

Polyhedral Predictive Regions For Power System Applications

Faranak Golestaneh, Pierre Pinson, *Senior Member, IEEE*, and Hoay Beng Gooi, *Senior Member, IEEE*

Abstract—Despite substantial improvement in the development of forecasting approaches, conditional and dynamic uncertainty estimates ought to be accommodated in decision-making in power system operation and market, in order to yield either cost-optimal decisions in expectation, or decision with probabilistic guarantees. The representation of uncertainty serves as an interface between forecasting and decision-making problems, with different approaches handling various objects and their parameterization as input. Following substantial developments based on scenario-based stochastic methods, robust and chance-constrained optimization approaches have gained increasing attention. These often rely on polyhedra as a representation of the convex envelope of uncertainty. In the work, we aim to bridge the gap between the probabilistic forecasting literature and such optimization approaches by generating forecasts in the form of polyhedra with probabilistic guarantees. For that, we see polyhedra as parameterized objects under alternative definitions (under L_1 and L_∞ norms), the parameters of which may be modelled and predicted. We additionally discuss assessing the predictive skill of such multivariate probabilistic forecasts. An application and related empirical investigation results allow us to verify probabilistic calibration and predictive skills of our polyhedra.

Index Terms—Probabilistic forecasting, box uncertainty sets, polyhedron, robust optimization, chance-constrained optimization

I. INTRODUCTION

THE VARIABILITY AND limited predictability of renewable power generation have introduced new challenges into power systems. With a large-scale uncertain generation, in order to reduce the gap between fail-safe and economical solutions of operational problems, advancement in two areas is essential. First, the development of highly scalable optimization techniques capable of accommodating considerable degree of uncertainty is required. Second, it is of utmost importance to develop adequate and high-quality representations of the uncertainties involved to be used as input to the aforementioned optimization techniques [1], [2].

Practitioners mostly use so-called deterministic or point forecasts as input to decision making today. These comprise single-valued prediction for the future realization of a variable of interest, disregard the actual range of potential outcomes. However, the solution to an optimization problem in a deterministic setup may be highly sensitive to small perturbations of uncertain quantities. Hence, ignoring uncertainty can result in suboptimal or infeasible solutions in practice [3].

Uncertainty forecasts can be represented in various forms such as scenario, probabilistic and ramp forecasts [4]. Since wind and PV power both show high cross-correlation in time and space, more recently, forecasting spatial/temporal scenarios has been of interest. For example, temporal uncertainty forecast is a key requirement for multi-period operational problems such as unit-commitment and state of charge of energy storage [5]. Stochastic programming as one of the most common optimization techniques in power systems applications uses scenarios as inputs to find optimal solutions in

uncertain environments [6], [7]. However, stochastic programming holds a number of pitfalls in a practical context including heavy computational burden and the need for hard-to-obtain probability distributions [8].

The issues with stochastic programming motivates to move towards more recent approaches to optimization under uncertainty, namely robust, chance-constrained and interval optimization. Recently, these optimization techniques have been deployed in power systems applications [9]–[11]. For these classes of decision-making problems, the required uncertainty representation takes the form of prediction regions rather than scenarios. Robust optimization is a computationally viable methodology providing solutions deterministically immune to any realization of uncertainty within a defined uncertainty set (another term for prediction regions). Interval optimization derives optimistic and pessimistic solutions based on boundaries of prediction regions. In chance-constrained optimization, the uncertainty sets give a probability guarantee for the coverage of observations from the stochastic process considered.

Prediction regions in univariate case, e.g. modeling uncertainty of a single wind farm in a particular time, can be adequately addressed by prediction intervals [4]. However, when modeling temporal/spatial or multivariable correlations is of interest, prediction regions take the form of multivariate ellipsoids, boxes and polyhedra. We refer to uncertainty sets as prediction regions to emphasize on the fact that they are predictions in nature.

Although the multivariate prediction regions have been used in several optimization applications, the literature has been almost silent on how to efficiently generate and evaluate them. The parameters of multivariate prediction regions are simply chosen based on assumptions or by trial-and-error without verification of those assumptions in practical applications. Uncertainty sets are constructed based on a Gaussian assumption in [11] for nodal load and in [10], [12] for wind power. The inadequacy of a Gaussian assumption in describing uncertainty of wind and PV power is discussed in e.g. [13]. A parameter named uncertainty budget is used to control the size and conservativeness of wind power uncertainty sets in the form of ellipsoids in [14] and in the form of polyhedra in [15]. As a different approach, in [16] convex hull of spatial/temporal scenarios is defined as a prediction region of wind/PV. In [17], temporal scenarios are used as input to produce multivariate prediction intervals (MPIs) to characterize the dependency of wind power forecast errors over a time horizon.

Robust optimization tends to produce conservative solutions. The conservativeness of a robust solution is directly linked to the size of uncertainty sets [7]. However, controlling size of uncertainty sets is not a trivial task to be determined arbitrarily. As any other type of prediction, uncertainty prediction should provide a certain level of required performance. Multivariate prediction regions are assessed based on their calibration and sharpness. Calibration is linked to conserva-

tiveness and it shows how close the empirical coverage rate of a prediction region is to its nominal one. In contrast to [10]–[12], [14]–[16], we emphasize on generating prediction regions with predefined coverage rates. This helps the decision-maker to know in advance what the degree of constraint violation is upon obtaining the solution of the optimization problem. Sharpness relates to how small the spread of uncertainty is for the required probability guarantee. Too large prediction regions increase cognitive load.

In [13], we proposed a framework to produce skilled ellipsoidal prediction regions. However, various decision-making problems demand for different forms of uncertainty characterization. For example, the robust counterpart of a linear programming problem with polyhedral uncertainty sets is a linear programming problem while the same with the ellipsoidal uncertainty sets is a Second Order Cone Programming (SOCP) problem [18]. Although, SOCP problems are convex and computationally tractable, their nonlinearity can be a practical drawback. Consequently, in this work we focus on multivariate prediction polyhedra. Our underlying motivation is to propose a data-driven approach capable of generating highly skilled prediction polyhedra. We study evaluation methodologies for verification of the proposed methods using real data. Two formulations for prediction polyhedra are developed. In addition, due to recent interests in prediction convex hulls [16], their relevance and limitations are discussed and a verification framework for their quantitative assessment is developed. Because any forms of multivariate prediction is prone to be affected by outliers, we propose an idea to make convex hulls more robust to outliers. The robustness of the prediction regions to outliers is also examined and compared. All techniques output convex polyhedra and suit the requirements of robust and chance-constrained optimization. Also, theoretically they can be employed for both spatial and temporal uncertainty prediction. Their performances in practice, however, will be assessed over empirical results in Section V. The efficiency of the proposed frameworks is evaluated for wind and PV power. Temporal and spatial prediction polyhedra of dimensions 2, 3, 6, 12 and 24 with the probability of 5% to 95% in 5% increments are generated and evaluated.

The rest of the paper is organized as follows: in Section II, the proposed methodology and formulations to generate prediction polyhedra are discussed. The proposed skill assessment techniques are provided in Section III. The framework to estimate the parameters of the proposed prediction regions is explained in Section IV. Section V contains the empirical results and finally concluding remarks are given in Section VI.

II. METHODOLOGY

Due to growing interest in characterizing uncertainty information in forms of polyhedra and multivariate boxes, in this section, four frameworks to produce such prediction geometries are proposed.

A. Simple Prediction Polyhedra

At every time step t , one aims at predicting the random variable, e.g. wind/PV power, for future times $t+1$, $t+2$, ..., $t+K$ at Z contiguous locations. Denote \mathbf{X} as an uncertain variable of dimension $D = K \times Z$, $\mathbf{X}_t = [X_{t+1}, \dots, X_{t+D}]$. Denote $\mu = E(\mathbf{X})$ as the expected value \mathbf{X} and $\Sigma = E[(\mathbf{X}-\mu)(\mathbf{X}-\mu)^\top]$

as its the covariance matrix. Inspired by [18], we propose the following two formulations for prediction polyhedra.

$$P_{t,\alpha}^1 := \{\mathbf{X} \mid \|\Lambda_t(\mathbf{x}_t - \mu_t)\|_1 \leq \Gamma_t^\alpha\} \quad (1)$$

$$P_{t,\alpha}^\infty := \{\mathbf{X} \mid \|\Lambda_t(\mathbf{x}_t - \mu_t)\|_\infty \leq \Delta_t^\alpha\} \quad (2)$$

where α is the nominal coverage rate of prediction polyhedra, Δ and Γ are called scale or robust parameters. With the assumption that Σ^{-1} is a symmetric and positive definite matrix, Λ as the Cholesky decomposition of Σ^{-1} is an upper triangular matrix with positive diagonal elements. For vector $\mathbf{x} \in \mathbb{R}^{D \times 1}$, $\|\mathbf{x}\|_1$ denotes the first norm as $\sqrt{\sum_{d=1}^D |x_d|}$ and $\|\mathbf{x}\|_\infty$ denotes the infinity norm given by $\max_{d=1, \dots, D} |x_d|$. Henceforth, the upper case letters symbolize random variables while lower case letters express their realizations.

The polyhedron given by (1) is inscribed in an ellipsoid defined by the following formulation.

$$E_{t,\alpha} := (\mathbf{x}_t - \mu_t)^\top \Sigma^{-1} (\mathbf{x}_t - \mu_t) \leq (\Gamma_t^\alpha)^2 \quad (3)$$

The predictive performance of the uncertainty sets in forms P^1 , P^∞ is directly linked to how accurate and optimal their predicted parameters are. The parameters to be predicted include a location parameter μ (mean vector), a shaping parameter Σ (covariance matrix), and scaling parameters Γ^α and Δ^α (being a function of the nominal coverage rate). It is worth noting that even though the first and the second-order moment information (i.e., mean and covariance) are classically used for Gaussian objects, considering them as a basis for defining polyhedra does not necessarily means one can assume the underlining distribution is Gaussian.

In robust optimization literature, the scale parameter is commonly known as the uncertainty budget and it controls the conservativeness. The uncertainty budget is determined by the user arbitrarily based on his aversion to uncertainty. One does not expect to get uncertainty sets with predefined probability levels based on the common approaches available for determination of the uncertainty budget [14], [18].

Assuming equal values for Γ and Δ in (1)-(3), typical P^1 , P^∞ and E with probability level of 85% are illustrated in Fig. 1. In Fig. 1 the predicted scale parameter is 2.210 while the location μ , and shape Σ for P^1 , P^∞ and E are

$$\Sigma = \begin{bmatrix} 0.01762222 & 0.01135601 \\ 0.01135601 & 0.01265258 \end{bmatrix}, \quad \mu = [0.370 \quad 0.405] \quad (4)$$

It is to be emphasized that Γ and Δ are not expected to be equal in general. In Section IV, the proposed ideas to determine the correct values for Γ and Δ are explained. What Fig. 1 illustrates is that in case Γ and Δ take equal values in (1)-(3), how P^1 , P^∞ and E relate to one another. Robust or interval optimization go along the faces/edges to find the optimal solution. From Fig. 1, it can be inferred that P^1 and P^∞ impose similar computational cost in optimization because they actually have equal number of edges/facets with a difference that P^1 tends to be sharper. As shown in Fig. 1, the measurement is included in all P^1 , P^∞ and E . If having many more sample observations, one would expect close to 85% of the observations to be covered by the prediction regions.

B. Prediction Convex Hulls

The convex hull of a set of points, S , is the smallest convex set containing all the points. The idea is to find the convex hull of spatial/temporal scenarios [19]. Spatial/temporal scenarios

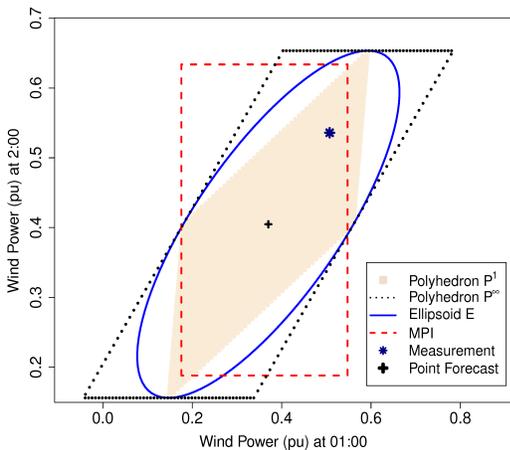


Fig. 1: Typical prediction geometries of dimension two.

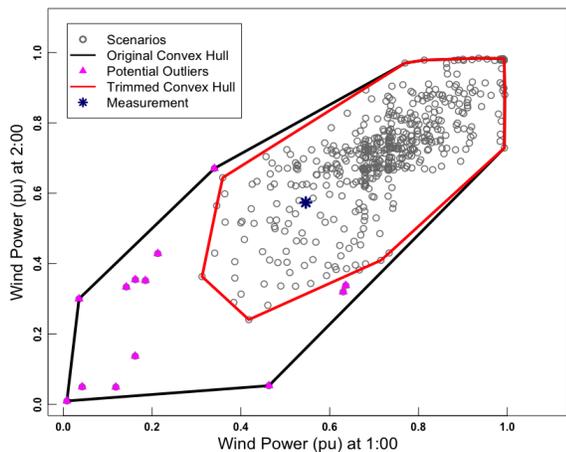


Fig. 2: The convex hull containing all the predicted scenarios. The original convex hull against the trimmed convex hull generated by excluding the potential outliers .

are generated by uniformly sampling from multivariate predictive distributions. At each time t , S scenarios are produced where each scenario is a vector of dimension D . Among few methods available to find a non-ambiguous and efficient representation of required convex hulls, we use Quickhull algorithm. Quickhull algorithm is fast and efficient in most cases and tends to perform well in practice [20]. Time complexity of this algorithm for most cases is $O(n \log n)$ and in the worst case is $O(n^2)$. Theoretically, Quickhull algorithm can work in high dimensions. For a straightforward explanation and implementation guide of Quickhull algorithm, one can refer to [21]. The original convex hull illustrated in Fig. 2 represents the convex hull of predicted temporal scenarios of wind power for a randomly selected day.

C. Trimmed Prediction Convex Hulls

As stated in subsection II-B, the predicted spatial/temporal scenarios are produced by uniformly sampling from multivariate predictive distributions. Although in the Monte Carlo-based analysis, all scenarios are considered to come with an equal likelihood of occurrence, some of them might be far from the center of cloud. We label those scenarios as outliers. Outliers are marked in Fig. 2. As can be observed, they grossly impact on the size of prediction convex hulls. Discarding the outliers

results in the trimmed convex hull in Fig. 2 which is much sharper than the original one.

Although there is a wealth of techniques available to detect outliers in univariate datasets, only limited options are at hand for multivariate data. It is to be noted that a multivariate outlier does not necessarily have to be an outlier in any of its univariate coordinates. Mahalanobis distance is one of the most widely used metrics for multivariate outlier identification. Basically, it quantifies how far away a point is from the center of the cloud, taking into account the shape of the cloud as well. Those scenarios with Mahalanobis distance larger than the critical chi-square values at a significance level of 0.001 are labeled as outliers [22].

D. Benchmark Method

The multivariate prediction literature still is in a primitive stage and there are not many data-driven benchmarks available to conduct a comparative study on the performance of the proposed techniques. Among few works available, the adjusted intervals approach is found to be a relevant benchmark [17], [23], [24]. This technique uses the marginal (univariate) prediction intervals and the multivariate scenarios as the inputs to generate Multivariate Prediction Intervals (MPIs). For the approach to generate MPIs, the reader is referred to [17]. Typical MPIs are illustrated in Fig. 1.

III. PREDICTIVE SKILL ASSESSMENT

The predictive performance of probabilistic forecasts are commonly examined based on their two properties, namely calibration and sharpness. Calibration is a joint property of forecasts and observations, and it is decided based on the statistical consistency between them. Sharpness refers to concentration of forecasts [25]. Following the probability and statistics literature, we refer to calibration as the proximity of the nominal coverage rate of a prediction region to its empirical coverage rate. The coverage rate of a prediction region is the proportion of times that the region contains materialized events (observations). Similarly, a nominal coverage rate refers to the expected coverage while an empirical coverage represents the empirical coverage of that region calculated based on real data. Sharpness is examined based on the size of prediction regions, e.g. area in dimension two, and volume in higher dimensions. The aim is to generate sharp and concentrated prediction polyhedra subjected to calibration.

A. Simple Prediction Polyhedra

As can be observed in Fig. 1, both P^1 and P^∞ are simple, convex and have few edges. In geometry, based on the definition, each vertex of a D -dimensional simple polyhedron is adjacent to exactly D edges. Robust and interval optimization go along the edges of uncertainty sets to find the optimal solution. In general, fewer number of faces is an advantage in the sense that it imposes less computation to optimization.

There is no limitation to represent uncertainty in higher dimensions in the form of P^1 and P^∞ as long as the correlation matrix Σ can be predicted and Cholesky decomposition of Σ^{-1} can be calculated. The proposed approach and formulations are competent at generating prediction polyhedra with any desired probability guarantees.

Volume: Since to the best of our knowledge there is no straightforward approach to calculate the volume of P^1 and

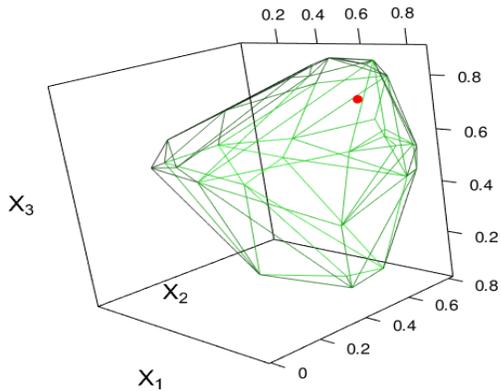


Fig. 3: A typical convex hull generated for PV power data. X_1 , X_2 and X_3 denote PV power output at zones 1, 2 and 3, respectively. The red dot represents the measurement.

P^∞ analytically, in Section V, a Monte Carlo-based technique is explained to estimate their volumes numerically.

Calibration: To evaluate calibration of prediction polyhedra, one needs to calculate the empirical coverage of each predicted polyhedron and compares that with the corresponding nominal coverage. Let $\xi_t^{\alpha_i}$ be a binary variable taking 1 if the prediction polyhedron with nominal probability α_i contains the observed value at time t and 0 otherwise. Then the empirical coverage is given by

$$\hat{\alpha}_i = \frac{1}{T} \sum_{t=1}^T \xi_t^{\alpha_i} \quad (5)$$

with T as the length of the evaluation set. P^1 and P^∞ include the measurement \mathbf{x}_t if it satisfies (1) and (2), respectively.

B. Prediction Convex Hulls

As one can notice in Fig. 2, the number of faces in convex hulls is higher than simple prediction polyhedra. This is more noticeable in higher dimensions as shown in Fig. 3. A higher number of faces/edges imposes a higher computation into optimization because as mentioned before robust or interval optimization go along the faces/edges to find the optimal solution. In addition, a major limitation of prediction convex hulls is that they cannot be generated for predefined nominal coverage rates. They just represent the smallest convex region including the predicted scenarios. When verifying them on real measurements, they can show any empirical coverage, ranging between zero to one. Prediction convex hulls comparing to P^1 and P^∞ have the complexity of generating multivariate scenarios as their input first.

Volume: One advantage of convex polyhedra is that their volumes can be calculated by subdividing them into smaller pieces. To do that a common approach is by triangulation methods where the polyhedron is decomposed into simplices. A simplex is a generalization of triangle to arbitrary dimensions. The volume of simplices can easily be computed. Eventually, the volume of polyhedron is computed by summing up the volumes of all simplices [20], [26].

Calibration: In convex geometry, given points $C = \{\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2, \dots, \hat{\mathbf{x}}_S\}$, the point $\theta_1 \hat{\mathbf{x}}_1 + \theta_2 \hat{\mathbf{x}}_2, \dots, \theta_S \hat{\mathbf{x}}_S$ is called their convex combination if $\theta_i \geq 0, i = 1, 2, \dots, S$ and $\sum_{i=1}^S \theta_i = 1$. Therefore, a convex combination of points can be viewed as a weighted average of the points, with θ as the weight of

each point in the mixture. The convex hull of set C ($\text{conv } C$) contains the arbitrary point \mathbf{y} , if \mathbf{y} is a convex combination of C [27]. We use this definition to identify if a prediction convex hull generated for time t includes the measurement recorded at the same time. $\mathbf{y} \in C$ if there is a solution for the following linear programming problem

$$\arg_{\theta} \min e^\top \theta \quad (6)$$

subject to

$$\begin{aligned} A\theta &= y \\ B\theta &= 1 \\ \theta &\geq 0 \end{aligned} \quad (7)$$

with $\mathbf{y} \in \mathbb{R}^D$, $e \in \mathbb{R}^D$ arbitrary cost vector, $B = (1, \dots, 1) \in \mathbb{R}^S$ $A = (\mathbf{x}_1, \dots, \mathbf{x}_S) \in \mathbb{R}^{D \times S}$ and $\theta \in \mathbb{R}^S$.

Following (6) and (7), one can determine if an arbitrary point \mathbf{y} is inside $\text{conv } C$ directly with no need to generate $\text{conv } C$ first. The observed coverage of prediction convex hulls can be computed according to (5) once the inclusion of each \mathbf{y} in the evaluation set is examined.

C. Multivariate Prediction Intervals

Calibration: For MPIs, the empirical coverage rate is computed by counting the number of measured scenarios which fully lie within their boundaries [17].

Volume: The volume V_t of a MPI at time t is calculated as

$$V_t = \prod_i (h_{i,t} - l_{i,t}) \quad \forall t \quad (8)$$

with $h_{i,t}$ and $l_{i,t}$ as the upper and lower bounds of the MPI at dimension i , respectively.

IV. PARAMETER ESTIMATION

The parameters of P^1 and P^∞ are determined as

u_t : The location parameter u_t is the center of prediction polyhedra and is considered to be the point forecasts for the multivariate random variable \mathbf{X} at time t . Denote $\hat{\mathbf{x}}_t = [\hat{x}_{1,t} \ \hat{x}_{2,t} \ \dots \ \hat{x}_{D,t}]$, 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. We refer to $\hat{\mathbf{x}}$ as predictions and \mathbf{x} as the measurement or materialized trajectory.

Σ_t : The Σ_t is defined as the covariance matrix of point forecast errors estimated using data up to time t . We suggest to use the Dynamic Conditional-Correlation-GARCH (GARCH-DCC) technique to predict the covariance matrix [28]. In econometrics literature, GARCH-DCC has been widely implemented and is shown to be capable of estimating time-varying covariance matrices [28], [29]. GARCH-DCC is mostly suitable for those random processes like forecast errors of renewable power generation for which the covariance matrix changes noticeably over time. In case the random process presents a slow-moving covariance matrix, rolling historical correlations and exponential smoothing as less complicated techniques can be deployed [30].

Γ_t^α : In [13], for ellipsoidal prediction regions, we proposed an approach to find the optimal scale parameters by making a compromise between volume of the ellipsoids and their calibration. To the best of our knowledge, for P^1 and P^∞ polyhedra, there is no straightforward closed form formulation to calculate the volume. Therefore, we propose a data-driven technique to find the minimum scale parameter which provides the required coverage rate over the most recent historical data. The scale parameter is updated whenever new measurements

are received. In the proposed method, a window of size ω of the most recent measurements, point forecasts and predicted Cholesky decomposition of covariance matrices are input in the following equation and Υ for those values is calculated.

$$\Upsilon_i = \{\mathbf{X} \|\Lambda_i(\mathbf{x}_i - \mu_i)^\top\|_1\} \quad i = t - \omega, \dots, t - 2, t - 1 \quad (9)$$

where μ_i is $\hat{\mathbf{x}}_i$ and $\mathbf{x}_i \forall i$ are the measured trajectories. Then, $\Upsilon_i, \forall i$ are sorted ascending. For the desired probability level α , Γ_t^α is considered as the N^{th} smallest $\Upsilon_i, \forall i$ or in other words the N^{th} element of the sorted vector, where N is

$$N = \text{round}(\omega \times \alpha) \quad (10)$$

with $\text{round}(x)$ as a function which returns the closest integer to x . Following the proposed method, Γ_t^α is updated for each t on a rolling base. This technique is based on this expectation that if prediction polyhedra envelop a window of most recent historical data, they should present a similar coverage for the future observations. After obtaining u_t, Σ_t and $\Gamma_t^\alpha, P_{t,\alpha}^1$ is readily available.

Δ_t^α : The Δ_t^α can be estimated similar to Γ_t^α as explained above, with the only difference that $\Upsilon_i, \forall i$ are calculated as

$$\Upsilon_i = \{\mathbf{X} \|\Lambda_i(\mathbf{x}_i - \mu_i)^\top\|_\infty\} \quad i = t - \omega, \dots, t - 2, t - 1 \quad (11)$$

For techniques to generate spatial/temporal scenarios, MPIs and convex hulls, the reader is referred to [19], [17] and [21], respectively.

V. EMPIRICAL RESULTS

In order to evaluate the proposed forecasting frameworks, the wind power and PV power datasets provided for the Global Energy Forecasting Competition (GEFCom) 2014 are used here. The datasets are available online [31]. We use wind power data to predict temporal dependency and PV power data to study spatial dependency. The wind power dataset includes wind power measurements of 10 wind farms in Australia. The data for farm three is used here for analysis. The data includes four explanatory variables which are zonal and meridional wind components forecasts at two heights, 10 and 100 m above ground level provided by the European Centre for Medium-range Weather Forecasts (ECMWF). Weather forecasts are issued every day at midnight. The resolution of data is of one hour and forecast horizons are 1- to 24-hour ahead. Data for January 2012 to the end of April 2013 is used to train the models while the out of sample subset covers May 2013 to December 2013. PV power data includes 12 independent variables as the output of Numerical Weather Predictions (NWP) used as predictors and PV power generation as predictand. The available data covers the period from April 2012 to the end of June 2014 for three contiguous zones. Data for April 2012 to the end of May 2013 is used to train the model and the evaluation subset covers data from June 2013 to the end of May 2014. Analysis are carried out to predict the simultaneous stochastic behavior of PV power at three zones at 12:00 pm for spatial dependency studies. Power measurements are normalized by the nominal capacity of their corresponding generation unit.

A support vector machine (SVM) from package "e1071" in R whose parameters are tuned based on 5-fold cross-validation is used to generate wind/PV power point forecasts. It yields 15.43%, 14.1% root mean square error for wind and PV power (12:00 pm only), respectively. Because all wind farms are adjacent to each other, the weather forecasts available

for the first six wind farms are used as the explanatory variables to generate forecasts for farm 3. The covariance matrices are predicted using DCC-GARCH functions from "rmgarch" package in R. Univariate quantiles with the nominal probability 2.5% to 97.5% in 2.5% increments are produced by quantile regression. 500 scenarios [19] are generated as the inputs for adjusted interval technique. The upper limit of intervals is considered to be 99.5% quantile given by quantile regression and the lower limit is considered to be zero.

Prediction polyhedra of dimensions 2, 6, 12 and 24 for wind power data are generated and evaluated. Dimension 2 includes wind power data at 01:00 am and 2:00 am. Dimension 6 covers 1- to 6-hour head predictions from 01:00 am to 6:00 am. Dimension 12 represents data from 01:00 am to 12:00 pm and dimension 24 includes all 24 hourly lead times from 01:00 am to 24:00 midnight. Prediction polyhedra produced for PV power data are of dimension 3, describing the correlated uncertainty of PV power at three zones under the study at 12:00 pm. Throughout this section, all the analysis in dimension 3 are based on spatial prediction polyhedra produced for PV power data while the results provided in other dimensions are based on temporal prediction polyhedra of wind power.

Fig. 4 shows the prediction polyhedra for three randomly selected days for out of sample data. The regions are limited to the feasible range of normalized wind power data [0,1]. Comparing MPIs with P^1, P^∞ and convex hulls, one can notice that the later ones present a correlated pattern between generation at two successive hours while MPIs show a uniform relation between them. All polyhedra have a fairly reasonable size and follow the variations in wind power generation.

In the following, we will compare the various prediction polyhedra in dimensions $D \geq 2$. The following simulations results suggest that both P^1 and P^∞ have better predictive skill than MPIs in terms of both calibration and sharpness (volume). In addition, P^1 tends to be sharper and less conservative than P^∞ . It is to be noted that one should expect to see more improvements in the area of verification of such forecasts in the future. We still have a minimum sound basis here to analyze our forecasts and conclude.

Fig. 5 reports the deviations between empirical coverage rate and nominal coverage rate of prediction polyhedra for dimensions 2, 3, 6 and 24. The nominal coverage rates ranging from 0.05 to 0.95 in 0.05 increments are included in the figure. As it is expected, UPIs fail to capture the dependent and correlated uncertainty of wind/PV power output over successive hours and at adjacent locations. The calibration and reliability of UPIs decline as the dimension increases. The calibration of MPIs is also woefully inadequate. The P^1 and P^∞ polyhedra maintain a fairly stable calibration in all dimensions and for all nominal coverage rates. Convex hulls are not covered in this figure because as discussed in Section III, they do not provide prediction regions with predetermined nominal coverage rates. When using prediction convex hulls, one expects to get the smallest convex region with the highest probability guarantee. The untrimmed temporal convex hulls return 90.4%, 56%, 10% and almost 0% empirical coverage rates in dimensions 2, 6, 12 and 24 for wind power, respectively. The spatial prediction convex hulls contain 90% of the PV power measurements. Our empirical results suggest that the prediction convex hulls perform poorly in higher dimensions. We produced temporal prediction con-

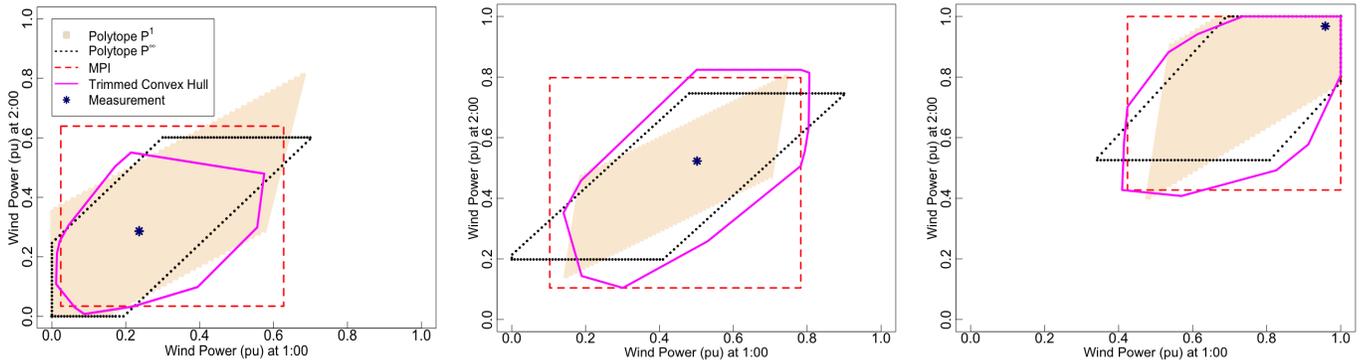


Fig. 4: Visual comparison of temporal prediction polyhedra of dimension 2 for wind power data.

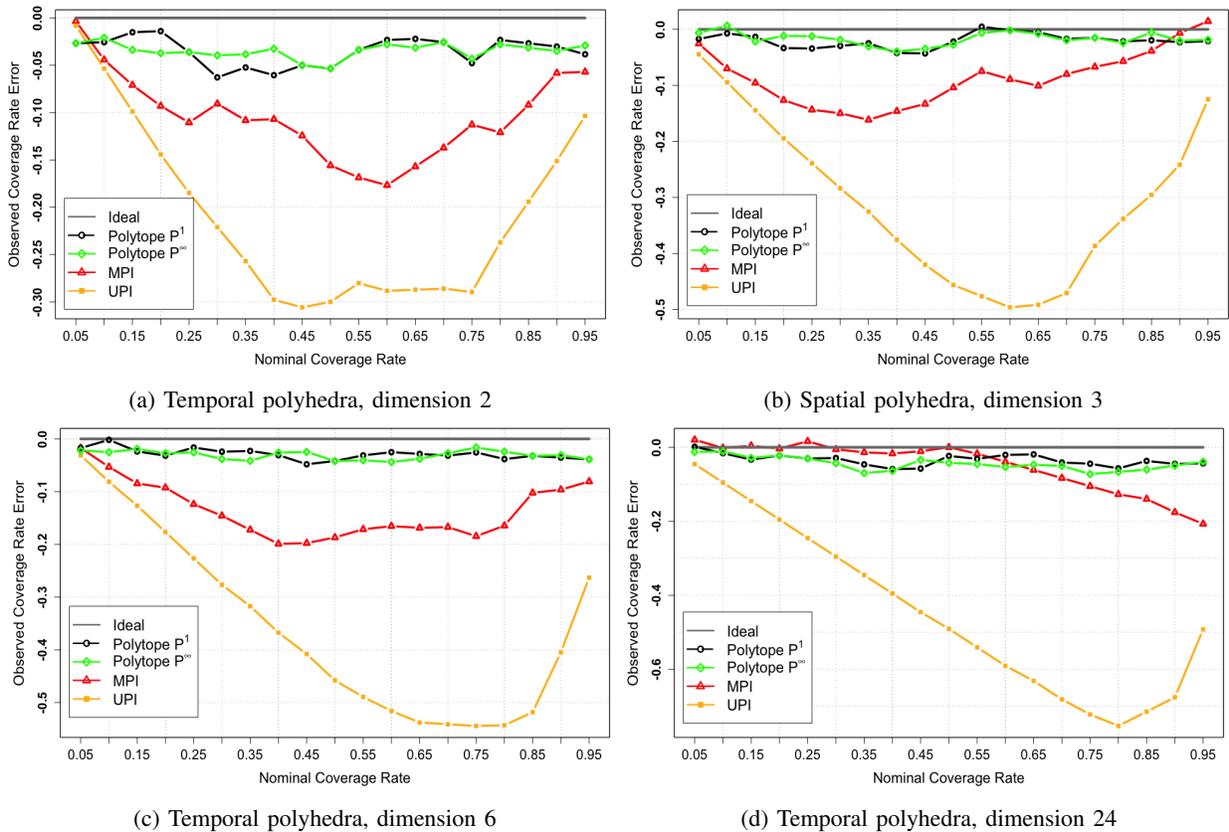


Fig. 5: The difference between empirical coverage and nominal coverage rates (Empirical-Nominal).

vex hulls of dimension 4 for wind power data (1:00 am to 4:00 am) and obtained 72% empirical coverage rate. Thus, we do not recommend prediction convex hulls in dimensions higher than 4. Additionally, although it is straightforward and computationally efficient to compute calibration of convex hulls following (6) and (7), the algorithm to find convex hulls themselves becomes very slow for dimensions higher than 8 and it does not converge in dimensions more than 9.

For the bounded random variables, the size of the prediction polyhedra is determined by the intersection of two polyhedra. The first one is the prediction polyhedron itself and the second one is formed by the feasible range of the random variable. For the case of wind/PV power, the second polyhedron is a hypercube with edges of length equal to the maximum capacity of generation unit. Because there is no simple formulation to calculate the intersection analytically, we use a Monte Carlo-based method for estimation of the volume of prediction

regions [13]. The idea is to generate N' random samples in the feasible range and then calculate the proportion of those points which lie in the prediction polyhedron. The volume of prediction polyhedra V^P is calculated as

$$V^P = N''V^c/N' \quad (12)$$

with N'' as the number of D -dimensional points enveloped by the prediction polyhedron and V^c is the volume of the bounded hyper-cube. Fig. 6 illustrates the size of the proposed prediction polyhedra in comparison with MPI for ten randomly selected days from the evaluation data. The prediction polyhedra of sizes 2, 3, 6 and 12 and nominal probabilities 95%, 90%, 85% and 80% are included in the figure. It is to be noted that to study the results for different days, the selected days are not the same for all dimensions shown in the figure. The vertical axes are logarithmic for a clearer illustration. The empirical coverage of original prediction convex hulls is 90% and it reduces to 87% for trimmed convex hulls.

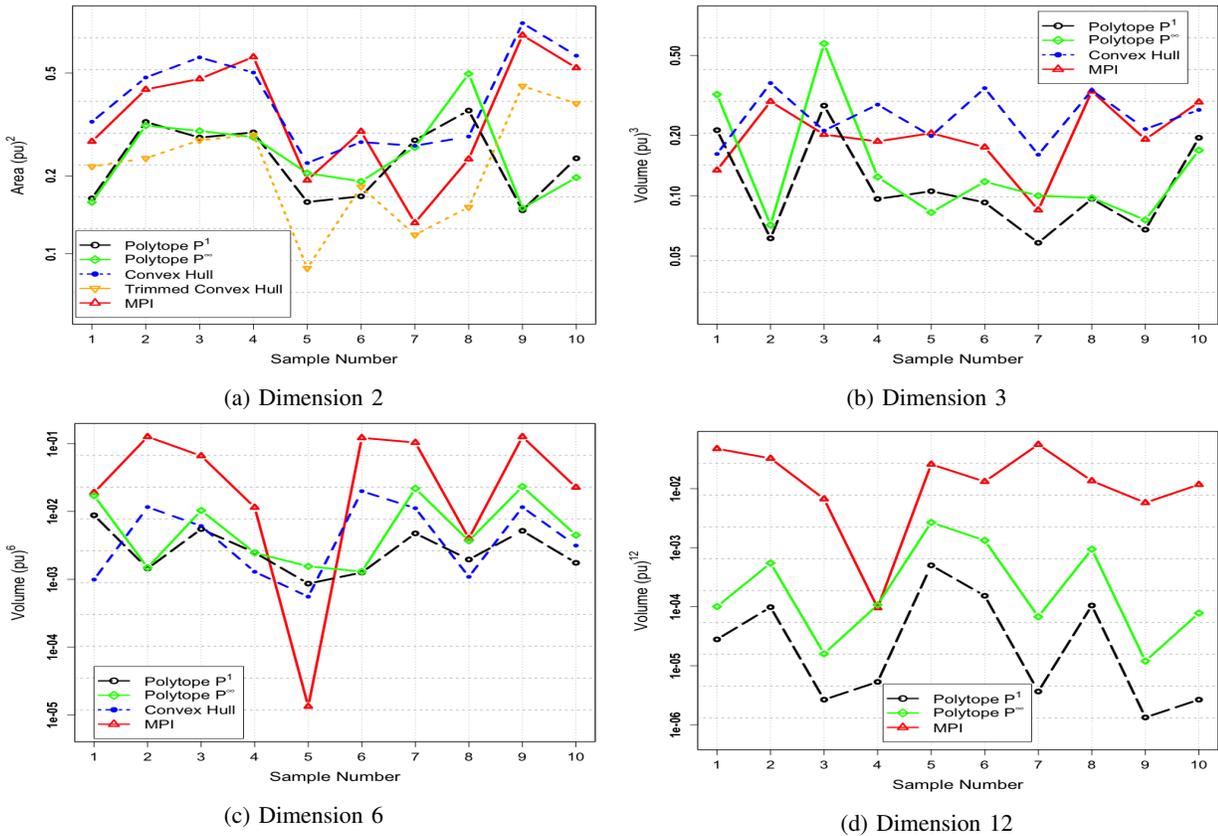


Fig. 6: Estimated volume of prediction polyhedra with nominal coverage rates (a) 95%, (b) 90%, (c) 85% and (d) 80%.

However, as trimming the convex hulls in dimension 2 reduces their sizes significantly, we recommend discarding outliers regardless. Among the four techniques, P^1 shows the overall best performance in terms of conservativeness and sharpness. As the dimension increases, the MPIs become wider and more conservative. For example, in dimension 12 for the first day, 80% MPI is more than 100 times larger than 80% P^1 . In order to better visualize the size of MPIs in higher dimensions, Fig. 7 provides the MPIs for a randomly selected day from the evaluation data along with the marginal prediction intervals and multivariate trajectories used as their inputs. As shown in the figure, MPIs specially those with low coverage rates are wide and low in sharpness.

In almost all real world datasets, there is a possibility of outlier occurrence. Outliers might come from error in measurement, collection or communication of data. Outliers can pose serious problems in statistical analysis and grossly distort and mislead them. The first measure to deal with outliers is to identify them, then either they can be discarded or replaced with more consistent data. Detection of multivariate outliers is discussed in subsection II-C. It is important to assess the robustness of various regression models to outliers. Fig. 8 shows historical wind power measurements available in the training subset recorded at 1:00 am and 2:00 am. Following the approach discussed in subsection II-C, those observations with Mahalanobis distance higher than $1.1\chi_2^2(\alpha = 0.001)$ are detected as outliers. We generate six more synthetic outliers as shown in Fig. 8 and substitute them for 6 randomly selected measurements. Then, we compare the performance of prediction polyhedra with and without the presence of those synthetic outliers. Based on empirical results, outliers on average change the empirical coverage of MPIs, P^1 , P^∞ and convex hulls by 1.5%, 0.7%, 0.9% and 3%, respectively.

The occurrence of outlines also changes the volumes of 90% prediction polyhedra by 0.048, 0.012, 0.020 and 0.044 for MPIs, P^1 , P^∞ and convex hulls, respectively. The results indicate that P^1 provides the highest robustness to outliers followed by P^∞ .

VI. CONCLUSION

In order to facilitate the transition from deterministic forecasts to the point where end-users can confidently harness uncertainty information, it is required to develop frameworks allowing to characterize uncertainty in forms that suit best the needs of various decision-making communities. Due to growing interests in polyhedral uncertainty sets, this work proposed frameworks to generate, calibrate and evaluate uncertainty information in the form of multivariate polyhedra for PV and wind power and within various temporal and spatial scales. Two of the proposed techniques use point and correlation matrix forecasts as inputs and predict the uncertainty budget such that prediction polyhedra provide the desired probability levels and conservativeness. Two other techniques work based on finding convex hulls of spatial/temporal scenarios. The proposed approaches together with multivariate prediction intervals as a benchmark are compared based on their calibration and conservativeness. The empirical results suggest that prediction convex hulls are not recommended for wind/PV predictions in dimensions higher than four. P^1 shows overall the best performance followed by P^∞ . Both P^1 and P^∞ are promising formulations for skilled uncertainty characterization in convex forms and their performance does not degrade as the dimension increases, provided that their parameters are predicted appropriately.

REFERENCES

- [1] D. Bertsimas, E. Litvinov, X. A. Sun, J. Zhao, and T. Zheng, "Adaptive robust optimization for the security constrained unit commitment prob-

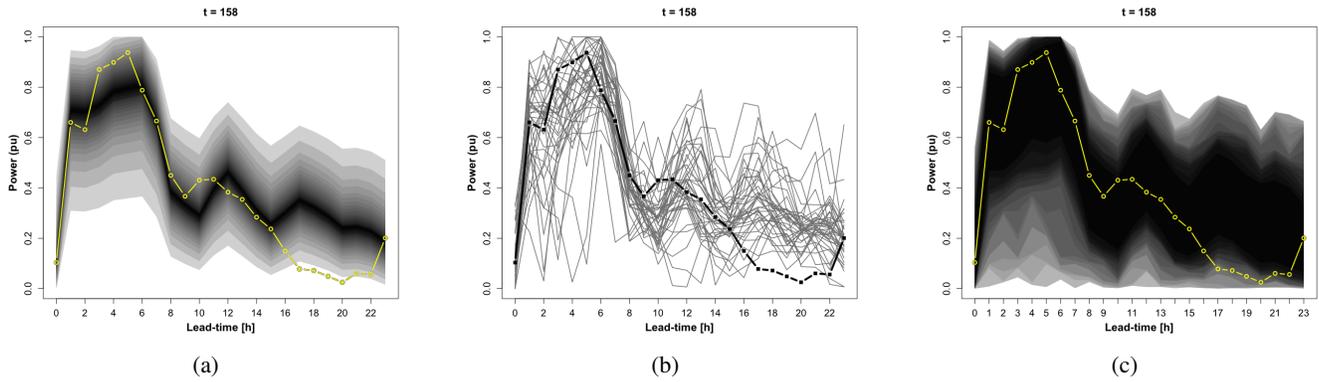


Fig. 7: (a) PV observations (yellow color curves) along with 19 univariate prediction intervals with coverage ranging from 0.05 to 0.95 in 0.05 increments (from the darkest to the lightest), (b) PV observations (dark black color curves) along with 40 generated space-time trajectories (gray color curves), (c) PV power observations (yellow color curves) along with 19 MPIs with nominal coverage ranging from 0.05 to 0.95 in 0.05 increments (from the darkest to the lightest)

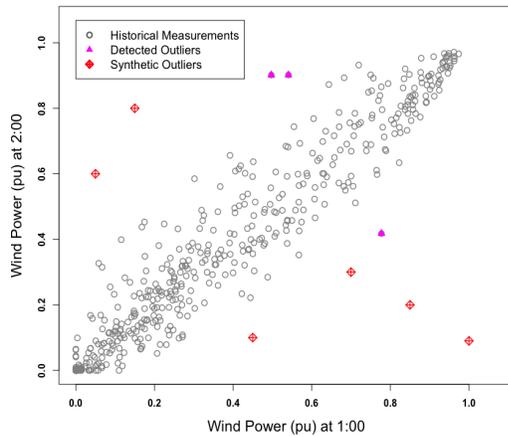


Fig. 8: Temporal scenarios predicted for a randomly selected day from the wind power data. Few scenarios are detected as outliers. 6 Synthetic outliers are generated for analysis.

- lem,” *IEEE Transactions on Power Systems*, vol. 28, no. 1, pp. 52–63, 2013.
- [2] R. J. Bessa, C. Möhrlen, V. Fundel, M. Siefert, J. Browell, S. Haglund El Gaidi, B.-M. Hodge, U. Cali, and G. Kariniotakis, “Towards improved understanding of the applicability of uncertainty forecasts in the electric power industry,” *Energies*, vol. 10, no. 9, p. 1402, 2017.
 - [3] A. Ben-Tal and A. Nemirovski, “Robust solutions of linear programming problems contaminated with uncertain data,” *Mathematical programming*, vol. 88, no. 3, pp. 411–424, 2000.
 - [4] 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.
 - [5] P. Haessig, B. Multon, H. B. Ahmed, S. Lascaud, and P. Bondon, “Energy storage sizing for wind power: impact of the autocorrelation of day-ahead forecast errors,” *Wind Energy*, vol. 18, no. 1, pp. 43–57, 2015.
 - [6] C. Lowery and M. O’Malley, “Reserves in stochastic unit commitment: An Irish system case study,” *IEEE Transactions on Sustainable Energy*, vol. 6, no. 3, pp. 1029–1038, 2015.
 - [7] Q. P. Zheng, J. Wang, and A. L. Liu, “Stochastic optimization for unit commitment? A review,” *IEEE Transactions on Power Systems*, vol. 30, no. 4, pp. 1913–1924, 2015.
 - [8] W. Wiesemann, D. Kuhn, and M. Sim, “Distributionally robust convex optimization,” *Operations Research*, vol. 62, no. 6, pp. 1358–1376, 2014.
 - [9] M. Doostizadeh, F. Aminifar, H. Ghasemi, and H. Lesani, “Energy and reserve scheduling under wind power uncertainty: An adjustable interval approach,” *IEEE Transactions on Smart Grid*, vol. 7, no. 6, pp. 2943–2952, 2016.
 - [10] W. Wei, F. Liu, and S. Mei, “Distributionally robust co-optimization of energy and reserve dispatch,” *IEEE Transactions on Sustainable Energy*, vol. 7, no. 1, pp. 289–300, 2016.
 - [11] B. Hu and L. Wu, “Robust SCUC with multi-band nodal load uncertainty set,” *IEEE Transactions on Power Systems*, vol. 31, no. 3, pp. 2491–2492, 2016.
 - [12] A. Venzke, L. Halilbasic, U. Markovic, G. Hug, and S. Chatzivasileiadis, “Convex relaxations of chance constrained ac optimal power flow,” *IEEE Transactions on Power Systems*, 2017.
 - [13] F. Golestaneh, P. Pinson, R. Azizipanah-Abarghooee, and H. B. Gooi, “Ellipsoidal prediction regions for multivariate uncertainty characterization,” *IEEE Transactions on Power Systems*, 2018.
 - [14] 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.
 - [15] C. Shao, X. Wang, M. Shahidehpour, X. Wang, and B. Wang, “Security-constrained unit commitment with flexible uncertainty set for variable wind power,” *IEEE Transactions on Sustainable Energy*, vol. 8, no. 3, pp. 1237–1246, 2017.
 - [16] T. Soares, R. J. Bessa, P. Pinson, and H. Morais, “Active distribution grid management based on robust AC optimal power flow,” *IEEE Transactions on Smart Grid*, 2017.
 - [17] 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.
 - [18] D. A. Pachamanova, “A robust optimization approach to finance,” Ph.D. dissertation, Massachusetts Institute of Technology, 2002.
 - [19] 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.
 - [20] C. B. Barber, D. P. Dobkin, and H. Huhdanpaa, “The quickhull algorithm for convex hulls,” *ACM Transactions on Mathematical Software (TOMS)*, vol. 22, no. 4, pp. 469–483, 1996.
 - [21] E. Mücke, “Quickhull: Computing convex hulls quickly,” *Computing in Science & Engineering*, vol. 11, no. 5, pp. 54–57, 2009.
 - [22] M. Werner, “Identification of multivariate outliers in large data sets,” Ph.D. dissertation, Citeseer, 2003.
 - [23] D. Kolsrud, “Time-simultaneous prediction band for a time series,” *Journal of Forecasting*, vol. 26, no. 3, pp. 171–188, 2007.
 - [24] J. S.-H. Li and W.-S. Chan, “Simultaneous prediction intervals: An application to forecasting US and Canadian mortality,” 2011.
 - [25] T. Gneiting, F. Balabdaoui, and A. E. Raftery, “Probabilistic forecasts, calibration and sharpness,” *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 69, no. 2, pp. 243–268, 2007.
 - [26] B. Büeler, A. Enge, and K. Fukuda, “Exact volume computation for polytopes: a practical study,” in *Polytopes: combinatorics and computation*. Springer, 2000, pp. 131–154.
 - [27] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.
 - [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] R. F. Engle and K. Sheppard, “Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH,” National Bureau of Economic Research, Tech. Rep., 2001.
 - [30] V. Zakamulin, “A test of covariance-matrix forecasting methods,” *The Journal of Portfolio Management*, vol. 41, no. 3, pp. 97–108, 2015.
 - [31] Global energy forecasting competition 2014 probabilistic solar power forecasting. [Online]. Available: <https://crowdanalytix.com>