varying coefficients and hard bound on the controls and initial states. The approximations are described through ordinary differential equations with coefficients expressed in an explicit analytical form [15], [16].

As for univariate probabilistic forecasts, prediction ellipsoids ought to provide probabilistically reliable and skillful information about multivariate uncertainty. To the best of our knowledge, there is no established practice so far to generate and evaluate prediction ellipsoids with predefined probability to be used as inputs to optimization problems. In practice, parameters of the ellipsoids are chosen based on expert knowledge, assumptions or trial and error. For instance in [17], a framework is described where the conservativeness of the ellipsoids is controlled by a parameter called uncertainty budget. This parameter is decided through trial and error, with a higher uncertainty budget resulting in a higher probability and conservativeness. A clear disadvantage is that the probability associated with the ellipsoids (their coverage rates) cannot be determined in advance. This is while, in practice, one is most likely interested in having prediction ellipsoids with various predefined nominal coverage rates, e.g. 90%, 95% or 99%.

In this paper, an optimization-regression framework is developed to generate prediction ellipsoids with predefined probability and high performance. The most straightforward assumption about the properties of prediction ellipsoids is to consider them as Gaussian geometries. In this case, the prediction ellipsoids can be considered as contours of constant density in multivariate Gaussian distributions where the density is determined by the percentiles of χ^2 distribution.

The non-Gaussian nature of wind and Photovoltaic (PV) power and electricity prices has been studied in the literature. A Gaussian distribution is unlimited on both tails while both PV and wind power are double-bounded random variables limited between zero and the maximum capacity of the installation [18], [19]. In [20], the Kolmogorov-Smirnov (KS) test is conducted to investigate whether Gaussian, Rayleigh, Beta, Gamma and Weibull are suitable parametric distributions to characterize the forecast errors in two PV power datasets. The KS test rejected the validity of the mentioned distributions at 5% significance level. A statistical analysis of wind generation forecast errors is presented in [21]. In this work, it is argued that distributions of wind power forecasting errors exhibit large and variable kurtosis and have fat-tails which cannot be characterized by a Gaussian distribution. The same argument is made in [22], [23]. Day-ahead and hour-ahead wind power forecasting errors in six different countries are studied in [24]. In [24], it is shown that large forecast errors are not represented by Gaussian distributions while they have a large impact on operation and planning in power systems. In [25], it is indicated that electricity prices like any financial asset returns are not normally distributed and are heavy-tailed and exhibit excess kurtosis. Similar conclusions are made in [26], [27].

Our empirical investigations revealed that Gaussianity assumption of prediction ellipsoids is not valid for the important random variables in power systems, namely, PV and wind power and electricity price. Those prediction ellipsoids designed based on Gaussianity assumption show very low calibration and reliability. Therefore, in this work, prediction ellipsoids are generated without restrictive assumptions to skillfully mimic the true underling stochastic process. The proposed prediction ellipsoids are called Ellipsoidal Prediction Regions (EPRs).

In the proposed framework, the centers of EPRs are point forecasts while the covariance matrix of the ellipsoids are found by either exponential smoothing or Dynamic Conditional-Correlation-GARCH (DCC-GARCH) [28]. The choice of covariance matrix forecast technique depends on the inherent uncertainty of random variables. We use exponential smoothing for those random variables with slow-moving covariance matrix while DCC-GARCH performs much better to forecast a time-varying covariance matrix. The scale parameters are determined through an optimization procedure using the historical data. An objective function is proposed for the optimization phase, leading the EPRs to have lower conservativeness and better probability guarantees. The proposed model is capable of outputting calibrated EPRs with predefined coverage rates.

Because the literature on multivariate prediction regions is at a primitive stage, there is no established evaluation framework for this class of forecasts. A scoring rule is proposed here for quantitative assessment of the prediction ellipsoids based on the essential characteristics required for skilled forecasts, namely reliability (calibration) and sharpness (low conservativeness). A set of empirical experiments are designed to examine the ability of the proposed scoring rule in discriminating possible prediction misspecification in a multivariate context. Additionally, a formulation is proposed to estimate the size of ellipsoids for bounded random variables. The efficiency of the proposed framework is evaluated for wind and PV power and electricity price. Temporal prediction ellipsoids of dimensions 2, 11 and 24 with the probability of 5% to 95% with 5% increments are generated and evaluated.

The rest of the paper is organized as follows: In Section II, the basic definition and formulation of prediction ellipsoids are given. The proposed scoring rule is explained in Section III. The proposed framework to generate EPRs is described in Section IV. Section V contains the empirical results and finally concluding remarks are given in Section VI.

II. Ellipsoidal Prediction Regions: Basics and Formulation

Let $\mathbf{X}$ be a multivariate random variable of dimension D. For simplicity and without loss of generality, here, we consider temporal dependency only. However, the same framework can be employed for spatio-temporal dependency characterization. In case of temporal dependency, $\mathbf{X}$ can be described as $\mathbf{X}_t = [X_{t+k_1} \quad ... \quad X_{t+k_D}]$ with $k_i \ \forall i$ as the forecast horizons. To simplify the notation, hereafter $\mathbf{X}_t$ is denoted as $\mathbf{X}_t = [X_{t+1} \quad ... \quad X_{t+D}]$.

Let PE^α be the prediction ellipsoid with the nominal coverage rate as α, where α represents the proportion of realizations of $\mathbf{X}$ inside PE^α [29].

$$PE^\alpha : (\mathbf{x}-\mu)^\top \Sigma^{-1}(\mathbf{x}-\mu) \leq \Upsilon^\alpha \tag{1}$$

where $\mu = E(\mathbf{X})$ is the mean vector of the random variable and the center of the ellipsoid. $\Sigma = E[(\mathbf{X}-\mu)(\mathbf{X}-\mu)^\top]$ is the covariance matrix. Υ^α is called the scale or robust parameter

for an ellipsoid with the nominal coverage α. It should be noted that hereafter, upper case letters symbolize random variables while lower case letters express their realizations.

When $\mathbf{X}$ follows a multivariate Gaussian distribution, $\mathbf{X} \sim \mathcal{MVN}(\mu, \Sigma)$, $\mathbf{x}$ in (1) describes the contours of constant density for a D-dimensional normal distribution. In this case, the scale parameters are the percentiles of χ^2 distribution as

$$PE^{\alpha} : (\mathbf{x}-\mu)^{\top}\Sigma^{-1}(\mathbf{x}-\mu) \leq \chi^2_D(\alpha) \tag{2}$$

where $\chi^2_D(\alpha)$ is the lower 100^{th} percentile of χ^2 with D degrees of freedom. The ellipsoid in (2) has probability of α.

In robust optimization, the scale parameter is also called the uncertainty budget and it controls the trade-off between robustness and performance. The value of uncertainty budget usually is selected arbitrarily or by trial and error in the range of $[0, D^2]$ [17]. Let us call the ellipsoids characterized by the uncertainty budget as the robust ellipsoids.

In probability and statistics, the coverage rate of a prediction region is the proportion of times that the region contains the realizations from a random variable of interest. Similarly, the nominal coverage rate represents the expected coverage of that region while the observed coverage expresses the empirical coverage of the region calculated using actual data. The EPRs should not only be calibrated, i.e., providing observed coverage rate close to their nominal ones, but also need to be as concentrated (sharp) as possible. Basically, for the same coverage probability, a smaller prediction region is much more informative because it facilitates decision-making by giving a smaller range for uncertainty. In this paper, the idea is to design the EPRs by making a compromise between conservativeness and calibration. We propose a systematic approach to predict the parameters of the prediction ellipsoids in (1). These parameters include a location parameter μ (mean vector), a shaping parameter Σ (covariance matrix), and a scaling parameter Υ^{α} (being a function of the nominal coverage rate). In Section IV, we explain how these parameters are predicted. Here, we give a simple example to illustrate the importance of an accurate estimation of EPR parameters for a skillful characterization of multivariate uncertainty.

Fig. 1 illustrates two typical bivariate EPRs. For EPR 1, the predicted scaling $\Upsilon^{0.95}$ is equal to 5.44, while location μ, and shape Σ are

$$\Sigma = \begin{bmatrix} 0.000925 & 0.000673 \\ 0.000673 & 0.000629 \end{bmatrix}, \quad \mu = \begin{bmatrix} 0.219 & 0.192 \end{bmatrix} \tag{3}$$

Now, let us assume that three different multivariate observations (here, of dimension 2) are collected: $\mathbf{x}_1 = [0.23 \;\; 0.20]$, $\mathbf{x}_2 = [0.20 \;\; 0.15]$ and $\mathbf{x}_3 = [0.24 \;\; 0.16]$. Given the defined values for μ, Σ, Υ^{α}, by inputing $\mathbf{x}_1$ on the left hand-side of the inequality in (1), one obtains a value 0.184. The fact that this value is less than $\Upsilon^{0.95} = 5.44$ tells this observation lies within the EPR. In contrast for $\mathbf{x}_2$ and $\mathbf{x}_3$, the values obtained are 17.15 and 5.44, placing these two observations outside and on the boundary of the EPR, respectively. If having many more sample observations, one would expect close to 95% of the observations to be covered by the EPR. The scaling parameter $\Upsilon^{0.95}$ is therefore crucial.

In Fig. 1, EPR 2 has the same location μ and scale $\Upsilon^{0.95}$ but different shape Σ, which is given as

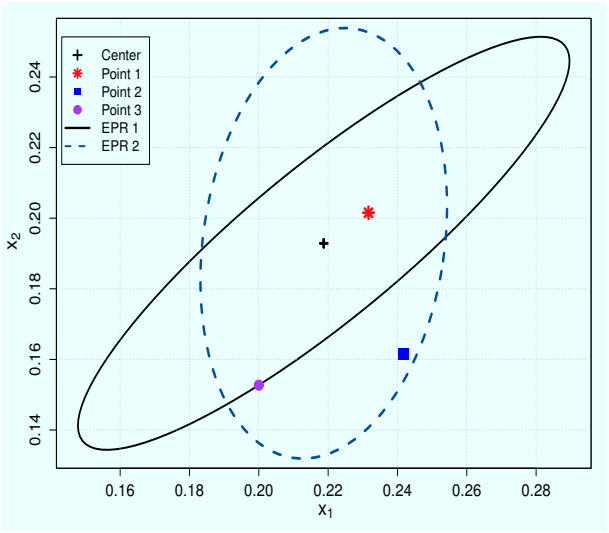


Fig. 1: Example bivariate EPRs and related observations.

$$\Sigma = \begin{bmatrix} 0.000673 & 0.000315 \\ 0.000315 & 0.000647 \end{bmatrix} \tag{4}$$

Σ highly impacts both size and orientation of the EPRs. Therefore, it is of utmost importance to predict the shaping parameter Σ accurately.

III. Evaluating the Skill of Ellipsoidal Prediction Regions

The aim of designing a scoring rule is to provide a theoretically principled framework for quantitative assessment of predictive performance of ellipsoidal prediction regions. In general, two properties are required for probabilistic forecasts, namely calibration and sharpness.

To assess the quality of univariate quantile forecasts, the Negative Quantile-based Score (NQS) as a proper score has been used widely in theory and practice [30]. NQS is formulated as

$$NQS^{\alpha_i}_{k_i} = -\frac{1}{T}\sum_{t=1}^{T}\left(\xi^{\alpha_i}_{t,k_i} - \alpha_i\right)\left(x_{t+k_i} - \hat{q}^{\alpha_i}_{t+k_i}\right) \;\; \forall \alpha_i \tag{5}$$

where $\hat{q}^{\alpha_i}_{t+k_i}$ is defined as a set of univariate quantile forecasts with the nominal coverage rate α_i for the lead-time k_i. x_{t+k_i} is the measurement at time t and T is the length of evaluation set. $\xi^{\alpha_i}_{t,k_i}$ is a binary variable indicating whether the quantile forecasts cover the measured random variable or not. $\xi^{\alpha_i}_{t,k_i}$ is defined as

$$\xi^{\alpha_i}_{t,k_i} = \mathbf{1}\{x_{t+k_i} \leq \hat{q}^{\alpha_i}_{t+k_i}\} = \begin{cases} 1, & \text{if } x_{t+k_i} \leq \hat{q}^{\alpha_i}_{t+k_i} \\ 0, & \text{otherwise} \end{cases} \tag{6}$$

In (5), $\left(\xi^{\alpha_i}_{t,k_i} - \alpha_i\right)$ is a measure of reliability/calibration of quantiles while $\left(x_{t+k_i} - \hat{q}^{\alpha_i}_{t+k_i}\right)$ is an expression of sharpness/size of the quantile forecasts and it assesses the distance between quantile forecasts and the measurement. NQS evaluates the sharpness and calibration of quantile forecasts jointly. We generalize NQS to propose a scoring rule for evaluation of multivariate prediction ellipsoids.

In the context of ellipsoidal regions similar to the case of univariate quantiles, calibration is referred to the proximity

of the nominal coverage rate of an ellipsoid to its observed coverage. The definition of the sharpness though is more challenging in this new context. One can consider the volume of an ellipsoid as the most straightforward representation of sharpness. In the following subsection, a scoring rule is proposed for verification of predictive performance of EPRs as a generalized form of NQS in a multivariate context. In the second subsection, a number of empirical experiments are designed to evaluate the capability of the proposed scoring rule to discriminate various parameter misspecification in the generation of multivariate prediction ellipsoids.

A. Formulation

The forecaster is always looking for reliable and calibrated prediction regions with a minimal area or volume possible to reduce the conservativeness. Sharpness and calibration can be assessed simultaneously through a skill score. A negatively-oriented skill score is expected to assign the lowest score value to the actual (true) ellipsoid. The proposed ellipsoidal skill score is given by

$$Sc^{\alpha_i} = \left| \frac{1}{T} \sum_{t=1}^{T} \left(\xi_t^{\alpha_i} - \alpha_i\right) \left(V_t^{\alpha_i}\right) \frac{1}{D} \right| \tag{7}$$

where T is the number of multivariate ellipsoids available in the evaluation data. $\xi_t^{\alpha_i}$ is an indication variable which is equal to 1 if the observed trajectory is inside the predicted geometrical region and is 0 otherwise. The observed trajectory is inside the ellipsoid if it satisfies (1). α_i shows the nominal coverage rate of the predicted geometrical region. $V_t^{\alpha_i}$ is the volume of multivariate ellipsoid with nominal coverage rate α_i at time t and it is calculated by

$$V_t^{\alpha} = \frac{\pi^{\frac{D}{2}}}{\Gamma(\frac{D}{2}+1)} \sqrt{(\Upsilon_t^{\alpha})^D \det(\Sigma_t)} \tag{8}$$

where Γ represents Gamma function [17] and $\det(.)$ is the determinant function.

The first pair of parentheses in (7) as the representation of calibration is similar to the same in (5) with the only difference in the definition of $\xi_t^{\alpha_i}$. The second pair of parentheses in (5) is replaced with the D^{th} root of volume in (7). We include the D^{th} root of volume instead of volume itself for two reasons. First, to be able to compare the skill scores of EPRs with different dimensions. Second, in case variables are in per-unit order, including the D^{th} root of volume makes the values from both pairs of parentheses in a fairly similar range. Therefore, the impact of calibration and size on overall skill would be comparable.

To get a single score for all prediction ellipsoids with nominal coverage rates $\alpha_i, i = 1, ..., m$, one can sum individual scores as

$$Sc = \sum_{i=1}^{m} \left(Sc^{\alpha_i}\right) \tag{9}$$

In order to assess calibration only, as for univariate probabilistic forecasts [30], one can calculate the observed coverage rate and compare it with nominal one. The observed coverage rate can be calculated as

$$\hat{\alpha}_i = \frac{1}{T} \sum_{t=1}^{T} \xi_t^{\alpha_i} \tag{10}$$

It is to be noted that the formulation given in (8) is accurate if the ellipsoids do not exceed the feasible limits of random variables. The prediction ellipsoids for PV and wind power are bounded between zero and nominal capacity of the corresponding wind or PV installation. Therefore, the feasible volume of each D-dimensional prediction ellipsoid is the intersection of that ellipsoid and a D-dimensional polyhedron. Calculation of the volume of the intersection analytically is intractable. However, one can use a Monte Carlo based approach to estimate the feasible volume numerically [17]. The proposed methodology to estimate the feasible volumes of EPRs is explained in Appendix A.

B. Evaluation of discrimination ability

The possible prediction errors in ellipsoidal context are the errors in prediction of the location parameter (center), the correlation (covariance) matrix, variance in each dimension and the scale parameter. To investigate the ability of the skill score proposed in (7) to detect possible prediction errors, the following experiments are designed. In all experiments $T = 10,000$ vectors of realizations of random variable $\mathbf{X}$ are generated from the actual Gaussian density. Let $\mathbf{x}_t$ be the realization of $\mathbf{X}$ at time t with x_i as its element at the i^{th} dimension. Let the actual density be defined with zero mean and unit variance of dimension $D = 24$, and covariance function as

$$\Sigma(x_i, x_j) = \sigma_i \sigma_j \exp(-\frac{|i-j|}{4}) \quad i, j = 1, ..., D \tag{11}$$

with $\sigma_i \, \forall i$ as the variance of $\mathbf{X}$ in its i^{th} dimension.

1) *Misspecified location parameter* (*center*): In this scenario, the prediction ellipsoids are assumed to have the correct covariance matrix as described in (11) and the correct scale parameters as given in (2) but erroneous center as $\hat{\mu}_i = \Xi(-1, 1), i = 1, ..., D$. $\Xi(a, b)$ is a function which generates decimal values between a and b from the Uniform distribution. For 10,000 successive times, 10,000 mean vectors are generated and assumed to be the center of 10,000 prediction ellipsoids.
2) *Misspecified variance*: The prediction ellipsoids for this case are formulated with true location and scale parameters but with wrong variance as $\hat{\sigma}_i = \sigma_i + \Xi(-0.15, 1), i = 1, ..., D$.
3) *Misspecified covariance model and strength*: The prediction ellipsoids are modeled with the actual center, variance and scale parameters but with misspecified correlation models and correlation strengths as

$$\hat{\Sigma}(x_i, x_j) = \sigma_i \sigma_j (1 + \frac{|i-j|}{r})^{-1} \quad i, j = 1, ..., D \tag{12}$$

with $r = \Xi(2, 6)$.
4) *Misspecified scale parameter*: Prediction ellipsoids in this case have the actual location parameter, variance, covariance matrix but they are characterized with wrong scale parameters Υ_t^{α}.

$$\Upsilon_t^{\alpha} = \Xi(0.01\chi_D^2(\alpha), 3\chi_D^2(\alpha)) \tag{13}$$

Subject to: if $\alpha_i > \alpha_j$, then $\Upsilon^{\alpha_i} > \Upsilon^{\alpha_j}$

Fig. 2 demonstrates the discrimination ability of the skill score introduced in (7) in detecting the various types of

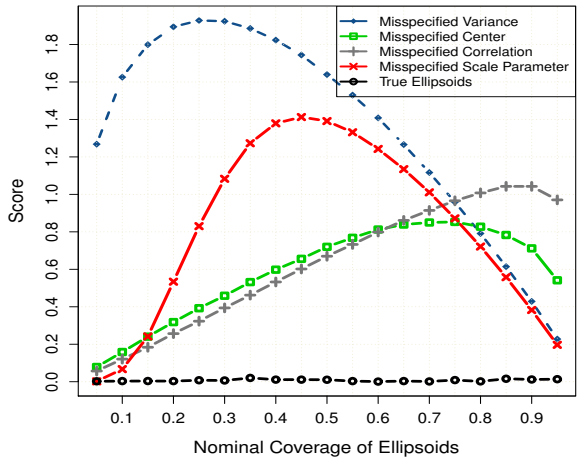


Fig. 2: Skill scores for true Gaussian ellipsoids versus those for various misspecified ellipsoids.

misspecification. The scores are calculated per α ranging from 0.05 to 0.95 with 0.05 increments. As it is shown in Fig. 2, the best scores are obtained for the true ellipsoids. The relatively lower scores of true ellipsoids with respect to the other four misspecified ellipsoids allow for relying on the proposed skill score to discriminate a good ellipsoidal representation of uncertainty against an incorrectly specified one.

IV. Ellipsoidal Prediction Regions

The proposed EPRs are constructed through an optimization-regression framework. The optimization phase is conducted only once for historical data to find the scale parameters. The first regression phase deals with univariate point forecasting. The second regression phase updates the covariance matrix of point forecast errors using the forecast errors calculated up to time t. The goal is to generate m EPRs with nominal coverage rates $A = [\alpha_1\ \alpha_2\ ...\ \alpha_m]$.

1) Parameters specifications: The EPRs are given by

$$(\mathbf{x}_t - \mu_t)^\top \Sigma_t^{-1} (\mathbf{x}_t - \mu_t) \leq \Upsilon^\alpha \tag{14}$$

μ_t: Considered to be the point forecasts for time t. Denote $\hat{\mathbf{x}}_t = [\hat{x}_{1,t}\ \ \hat{x}_{2,t}\ \ ...\ \ \hat{x}_{D,t}]$, with $\hat{x}_{i,t}$, $\forall i$ as the point forecast for time t and dimension i where $\hat{x}_{i,t}$ for each dimension is generated independently. Let us call $\hat{\mathbf{x}}$ as the predicted or estimated trajectory and $\mathbf{x}$ as the measurement or true trajectory. With this definition, μ_t is considered to be equal to $\hat{\mathbf{x}}_t$.

Σ: The covariance matrix of point forecast errors. Covariance matrix is a critical input in multivariate dependency characterization. To generate skilled predictive ellipsoids, if correlations between the random variables of interest and/or their variances change over time, forecasting the future covariance/correlations is required.

When forecasting the covariance matrix, one should determine whether the covariance matrix is constant, very slow-moving or time-varying. If the covariance matrix is constant or varies very slowly over time, the rolling historical correlations and exponential smoothing are two options to be employed [31]. These two methods have been widely used in the literature and practice because of their simplicity.

To estimate time-varying conditional covariance matrices, various multivariate GARCH methods can be used. Three important factors in performance evaluation of a covariance matrix forecasting technique are whether it can make the covariance matrix parsimonious enough, but still maintain flexibility and ensure to provide positive definite covariance matrices [32].

The problem with the conventional multivariate GARCH models is that the parameters of the conditional covariance matrix grow very rapidly as the dimension increases. This makes the conventional multivariate GARCH models computationally highly demanding. The Constant Conditional Correlation-GARCH (CCC-GARCH) model is an alternative approach which is computationally very attractive. It assumes the conditional correlation matrix is constant over time and only conditional standard deviations are time-varying. This is very restrictive as the assumption of constant conditional covariance matrix is not true all the time [32]. The DCC-GARCH model is an extension of CCC-GARCH which allows for time-varying conditional correlation matrix. DCC-GARCH is a computationally efficient multivariate model, which has the flexibility of the univariate GARCH models while it parameterizes the conditional correlations directly [28]. In DCC-GARCH, the number of parameters to be estimated is independent on the dimension of the times series to be correlated. This allows for estimation of large covariance matrices [28], [32], [33].

In this work, the covariance matrix is updated any time a new set of measurement/point forecast trajectories is available.

Υ^α: It is a function of the nominal coverage rate (α) and is obtained through an optimization framework. Denote $\boldsymbol{\Upsilon}$ as the compact form of Υ^α as $\boldsymbol{\Upsilon} = [\Upsilon^{\alpha_1}\ \Upsilon^{\alpha_2}\ ...\ \Upsilon^{\alpha_m}]$.

2) Utility function and constraints: The idea is to optimize $\boldsymbol{\Upsilon}$ such that the EPRs present desired probability guarantees and conservativeness. This implies that the observed coverage rates of the predictive geometrical regions should be as close as possible to the nominal coverage rates while their volumes are kept minimal. Therefore, the potential objective function for optimization can be introduced as

$$\arg_{\boldsymbol{\Upsilon}} \min \frac{1}{m} \sum_{i=1}^{m} \left(\left| \frac{1}{T} \sum_{t=1}^{T} \left(\xi_t^{\alpha_i} - \alpha_i \right) \left(V_t^{\alpha_i} \right) \frac{1}{D} \right| \right) \tag{15}$$

subject to

$$\text{if } \alpha_i > \alpha_j, \text{then } \Upsilon_t^{\alpha_i} > \Upsilon_t^{\alpha_j} \tag{16}$$

where $|.|$ is the absolute value function and T is the number of measurement trajectories in the training set. $V_t^{\alpha_i}$ is the volume of multivariate ellipsoid with nominal coverage rate α_i at time t. The constraint in (16) is considered to avoid crossing EPRs.

3) Using the optimized $\boldsymbol{\Upsilon}$ to generate EPRs: The optimization process is conducted only once for each stochastic process using the training data. Then, to generate prediction ellipsoids for each time $t' > T$, the point forecast trajectory is generated and is used as the centre of the ellipsoids. Covariance matrix is updated and then by having the optimized $\boldsymbol{\Upsilon}$ as the scale parameters, EPRs are readily available.

The flowchart in Fig. 3 summarizes the process of designing the model to generate EPRs and Fig. 4 demonstrates how

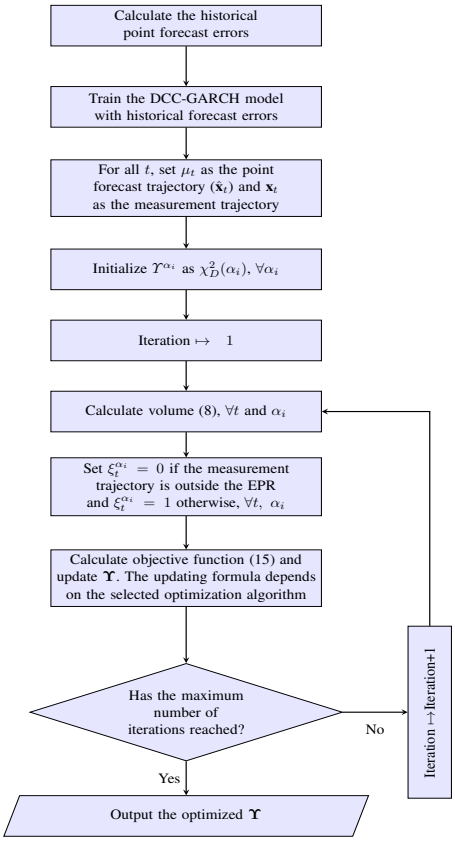


Fig. 3: A flowchart describing the process of designing the model to generate EPRs using the data up to time T.

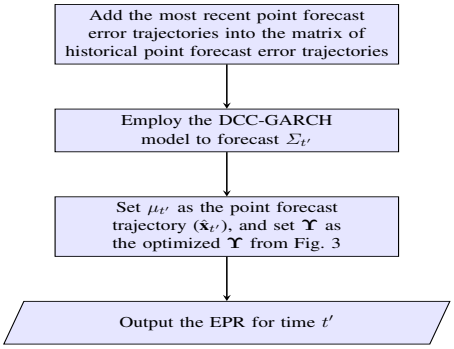


Fig. 4: A flowchart describing the process of generating EPRs for time $t' > T$.

the EPRs can be generated for the evaluation data using the designed model.

It is to be noted that for time series with time-varying dynamics, it is preferred to adapt and update the parameters of EPRs including the scale parameter continuously. The most straightforward approach to adapt to time-varying characteristics of the time series and underling processes is to estimate and optimize the parameters on sliding windows of historical data. Future works can focus on more sophisticated online learning approaches to design EPRs.

V. Results

In this section, the applicability of the proposed method for generation of skilled EPRs is investigated.

A. Data

As the basis for empirical investigations, three datasets are used here. The datasets include data for electricity price, wind and PV power. The datasets have been prepared for the Global Energy Forecasting Competition (GEFCom) 2014 and are available online [34]. For all three datasets, the resolution of data is of one hour and forecast horizons are 1- to 24-hour ahead. The datasets are briefly described below. For the full specifications, the reader is referred to [35].

- **Price data:** This dataset includes zonal load (MW) and system load forecasts (MW) as the predictors and locational marginal price ($/MW) as the predictand. The dataset covers about three years worth of data (from January 1st, 2011 to December 17th, 2013). The available data is divided into two parts including 550 and 532 days worth of data as the training and the evaluation sets, respectively. Price values are normalized by the maximum price available in the data.
- **Wind power data:** The wind data provides wind power output series from 10 wind farms in Australia. The data for the second wind farm is used in this paper. The predictors are zonal and meridional wind components forecasts at two heights, 10 and 100 m above ground level, generated by the European Centre for Medium-range Weather Forecasts (ECMWF). The predictand is wind power generation. The predictions were issued every day at midnight. The period for which both predictions and measurements are available is from January 2012 to December 2013. Training data covers the period from January 2012 to April 2013. The data from May 2013 to December 2013 is used for skill verifications. Power measurements are normalized by the nominal capacity of the corresponding wind farm.
- **PV power data:** Explanatory variables include 12 independent variables as the output of Numerical Weather Predictions (NWPs) provided by ECMWF and the predictand is PV power generation. Data for the period of April 2012 to the end of June 2014 for three zones is available. The date for the first zone is used here. Training data covers the period from April 2012 to the end of May 2013. The data from June 2013 to the end of January 2014 is used for skill evaluation. Power measurements are normalized by the nominal capacity of the corresponding PV installation.

B. Set-up

For wind power and electricity prices, the temporal correlations of 1- to 24-hour ahead predictions are studied. For PV power data, the temporal dependency of hourly PV generations from 7 am to 5 pm are taken into account.

The BOBYQA algorithm is deployed as the global search engine and the Generalized Simulated Annealing (GenSA) is used for local search. The functions “optimx” and “GenSA” in R are used for optimization. The name BOBYQA stands for Bound Optimization By Quadratic Approximation [36]. BOBYQA is an iterative and derivative-free algorithm to find the minimum value of an objective function ($F(X)$) subject to simple bounds $a \leq X \leq b$ on variables. This algorithm requires no derivatives of F explicitly and in each iteration, it

TABLE I: Point forecasts accuracy in percent form (%)

	RMSE (Train Data)	RMSE (Test Data)
Price	3.25	3.8
Wind Power	16.83	11.82
PV Power	8.73	10.68

constructs a quadratic approximation (Q) of F. In BOBYQA, the model is updated by the efficient approach of minimizing the Frobenius norm of the change to the second derivative of Q. It is to be noted that the application of the proposed approach to generate EPRs is not limited to the BOBYQA optimization solver. Employing BOBYQA is just a suggestion and the users can try other appropriate optimization solvers to optimize the objective function introduced in (15).

The window size w in rolling historical correlations and exponential smoothing methods is considered to be 500 for all simulations. The decay constant in exponential smoothing method is chosen to be 0.99. In the following, all the analyses are conducted based on the results obtained for the evaluation data.

The point and covariance matrix foretasting set-ups are as follows:

PV Power: Data is preprocessed as it is explained in [37]. The k-nearest neighbors (KNN) algorithm and Support Vector Machines (SVM) are deployed to provide forecasts. In order to predict PV power for a particular time, first, the k nearest neighbors of the explanatory variables available for that time are found. Then, those neighbors are considered as the training set to formulate the forecasting model. In other words, for each particular time, a new forecasting model is trained using the k nearest neighbors found within the historical data. k in KNN is considered to be equal to 300. The cost value and the gamma parameters in SVM are found by 5-fold cross-validation.

DCC-GARCH(1,1) rolling forecast is used to predict the time-varying covariance matrix of PV power because it is found to be more efficient than rolling historical correlations and exponential smoothing methods. A moving window of size 300 is used in the rolling estimation. An ARMA(1,0)-GARCH(1,1) is used as the univariate estimator for the conditional mean and variance in the DCC model.

Wind Power: Forecasts are provided by SVM because it is found to be more efficient than the combination of KNN and SVM. Since all wind farms are adjacent to each other, the NWPs available for the first six wind farms are used as the explanatory variables to generate forecasts for farm 2.

The set-up for covariance matrix forecasting is similar to the PV power case.

Electricty Price: Generalized linear regression is chosen because it shows better performance than SVM, ELM [20] and KNN for the price data. We found the price values less volatile comparing to PV and wind power data. The covariance matrix of price data varies very slowly in time. Therefore, here the exponential smoothing method is found to be more efficient to forecast the covariance matrix.

The point forecast accuracy in terms of Root Mean Score Error (RMSE) for all three datasets is summarized in Table I.

C. EPRs visualization

In order to visualize the EPRs, two dimensional prediction ellipsoids are generated for all three datasets. For the electricity price data, prediction ellipsoids describing the joint uncertainty of the price at 8:00 pm and 9:00 pm are generated. For wind power, the bivariate prediction ellipsoids are obtained for 3:00 am and 4:00 am while for PV power data they are generated for 2:00 pm and 3:00 pm.

In Fig. 5, 19 EPRs with probabilities ranging from 0.05 to 0.95 by 0.05 increments, for three randomly selected days are illustrated. One can notice that as the nominal coverage rates of the ellipsoids increase, the EPRs become larger. The blue dotted line in Fig. 5 for wind power describes the actual generation limits, bounded between 0 and 1 pu. When using prediction ellipsoids as constraints in interval or robust optimization, these practical limits should be added to the optimization framework as the additional constraints.

As is shown in Fig. 5, although 95% Gaussian ellipsoids for PV and wind power are smaller than the 95% EPRs, 5% Gaussian ellipsoids are larger than EPRs with the same nominal probability. This happens because Gaussian ellipsoids for these variables tend to underestimate uncertainty for higher nominal coverage rates and overestimate it for lower nominal coverage rates.

In Figs. 6 and 7, 95% and 70% EPRs along with robust ellipsoids with budget uncertainty $\Upsilon = D^2 = 4$, and 95% and 70% Gaussian ellipsoids for PV power and price data are depicted. By comparing the EPRs in Figs. 5, 6, it is clear that the shape, rotation and the ratio of semi-major to semi-minor axes of the ellipsoids vary depending on the stochastic process of interest as well as the underling non-stationery uncertainty level in various days. As one can notice, the sizes of the price EPRs for different days are very close while the those of PV EPRs vary in various days. This happens because the covariance matrix of the electricity price forecasting errors changes very slowly in time while the rate is much faster for that of the PV power. The robust ellipsoids with $\Upsilon = 4$ give the observed coverage rates of 90%, 82% and 82% for the price, wind and PV power data, respectively.

For the robust ellipsoids, as the dimension of multivariate random variable increases, the ratio of their volumes to those of EPRs increases. For bivariate case as $D^2 = 4$ is small and close to $\chi^2_2(0.95) = 5.6$, the differences between the areas of the ellipsoids are not very large. However, for $D = 24$, $\chi^2_{24}(0.95) = 36.4$ comparing to $24^2 = 576$, the difference between volumes can be substantially larger.

D. Skill verification and evaluation

To get a sense of the values of the optimized scale parameters, Fig. 8 is provided. In this figure, the optimized Υ for dimensions 24, 11 and 2 are shown. The maximum Υ^α for dimension 24 is about 70 which is much lower than 24^2.

Comparing Υ for various coverage rates shows as it is expected, the ellipsoids with higher nominal coverage rates are more conservative and larger than those with lower nominal coverage rates. Based on (8), with the same location parameter (centre) and covariance matrix, the value of Υ intensely impacts on the volume of a multidimensional ellipsoid.

In order to evaluate the robustness and reliability of the EPRs, Fig. 9 is provided. In this figure, calibration of the EPRs

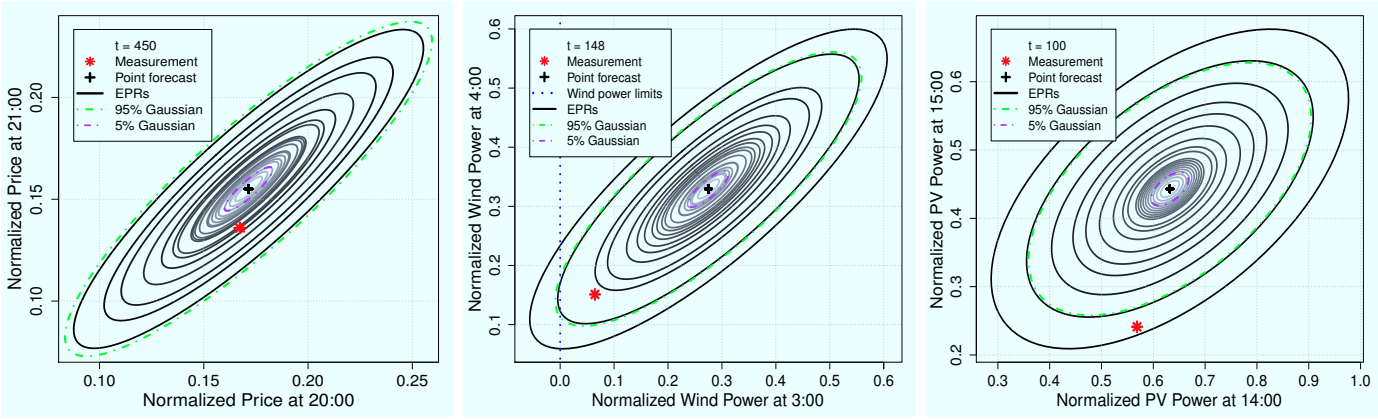


Fig. 5: 19 Optimal EPRs with probabilities ranging from 0.05 to 0.95 by 0.05 increments (from the lightest to the darkest), for three randomly selected days from the evaluation data of electricity prices, wind and PV power. Character t denotes the day number in the evaluation datasets.

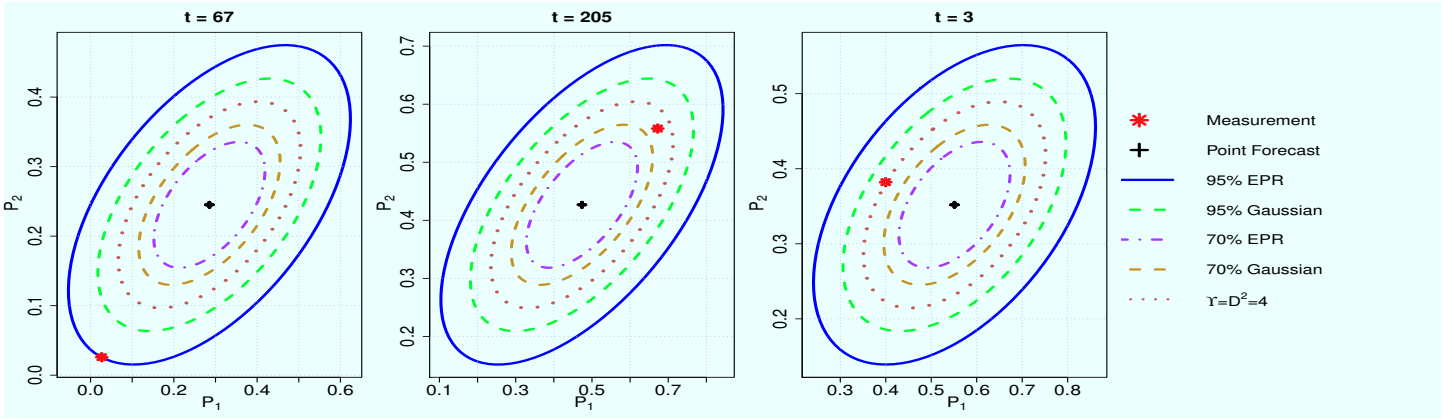


Fig. 6: PV power: Prediction ellipsoids of dimension two for three days from the evaluation dataset. P_1 and P_2 represent normalized predicted PV power for 14:00 and 15:00, respectively. Character t denotes the day number in the evaluation dataset.

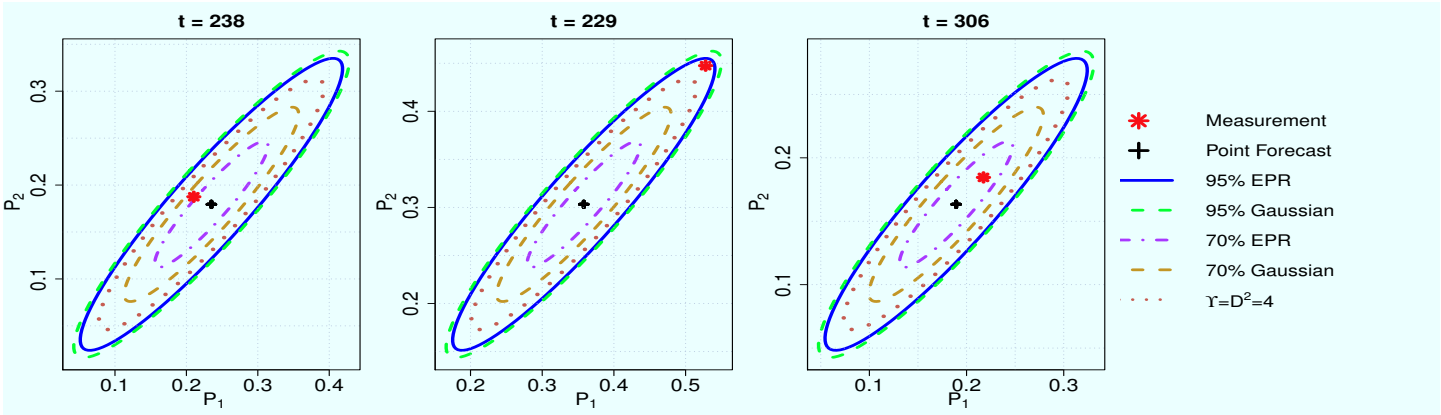
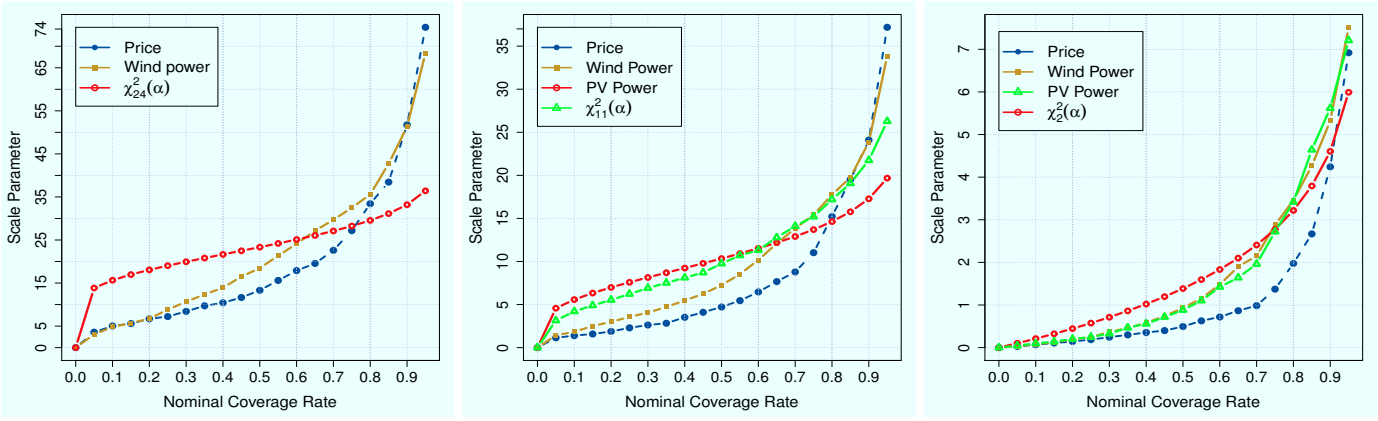


Fig. 7: Price: Prediction ellipsoids of dimension two for three randomly selected days from the evaluation dataset. P_1 and P_2 represent normalized predicted prices of energy for 20:00 and 21:00, respectively. Character t denotes the day number in the evaluation dataset.

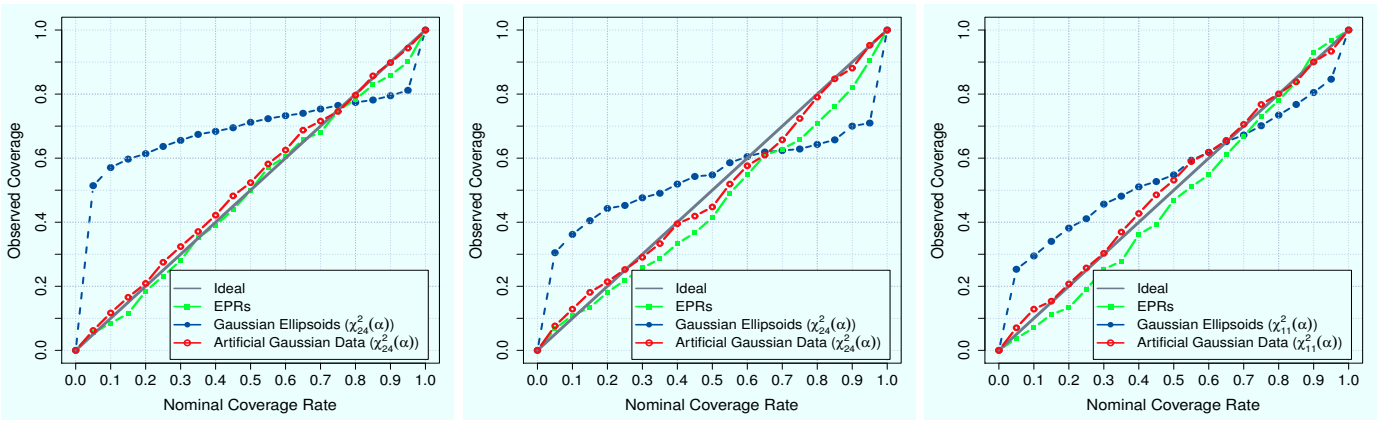
along with that of the Gaussian ellipsoids obtained for wind power and prices of dimension 24 and PV power of dimension 11 (from 7 am to 5 pm) are illustrated.

The artificial Gaussian data in Fig. 9 are generated by random draw from multivariate Gaussian distributions with the same location parameter μ_t and scale Σ_t, $\forall t$ as those used for estimating the EPRs. The scale parameters of the predicted ellipsoids for this data are equal to $\chi^2_D(\alpha_i)$, $\forall \alpha_i$. The highly calibrated ellipsoids fitted to the artificial Gaussian data in Fig. 9 reveals that if the multivariate random process is normally



(a) Dimension 24 (b) Dimension 11 (c) Dimension 2

Fig. 8: The scale parameters for the prediction ellipsoids with nominal coverage rates ranging from 0.05 to 0.95 by 0.05 increments.



(a) Price, dimension 24 (b) Wind Power, dimension 24 (c) PV Power, dimension 11

Fig. 9: Calibration of the prediction ellipsoids with nominal coverage rates ranging from 0.05 to 0.95 by 0.05 increments.

distributed, the prediction ellipsoids can be fully characterized by (2).

Looking at Fig. 9, it is inferred that the EPRs present close to ideal calibration by making a compromise between robustness (calibration) and performance (conservativeness) based on (15). The Gaussian ellipsoids tend to overestimate joint uncertainty for low nominal coverage rates but underestimate it for higher nominal coverage rates. This also can be inferred from Fig. 8, where the scale parameters given by $\chi^2_D(\alpha)$ for the Gaussian ellipsoids with lower coverage rates are much larger than those of EPRs with the same coverage rates while the relationship is opposite for the higher coverage rates. From Fig. 9, one can easily perceive very low robustness of the Gaussian prediction ellipsoids. The maximum deviations as high as 47% makes the Gaussian ellipsoids very unreliable to be used in decision-making as they lead to biased analyses. On the other hand, the EPRs generated for all three datasets offer reasonably high calibration.

In Table II, the calculated skill scores are given for three dimensions. Much better skill scores of EPRs with respect to the Gaussian ellipsoids confirm higher predictive performance of the proposed ellipsoids in terms of both conservativeness and probability guarantees.

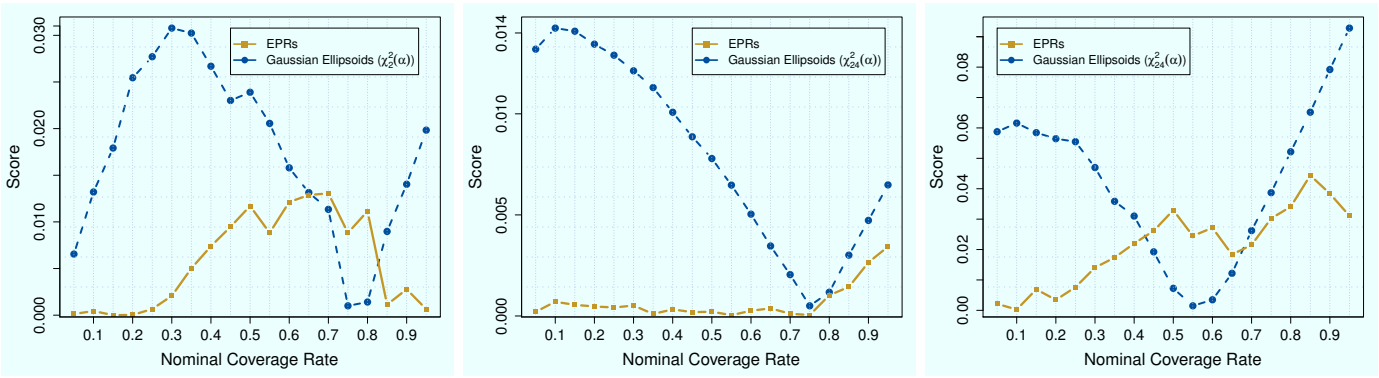
Fig. 10 shows the skill score for the EPRs of dimensions 2 and 24 for PV power, prices and wind power. As one can observe in this figure, for most of the nominal coverage rates the EPRs are more skilled than the Gaussian prediction ellipsoids.

TABLE II: Scores of EPRs and Gaussian prediction ellipsoids

		D=2	D=11	D=24
Price	**EPRs**	0.119	0.085	0.069
	Gaussian	0.914	1.034	0.794
Wind	**EPRs**	1.023	1.369	2.119
	Gaussian	1.483	4.228	4.223
PV	**EPRs**	0.570	0.562	-
	Gaussian	1.745	2.411	-

VI. Conclusion

We propose an approach to construct EPRs with predefined probability levels and optimal conservativeness. These multivariate ellipsoidal uncertainty sets provide essential information for the problems which are temporally/spatially coupled. In order to verify the applicability of the proposed method in characterizing multivariate uncertainty information for the stochastic processes with different underling stochasticity, three different datasets including data for wind and PV power and electricity prices are deployed. It is proposed to use exponential smoothing method to estimate the covariance matrix for



(a) PV Power, dimension 2 (b) Price, dimension 24 (c) Wind Power, dimension 24

Fig. 10: The skill scores of the prediction ellipsoids with nominal coverage rates ranging from 0.05 to 0.95 by 0.05 increments.

those stochastic processes like electricity price with either constant or slow-moving covariance matrix. DCC-GARCH model is preferred for those stochastic processes like wind and PV power with time-varying covariance. The simulation results showed that for all three case studies, Gaussian ellipsoids do not characterize the inherent joint uncertainty. They present very low calibration as a result of either overestimation or underestimation of the uncertainty level. This work provides a comprehensive framework for both generation and evaluation of ellipsoidal uncertainty sets which have been used in robust optimization. The proposed scheme is able to track and predict the existing uncertainty level in time and characterize the EPRs such that they provide the desired probability level with optimal sharpness. The proposed skill score is used to evaluate the predictive performance of the EPRs. The results confirm that the proposed approach is able to generate EPRs with acceptable reliability and conservativeness. The proposed framework can be applied to a variety of the decisions-making problems which involve correlated random variables.

APPENDIX A
ESTIMATING VOLUME OF EPRS FOR BOUNDED VARIABLES

The idea of estimating volumes of EPRs is to generate N random samples of dimension D in the feasible range and then calculate the proportion of the samples which lie in the ellipsoids. If the feasible limits in all dimensions are the same, then the feasible range forms a hyper-cube. The estimated volume (V^e) of the part of an ellipsoid inscribed in the feasible hyper-cube is given by

$$V^e = N' V^c / N \tag{17}$$

with N' as the number of D-dimensional trajectories (samples) enveloped by the ellipsoid and V^c is the volume of the bounded hyper-cube.

The Monte Carlo method converges very slowly and it requires a large N to allow for a reasonable estimation. To increase computational efficiency, one can generate the random samples from a smaller geometry enclosing the ellipsoid. In [17], it is suggested to generate samples from a hyper-cube with edges of equal to twice the largest semi-axis of the ellipsoid as

$$L = 2\max\{\lambda_i^{-1/2} \quad i = 1, ..., D\} \tag{18}$$

with $\lambda_i, \quad i = 1, ..., D$ as the eigenvalues of $\Sigma^{-1}/\Upsilon^\alpha$. In this case, V^c in (17) is equal to $(L)^D$. The idea is illustrated in Fig. 11 for an ellipse with S_1 and S_2 as its semi-minor and semi-major axes. Although generating samples in $(L)^D$ hyper-cube reduces the computational burden, still the method is computationally extensive specially in higher dimensions. The Monte Carlo method will be improved if the samples are generated from the smallest hyper-rectangular circumscribed the ellipsoid. To find the minimum-volume hyper-rectangular, it is sufficient to find the outermost point for each coordinate of the ellipsoid.

If $\mathbf{x}$ is located on the ellipsoidal surface, it satisfies $(x-\mu)^t\Sigma^{-1}(x-\mu) = \Upsilon^\alpha$. For $\mathbf{x}$, outward normal is pointing towards the direction

$$\nabla\left[(\mathbf{x}-\mu)^t\Sigma^{-1}(\mathbf{x}-\mu) - \Upsilon^\alpha\right] \propto \Sigma^{-1}(x-\mu) \tag{19}$$

In order to maximize $x \cdot n = n^t x$ along direction n, it is required to have

$$\Sigma^{-1}(\mathbf{x}-\mu) \propto n \quad \Longleftrightarrow \quad \mathbf{x} - \mu = \kappa\Sigma n \quad \text{for some } \kappa > 0 \tag{20}$$

Substituting $\kappa\Sigma n$ into the equation of the ellipsoidal surface gives

$$\Upsilon^\alpha = (\mathbf{x}-\mu)^t\Sigma^{-1}(\mathbf{x}-\mu) = \kappa^2 n^t\Sigma\Sigma^{-1}\Sigma n = \kappa^2 n^t\Sigma n$$
$$\Longrightarrow \quad \kappa^2 = \frac{\Upsilon^\alpha}{n^t\Sigma n} \tag{21}$$

Therefore, the point which maximizes $\mathbf{x}.n$ is given by

$$\mathbf{x}^{\max} = \mu + \sqrt{\frac{\Upsilon^\alpha}{n^t\Sigma n}}\Sigma n \tag{22}$$

The extreme values of the ellipsoid in dimension i are given as

$$x_i^{\max\,/\,\min} = \mu_i \pm \sqrt{\Upsilon^\alpha \Sigma_{ii}} \quad \forall i \tag{23}$$

with Σ_{ii} as the i^{th} diagonal element of Σ.

The length of each edge of the proposed minimum-volume hyper-rectangular is $L_i = x_i^{\max} - x_i^{\min}$ and the V^c can be calculated by multiplication of the length of D sides of the hyper-rectangular. In Fig. 11, the proposed hyper-rectangular is depicted for a typical EPR. The difference between the volume of the hyper-cube proposed in [17] and the hyper-rectangular proposed here can be large in higher dimensions. To examine the efficiency of the proposed volume estimation

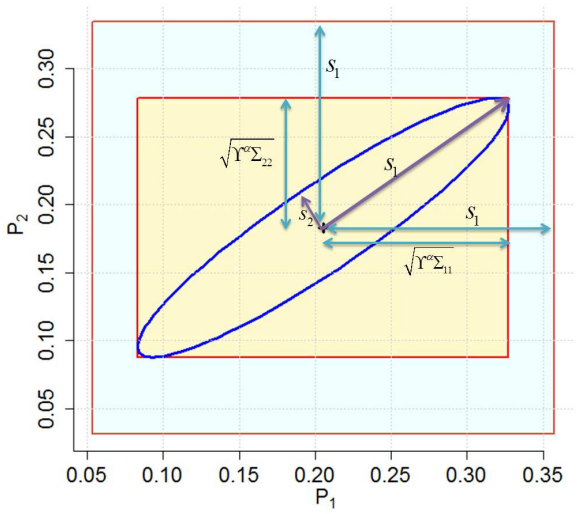


Fig. 11: A typical prediction ellipsoid of dimension two along with a hypercube and a hyper-rectangular enclosed it.

TABLE III: Absolute deviation percentage of estimated volumes of ellipsoids

	N=5,000	N=10,000	N=20,000	N=50,000	N=100,000
ADP (%)	0.614	0.550	0.297	0.207	0.095
Simulation Time (s)	0.0008	0.0028	0.0094	0.0079	0.0174

method explained in Appendix A, the reader is referred to Appendix B.

APPENDIX B
SIMULATION RESULTS OF VOLUME ESTIMATION FOR BOUNDED RANDOM VARIABLES

Wind and PV power both are double-bounded random variables between zero and 1 pu. Electricity price also can have upper and lower limits depending on power market regulations. As shown in Fig. 5, there is a chance that EPRs exceed the practical limits. Here, as explained in Appendix A, the volume of intersection of the EPRs and the polyhedrons are estimated using Monte Carlo method. By following [17], a metric named Absolute Deviation Percentage (ADP) is used to examine the efficiency of the volume estimation approach. ADP is given by

$$ADP = \frac{|V_N^c - V_{N+\triangle N}^c|}{V_N^c} \tag{24}$$

with V_N^c as the estimated volume with N random samples and $\triangle N = N/100$. ADP values along with system simulation time for five sets of random samples are given in Table III. In [17], $N = 1200$, $N = 20,000$ and $N = 500,000$ resulted in ADP equal to 0.952, 0.322 and 0.116, respectively. Comparing these values with those reported in Table III, one can conclude that the smaller circumscribed polyhedron proposed in Appendix A leads to a more efficient estimation of volumes of EPRs. The results in Table III are based on estimation of 90% EPRs.

REFERENCES

[1] J. M. Morales, A. J. Conejo, H. Madsen, P. Pinson, and M. Zugno, *Integrating renewables in electricity markets: Operational problems*. Springer Science & Business Media, 2013, vol. 205.

[2] Y. Zhang, J. Wang, and X. Wang, "Review on probabilistic forecasting of wind power generation," *Renewable and Sustainable Energy Reviews*, vol. 32, pp. 255–270, 2014.

[3] J. Tastu, "Short-term wind power forecasting: probabilistic and space-time aspects," *PhD thesis*, 2013.

[4] F. Golestaneh, H. B. Gooi, and P. Pinson, "Generation and evaluation of space–time trajectories of photovoltaic power," *Applied Energy*, vol. 176, pp. 80–91, 2016.

[5] H. Zhang and P. Li, "Chance constrained programming for optimal power flow under uncertainty," *IEEE Transactions on Power Systems*, vol. 26, no. 4, pp. 2417–2424, 2011.

[6] A. A. Sousa, G. L. Torres, and C. A. Canizares, "Robust optimal power flow solution using trust region and interior-point methods," *IEEE Transactions on Power Systems*, vol. 26, no. 2, pp. 487–499, 2011.

[7] R. A. Jabr, "Adjustable robust opf with renewable energy sources," *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 4742–4751, 2013.

[8] L. Wu, M. Shahidehpour, and Z. Li, "Comparison of scenario-based and interval optimization approaches to stochastic scuc," *IEEE Transactions on Power Systems*, vol. 27, no. 2, pp. 913–921, 2012.

[9] R. J. Bessa, "From marginal to simultaneous prediction intervals of wind power," in *Intelligent System Application to Power Systems (ISAP), 2015 18th International Conference on*. IEEE, 2015, pp. 1–6.

[10] D. Kolsrud, "Time-simultaneous prediction band for a time series," *Journal of Forecasting*, vol. 26, no. 3, pp. 171–188, 2007.

[11] J. S.-H. Li and W.-S. Chan, "Simultaneous prediction intervals: An application to forecasting us and canadian mortality," 2011.

[12] D. Bertsimas, D. B. Brown, and C. Caramanis, "Theory and applications of robust optimization," *SIAM review*, vol. 53, no. 3, pp. 464–501, 2011.

[13] A. Chassein and M. Goerigk, "Min-max regret problems with ellipsoidal uncertainty sets," *arXiv preprint arXiv:1606.01180*, 2016.

[14] Y. Guan and J. Wang, "Uncertainty sets for robust unit commitment," *IEEE Transactions on Power Systems*, vol. 3, no. 29, pp. 1439–1440, 2014.

[15] A. A. Kurzhanskiy and P. Varaiya, "Ellipsoidal techniques for reachability analysis of discrete-time linear systems," *IEEE Transactions on Automatic Control*, vol. 52, no. 1, pp. 26–38, 2007.

[16] A. B. Kurzhanski and P. Varaiya, "On ellipsoidal techniques for reachability analysis. part i: external approximations," *Optimization methods and software*, vol. 17, no. 2, pp. 177–206, 2002.

[17] P. Li, X. Guan, J. Wu, and X. Zhou, "Modeling dynamic spatial correlations of geographically distributed wind farms and constructing ellipsoidal uncertainty sets for optimization-based generation scheduling," *IEEE Transactions on Sustainable Energy*, vol. 6, no. 4, pp. 1594–1605, 2015.

[18] P. Pinson, "Estimation of the uncertainty in wind power forecasting," 2006.

[19] S. Bofinger, A. Luig, and H. Beyer, "Qualification of wind power forecasts," in *2002 Global Windpower Conference, Paris*, vol. 2, no. 5, 2002.

[20] F. Golestaneh, P. Pinson, and H. B. Gooi, "Very short-term nonparametric probabilistic forecasting of renewable energy generationwith application to solar energy," *IEEE Transactions on Power Systems*, vol. 31, no. 5, pp. 3850–3863, 2016.

[21] H. Bludszuweit, J. A. Domínguez-Navarro, and A. Llombart, "Statistical analysis of wind power forecast error," *Power Systems, IEEE Transactions on*, vol. 23, no. 3, pp. 983–991, 2008.

[22] B.-M. Hodge and M. Milligan, "Wind power forecasting error distributions over multiple timescales," in *Power and Energy Society General Meeting, 2011 IEEE*. IEEE, 2011, pp. 1–8.

[23] A. Luig, S. Bofinger, and H. G. Beyer, "Analysis of confidence intervals for the prediction of regional wind power output," in *Proceedings of the European wind energy conference, Copenhagen, Denmark*, 2001, pp. 725–728.

[24] B.-M. Hodge, D. Lew, M. Milligan, H. Holttinen, S. Sillanpää, E. Gómez-Lázaro, R. Scharff, L. Söder, X. G. Larsén, G. Giebel *et al.*, "Wind power forecasting error distributions: An international comparison," in *11th Annual International Workshop on Large-Scale Integration of Wind Power into Power Systems as well as on Transmission Networks for Offshore Wind Power Plants Conference*, 2012.

[25] R. Weron *et al.*, "Heavy tails and electricity prices," in *The Deutsche Bundesbank?s 2005 Annual Fall Conference (Eltville)*, 2005.

[26] H. Zhou, B. Chen, Z. Han, and F. Zhang, "Study on probability distribution of prices in electricity market: A case study of zhejiang province, china," *Communications in Nonlinear Science and Numerical Simulation*, vol. 14, no. 5, pp. 2255–2265, 2009.

[27] T. Jónsson, P. Pinson, H. Madsen, and H. A. Nielsen, "Predictive densities for day-ahead electricity prices using time-adaptive quantile regression," *Energies*, vol. 7, no. 9, pp. 5523–5547, 2014.
[28] R. Engle, "Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models," *Journal of Business & Economic Statistics*, vol. 20, no. 3, pp. 339–350, 2002.
[29] S. B. Pope, "Algorithms for ellipsoids," *Cornell University Report No. FDA*, pp. 08–01, 2008.
[30] J. M. Morales, A. J. Conejo, H. Madsen, P. Pinson, and M. Zugno, *Integrating Renewables in Electricity Markets: Operational Problems*. Springer, 2014.
[31] V. Zakamulin, "A test of covariance-matrix forecasting methods," *The Journal of Portfolio Management*, vol. 41, no. 3, pp. 97–108, 2015.
[32] E. Orskaug, "Dcc-garch model-with various error distributions," 2009.
[33] R. F. Engle and K. Sheppard, "Theoretical and empirical properties of dynamic conditional correlation multivariate garch," National Bureau of Economic Research, Tech. Rep., 2001.
[34] Global energy forecasting competition 2014 probabilistic solar power forecasting. [Online]. Available: https://crowdanalytix.com
[35] T. Hong, P. Pinson, S. Fan, H. Zareipour, A. Troccoli, and R. J. Hyndman, "Probabilistic energy forecasting: Global energy forecasting competition 2014 and beyond," *International Journal of Forecasting*, vol. 32, no. 3, pp. 896–913, 2016.
[36] M. J. Powell, "The bobyqa algorithm for bound constrained optimization without derivatives," *Cambridge NA Report NA2009/06, University of Cambridge, Cambridge*, 2009.
[37] J. Huang and M. Perry, "A semi-empirical approach using gradient boosting and k-nearest neighbors regression for gefcom2014 probabilistic solar power forecasting," *International Journal of Forecasting*, vol. 32, no. 3, pp. 1081–1086, 2016.