Ensembling Wind Forecasting Models to Construct Data-Driven Uncertainty Sets in Robust Optimization

Panagiotis Andrianesis
Division of Systems Engineering
Boston University
Boston, MA, United States
panosa@bu.edu

Dimitris Bertsimas
Operations Research Center
Massachusetts Institute of Technology
Cambridge, MA, United States
dbertsim@mit.edu

Thodoris Koukouvinos
Operations Research Center
Massachusetts Institute of Technology
Cambridge, MA, United States
tkoukouv@mit.edu

Angelos Georgios Koulouras
Operations Research Center
Massachusetts Institute of Technology
Cambridge, MA, United States
angkoul@mit.edu

***Abstract*—Robust optimization (RO) has shown promising results in dealing with the uncertainty that renewable energy sources introduce in power systems. RO is based on uncertainty sets that contain all uncertain scenarios against which we want to be robust. Despite their importance in controlling the cost and conservativeness of the RO formulations, there has been little work in determining their size. In this paper, we propose a data-driven framework for constructing uncertainty sets using mixed integer optimization and machine learning. Specifically, we ensemble wind forecasting models to predict the available capacity of wind resources and we learn the uncertainty set size based on the confidence in our predictions. In addition, we introduce optional amendments to our framework that capture the spatial and temporal correlations in the data. We apply our framework on real-world forecasting models and wind generation data of a large US Independent System Operator and show empirically that ensembling, apart from improving the individual forecasting model predictions, yields uncertainty sets with small sizes and good empirical probabilities of constraint violation.**

***Index Terms*—Robust optimization, uncertainty sets, machine learning, ensemble models, wind generation uncertainty.**

I. Introduction

DESPITE the increasing presence of volatile and uncertain renewable generation in power grids and markets, deterministic approaches in the unit commitment problem remain prevalent in practice. Notably, recent advances in robust optimization (RO) [1] offer promising results in protecting against different types of uncertainty, e.g., in the load [2], [3], wind generation [4], [5] or both [6], [7], and have shown lower average dispatch and total costs, as well as lower volatility of the total costs [1]. RO is based on uncertainty sets containing all uncertain scenarios against which we want to be robust. So, the uncertainty sets are at the heart of RO, as their size and shape heavily affects the RO problem solution [8]. For example, if the uncertainty set is large, we protect against a large range of realizations of the uncertain parameter, so we are more conservative and the cost increases. If the uncertainty set is too small, we may exclude scenarios that are of interest and may have to make up for them with costly subsequent actions, e.g., with additional commitments and re-dispatching.

Most of the existing works feature widely used box and budget uncertainty sets [9] but they usually do not provide methods for selecting their size. The bounds of the uncertainty sets are usually selected either somewhat arbitrarily or with strong probabilistic assumptions. For example, [1], [3], which primarily introduce robust unit commitment formulations, only report results for different values of the hyperparameter controlling the size of the budget uncertainty set. While they offer useful insights into the trade-off between robustness and conservativeness, a principled framework for selecting the budget size is missing, indicating a gap in standardized methods for constructing uncertainty sets. Recent works have proposed data-driven approaches for constructing uncertainty sets that capture the intricate dependencies between the uncertain parameters, e.g., including constraints that capture the spatiotemporal correlations [2], [4], [6], or estimating the radius of ellipsoidal uncertainty sets using the inverse covariance matrix and excluding outliers [10]. Another line of works relies on historical data to construct the uncertainty set, such as previous realizations of the load or available wind capacity [5], [6], without considering related covarariates.

In this paper, we propose a data-driven framework for constructing uncertainty sets, using mixed integer optimization and machine learning (ML). We build a linear regression model for predicting the nominal values of the uncertain parameters, leveraging historical realizations along with covariates, while also learning the size of the uncertainty set. Specifically, we *ensemble* wind forecasting models and con-

Research partially supported by the Advanced Research Projects Agency with award number DE-AR0001282.

struct an uncertainty set around the predicted values based on our confidence in these predictions. Notably, although ensemble models are common in ML and have been used to predict the nominal values of the uncertain parameters, they have not been used for constructing uncertainty sets [11].

Our main contribution is three-fold. First, we provide a mixed integer optimization formulation for automating the construction of uncertainty sets by ensembling wind power forecasting models, while simultaneously learning their weight and the size of the uncertainty set. We further show that our data-driven uncertainty sets imply probabilistic guarantees bounded by the training error of the underlying ensemble model. Second, we augment the proposed formulation by introducing slowly-varying ensemble weights that capture the spatial and temporal correlations of the wind forecasts. Third, we demonstrate our method on real-world forecasting models and wind generation data of a large US Independent System Operator (ISO). We show that our approach constructs uncertainty sets with tight bounds and good empirical probabilities of constraint violation, whereas our ensemble models may improve the individual forecast predictions in terms of out of sample mean absolute error (MAE).

The remainder of the paper is organized as follows. In Section II, we introduce the basic RO concepts. In Section III, we construct uncertainty sets using regression models that ensemble different forecasting models, and in Section IV, we embed spatial and temporal correlations using slowly-varying ensemble weights. In Section V, we present numerical experiments, and in Section VI, we summarize our conclusions.

II. RO and Uncertainty Set Preliminaries

In this section, we briefly introduce the concepts and preliminaries on robust constraints, probabilistic guarantees, and uncertainty sets, which are useful for the rest of the paper.

Let $\mathcal{U}$ denote an uncertainty set. We consider the following robust constraint:

$$f(\boldsymbol{s}, \boldsymbol{y}) \leq 0, \quad \forall \boldsymbol{y} \in \mathcal{U},$$

where vectors $\boldsymbol{s}$ and $\boldsymbol{y}$ represent the optimization variables and uncertain parameters, respectively. The function f is linear or concave in $\boldsymbol{y}$, and $\mathcal{U}$ is convex, in which case the robust counterpart can be easily derived — see [8]. In our setup, $\boldsymbol{s}$ refers to the dispatch of wind resources, and $\boldsymbol{y}$ to their available capacity. We consider the robust constraint:

$$\boldsymbol{s} \leq \boldsymbol{y}, \quad \forall \boldsymbol{y} \in \mathcal{U},$$

ensuring that the wind dispatch does not exceed the available capacity for all possible realizations of $\boldsymbol{y}$ in $\mathcal{U}$.

In general, we assume that the uncertain parameters are distributed according to an *unknown* probability distribution $\mathbb{P}^*$. Hence, it is important to construct an uncertainty set that is both computationally tractable and implies a probabilistic guarantee for $\mathbb{P}^*$ [12]. For a given $\epsilon > 0$, any $\boldsymbol{s}$ and any function f concave in $\boldsymbol{y}$, the uncertainty set $\mathcal{U}$ implies a probabilistic guarantee at level ϵ, if the following holds:

$$f(\boldsymbol{s}, \boldsymbol{y}) \leq 0, \, \forall \boldsymbol{y} \in \mathcal{U} \implies \mathbb{P}^*(f(\boldsymbol{s}, \boldsymbol{y}) \leq 0) \geq 1 - \epsilon.$$

We consider norm-bounded uncertainty sets, where the deviation of the uncertain parameters $\boldsymbol{y}$ from their nominal values $\hat{\boldsymbol{y}}$ is at most δ, i.e.,

$$\mathcal{U}(\hat{\boldsymbol{y}}, \delta) := \{\boldsymbol{y} \in \mathbb{R}^n : \|\boldsymbol{y} - \hat{\boldsymbol{y}}\|_\ell \leq \delta\}, \quad \ell \in \{1, \infty\}.$$

Note that $\ell = 1$ yields the widely used budget uncertainty set, whereas $\ell = \infty$ yields the box uncertainty set.

III. Constructing Uncertainty Sets from Regression Models

In this section, we develop a standardized framework for constructing uncertainty sets from data.

In our setup, we consider n wind resources, with available capacity y_i, for $i = 1, \ldots, n$. The available data include K historical forecasts, deriving from m wind forecasting models, represented by the m-sized vector $\boldsymbol{x}_{i,k}$, for wind resource i and historical forecast $k = 1, \ldots, K$, along with the realized available capacity, $y_{i,k}$. The goal is to leverage the power of the separate forecasting models and ensemble them to make a single prediction $\hat{y}_i$ for the available capacity of wind resource i. To that end, we utilize a linear regression formulation where the coefficients of the model correspond to the weights given to each forecasting model in the ensemble. Our objective is to learn the regression coefficients, while also finding the tightest possible norm-bounded uncertainty set such that $p \times 100\%$ of our predicted values for each wind resource fall in that set.

Consider a box uncertainty set centered around the prediction for the available capacity, given by:

$$\mathcal{U} = \{\boldsymbol{y} \in \boldsymbol{R}^n : \, |y_i - \hat{y}_i| \leq \delta_i, \; \forall i \in [n]\},$$

whose size is parameterized by $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_n)$. The prediction is obtained by applying the m-sized vector $\boldsymbol{\beta}_i$ (regression coefficients) to ensemble the forecasting model output $\boldsymbol{x}_{i,k}$. We define the binary variable $z_{i,k}$ such that $z_{i,k} = 1$ if the prediction of the ensemble for wind resource i, historical forecast k, $\hat{y}_{i,k}$, is within δ_i from the realized available capacity, $y_{i,k}$, and $z_{i,k} = 0$ otherwise. We then enforce the empirical probabilistic guarantee requirement that at least p of the residuals are below δ_i, i.e., $\sum_{k=1}^{K} z_{i,k} \geq p \; n$.

The proposed mixed integer optimization formulation that simultaneously learns the regression coefficients, $\boldsymbol{\beta}$, and the uncertainty set size, $\boldsymbol{\delta}$, is as follows.

$$\min_{\boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{z}} \; \frac{1}{n} \sum_{i=1}^{n} \sum_{k=1}^{K} |y_{i,k} - \boldsymbol{\beta}_i^T \boldsymbol{x}_{i,k}| + \sum_{i=1}^{n} \delta_i, \tag{1a}$$

$$\text{s.t. } |y_{i,k} - \boldsymbol{\beta}_i^T \boldsymbol{x}_{i,k}| \leq \delta_i + M(1 - z_{i,k}), \quad \forall i \in [n], k \in [K], \tag{1b}$$

$$\sum_{k=1}^{K} z_{i,k} \geq p \; n, \quad \forall i \in [n], \tag{1c}$$

$$z_{i,k} \in \{0, 1\}, \quad \forall i \in [n], k \in [K], \tag{1d}$$

$$\boldsymbol{\beta}_i \geq 0, \quad \forall i \in [n], \tag{1e}$$

where M is a big number, and (1e) ensures non-negative weights. Note that Problem (1) can be solved for each resource i separately and in parallel, speeding up computations.

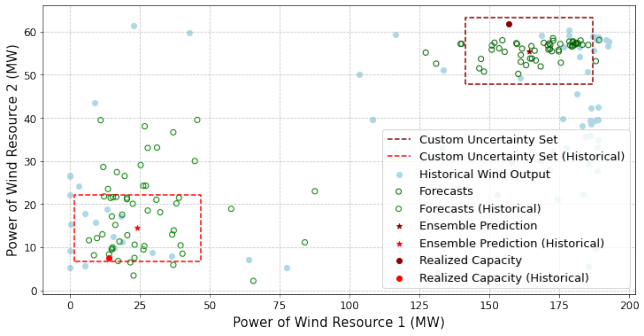

Fig. 1. Box uncertainty sets for two wind resources.

We illustrate our approach in Fig. 1, for two wind resources. We show historical realizations ($y_{i,k}$) along with historical forecasts (training set) and the historical ensemble prediction for one instance (bottom left), and the current forecasts and ensemble prediction (top right). Around the ensemble prediction, the box uncertainty sets (observe the same size of δ_i in both boxes), include a certain amount of forecasts in the training set (see indicatively the bottom left box), which on aggregate satisfy a certain probability p. Notably, our box uncertainty set includes the realization (even though it is of smaller size compared to a box that would include all forecasts).

Next, we show that a box uncertainty set constructed via Problem (1) satisfies a certain probabilistic guarantee. We define $\mathcal{U}_i = \{y_i \in \mathbb{R} : |y_i - \hat{y}_i| \leq \delta_i, \}$ the individual uncertainty set of wind resource i. We denote the joint distribution of $\boldsymbol{x}, y$ with $\mathbb{P}^*$ and the loss of the regression model with $L_{\mathcal{W}}(\boldsymbol{\beta})$ over a distribution $\mathcal{W}$ of the data.

Theorem 1. *Suppose that $\mathcal{D}$ is a distribution over $X \times Y$ such that with probability 1 we have that $\|\boldsymbol{x}\|_2 \leq R$ and $y \leq Q$. Let S be the training set of $\boldsymbol{x}_{i,k}$ for $i = 1, \ldots, n$ and $k = 1, \ldots, K$. Suppose $\boldsymbol{\beta}_i^*$, which is the optimal solution to Problem (1), is $\|\boldsymbol{\beta}_i^*\|_2 \leq B$ for all i. Then, for any $\epsilon \in (0, 1)$, with probability of at least $1 - \epsilon$ over the choice of an i.i.d. sample size of $K \to \infty$ and for all wind resources i*

$$\mathbb{P}^*(s_i \leq y_i) \geq p. \tag{2}$$

Proof. The MAE $\phi(a, y) = |a - y|$ is a ρ-Lipschitz continuous loss function and $\max_{a \in [-BR, BR]} \phi(a, y) \leq c$, for $c = BR + Q$ since y is bounded by Q. Then, using [13], for any $\epsilon \in (0, 1)$, with probability of at least $1 - \epsilon$,

$$L_{\mathcal{D}}(\boldsymbol{\beta}_i^*) \leq L_S(\boldsymbol{\beta}_i^*) + \frac{2\rho BR}{\sqrt{K}} + c\sqrt{\frac{2\log(2\ \epsilon)}{K}}.$$

So, for each wind resource i, with probability of at least $1 - \epsilon$,

$$\begin{aligned}
\mathbb{P}^*(s_i \leq y_i) &\geq \mathbb{P}^*(y_i \in \mathcal{U}_i) = \mathbb{P}^*(L_{\mathcal{D}}(\boldsymbol{\beta}_i^*) \leq \delta_i) \\
&\geq \mathbb{P}^*\left(L_S(\boldsymbol{\beta}_i^*) + \frac{2\rho BR}{\sqrt{K}} + c\sqrt{\frac{2\log(2\epsilon)}{K}} \leq \delta_i\right) \\
&\geq \mathbb{P}^*\left(L_S(\boldsymbol{\beta}_i^*) \leq \delta_i - \frac{2\rho BR}{\sqrt{K}} - c\sqrt{\frac{2\log(2\epsilon)}{K}}\right).
\end{aligned}$$

Also, $\mathbb{P}^*(s_i \leq y_i) \geq \mathbb{P}^*(y_i \in \mathcal{U}_i) \geq \mathbb{P}^*(L_S(\boldsymbol{\beta}_i^*) \leq \delta) = p$, for $K \to \infty$, which concludes the proof. □

Note that Theorem 1 also provides a guarantee of p_s for a finite number of training points K, if we let p_s be the ratio of the residuals in the training set less than or equal to $\delta_i - \frac{2\rho BR}{\sqrt{K}} - c\sqrt{\frac{2\log(2\ \epsilon)}{K}}$. Also, the assumption of bounded y is not too strong, because y correspond to the available capacities of wind resources, which are bounded by construction. This bound is based on the Radamacher complexity of linear models and is relatively loose but the previous probabilistic guarantee could be improved using tighter generalization error bounds.

IV. Slowly-Varying Ensemble Weights

In this section, we introduce the uncertainty set construction formulations that depend on time (in Subsection IV-A), which we further amend to capture spatial and temporal correlations (in Subsection IV-B).

A. Time-Dependent Uncertainty Set Construction

Motivated by the fact that the forecasts are more accurate on the nearest time periods and become less accurate after a large number of periods, it makes sense to adjust the size of the uncertainty set accordingly. We introduce the time parameter $t = 1, \ldots, T$, where T is the time horizon of the forecasts. Without loss of generality, t represents an hourly time period, in a daily horizon $T = 24$. Instead of assuming a constant size δ_i for all time periods, we allow it to adjust over time as $\delta_{i,t}$. In this case, Problem (1) is formulated as follows.

$$\min_{\boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{z}} \frac{1}{n} \sum_{i=1}^{n} \sum_{k=1}^{K} \sum_{t=1}^{T} |y_{i,k,t} - \boldsymbol{\beta}_i^T \boldsymbol{x}_{i,k,t}| + \sum_{i=1}^{n} \sum_{t=1}^{T} \delta_{i,t}, \tag{3a}$$

$$\text{s.t. } |y_{i,k,t} - \boldsymbol{\beta}_i^T \boldsymbol{x}_{i,k,t}| \leq \delta_{i,t} + M(1 - z_{i,k,t}), \quad \forall i \in [n], k \in [K], t \in [T], \tag{3b}$$

$$\sum_{k=1}^{K} z_{i,k,t} \geq p\ n, \qquad \forall i \in [n], \tag{3c}$$

$$z_{i,k,t} \in \{0, 1\}, \qquad \forall i \in [n], k \in [K], t \in [T], \tag{3d}$$

$$\boldsymbol{\beta}_i \geq 0, \qquad \forall i \in [n]. \tag{3e}$$

Note that we now have data-points for each time period and the bounds $\delta_{i,t}$ vary over time.

B. Spatial and Temporal Correlations

It is often the case that wind resources that are in nearby locations have similar forecasts (spatial correlations), whereas the generation of a resource at a specific time period is related to the generation at previous time periods (temporal correlations). To capture both types of correlations, we can embed in Problem (3) slowly-varying ensemble weights.

To account for spatial correlations, we require that the weights of the ensemble models vary slowly among nearby wind resources, i.e.,

$$\|\boldsymbol{\beta}_i - \boldsymbol{\beta}_j\| \leq b_{i,j}, \quad \forall i \in [n], j \in J(i), \tag{4}$$

where $J(i)$ is the set of wind resources that are close to wind resource i, and $b_{i,j}$ are parameters whose values can be determined from the correlations of different forecasting models in the training set for a pair (i, j) of wind resources.

To account for temporal correlations, we require that for each wind resource, the uncertainty set size varies slowly from time period t to $t+1$, i.e.,

$$|\delta_{i,t+1} - \delta_{i,t}| \leq d_{i,t}\delta_{i,t}, \quad \forall i \in [n], t \in [T], \tag{5}$$

where $d_{i,t}$ are parameters that can be determined as the relative difference of the standard deviation across different forecasting models in the training set between time periods t and $t+1$.

V. Numerical Experiments

In this section, we apply our method to forecasting models from a vendor that caters to a large US ISO. We used 24-hour forecasts for the following day, for a selection of $n = 18$ wind resources, and $m = 5$ forecasting models, along with realizations of the available capacity. For illustration purposes, we first constructed the uncertainty set solving Problem (3), using $p = 0.8$, and enforcing a constant size for all hours, i.e., $\delta_{i,t} = \delta_i$, $\forall t$. The training set included the three days prior to the day of interest, which was the testing set. The relatively short time horizon of the training set allows to determine the weights based on days that are close to the test day, however, the training set could well contain any type/number of historical forecasts that are selected to be representative, e.g., of the same weather type to the testing set.

Although the scope of this work is not forecasting in itself, we observe that the predictions of our ensemble model are on par with the best-performing model that may vary from hour to hour. Specifically, the MAE of the best performing model in three test days, were 11.2, 20.3, and 15.4, respectively, followed by the 11.8, 22.3, and 16.0, respectively, of the second best performing model, whereas our ensemble prediction MAE was 11.1, 21.6, and 12.4, respectively. Hence, our prediction, around which we center our uncertainty set, is close to the nominal value.

We further proceed with the illustration of the uncertainty set, and our attention shifts to the aggregate wind generation, whose uncertainty determines the amount of reserve requirements. In Fig. 2, we plot the sum of the predictions of the ensemble model for all wind resources at each hour of the first test day, the realizations, and the bounds of the uncertainty set for the hourly aggregate wind generation, i.e., $\delta^{\text{agg}} = \sum_{i=1}^{n} \delta_i = 237.67$. We observe that for all hours the realizations are within the lower and upper bounds obtained by our approach.

In Figure 3 (left), we illustrate the values of the uncertainty set size parameters, derived from (i) the sum of the individual uncertainty sets, and (ii) the aggregate prediction, i.e., solving Problem (3) for the aggregate wind generation, versus the probability p. Not surprisingly, the aggregate prediction achieves higher probabilities with lower uncertainty set sizes. For example, the probability $p = 0.8$ used earlier (see Fig. 2) for the individual uncertainty sets that yielded $\delta^{\text{agg}} = 237.67$,

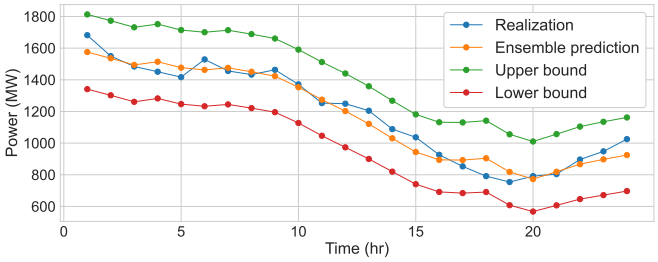

Fig. 2. Prediction and realization for the aggregate wind generation, and uncertainty set lower/upper bounds (δ^{agg}) using the ensemble model.

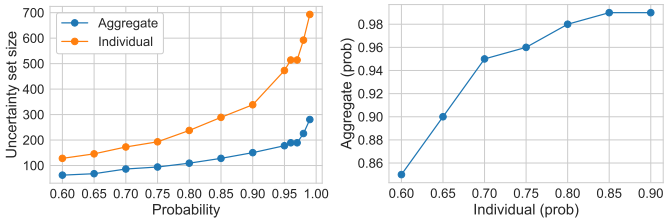

Fig. 3. Probabilities versus size (δ^{agg}) for the sum of individual and the aggregate uncertainty sets (left). "Iso-δ" curve (right).

TABLE I
Out of Sample Probabilities

Probability p		0.5	0.6	0.7	0.8	0.9
Individual	**Both bounds**	0.27	0.39	0.5	0.66	0.75
	Lower bound	0.61	0.66	0.72	0.8	0.88
Aggregate	**Both bounds**	0.79	0.96	1.00	1.00	1.00
	Lower bound	0.83	0.96	1.00	1.00	1.00

would only require less than half the size (109.18) for the aggregate uncertainty set. Notably, a size of 237.67 would correspond to a probability of 0.98 for the aggregate wind generation. In Fig. 3 (right), we illustrate an "iso-δ" curve, which relates the probabilities of individual uncertainty sets to the probabilities of the aggregate uncertainty set for the same size. For example, we observe that the probabilities of 0.65, 0.75, and 0.85 for the individual uncertainty sets, would translate to probabilities of 0.9, 0.96, and 0.99, respectively, for the aggregate wind generation.

In Table I, we report the probability of the available capacity belonging to the uncertainty set for different values of the parameter p. For both the individual uncertainty sets and the aggregate, we present the probability for the entire uncertainty set, i.e., the available capacity being within both lower and upper bounds (both bounds), as well as only the lower bound (to emphasize the wind falling worst case). We observe that the out of sample probability may be smaller for the individual uncertainty sets, when considering both bounds, whereas it is close if not better when considering the lower bound. For the aggregate uncertainty sets, the out of sample probability is significantly higher, and reaches 1 for $p \geq 0.7$.

We further experimented with constraints (4) to account for spatial correlations. For illustration purposes, we identified as spatially correlated wind resources those with pairwise

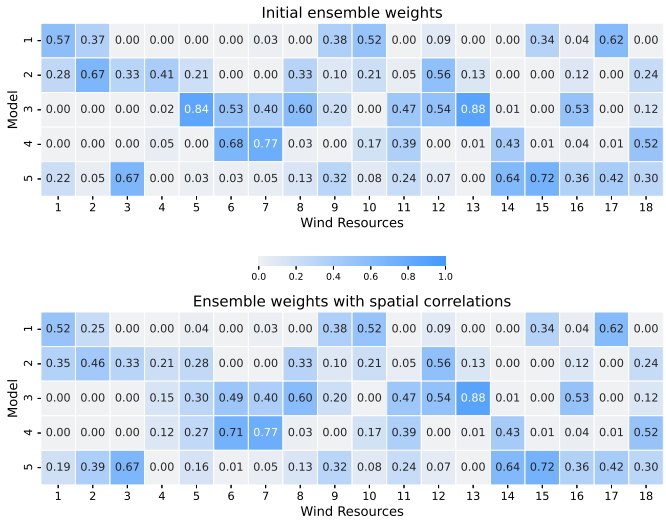

Fig. 4. Ensemble model weights: initial (top), and after including the slowly-varying constraints (4) for spatially correlated resources (bottom).

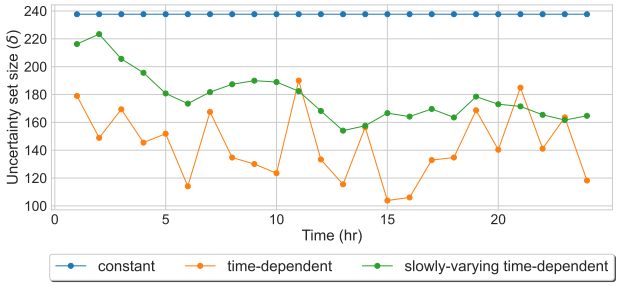

Fig. 5. Aggregate uncertainty set sizes (constant, time-dependent, and slowly-varying time-dependent).

distance, $\text{Dist}(i,j)$, smaller than 12 km, and we set the slowly-varying coefficient to $b_{i,j} = \frac{1}{m}\left(\frac{\text{Dist}(i,j)}{12} + \frac{1}{2}\right)$. In Fig. 4, we show a heatmap of the initial ensemble weights $\boldsymbol{\beta}$ for each forecasting model and wind resource (top), and the new weights (bottom) after including constraints (4). A first remark is the sparsity in the weights in both plots. A second remark refers to the correlated wind resource pairs $(1,2)$, $(2,3)$, $(4,5)$, and $(5,6)$ — see the first 6 columns — whose weights (bottom) are closer compared to the initial ones (top). The MAE is reduced by about 5% at the correlated resources, whereas the uncertainty set size increases by about 5%. We expect the size to increase, because we now have more constraints, so the in sample MAE is larger. However, the large differences in the weights at nearby locations might imply overfitting, which is mitigated by imposing slowly varying weights on highly correlated resources.

Lastly, we accounted for the time-dependent sizes, $\delta_{i,t}$, as well as the temporal correlations, experimenting with constraints (5). For each wind resource i and time period t, we computed the standard deviation in the training set $\sigma_{i,t} = std(\boldsymbol{x}_{i,t})$ and took $d_{i,t} = \frac{|\sigma_{i,t+1} - \sigma_{i,t}|}{\sigma_{i,t}}$. As a result, if the forecasts are "close" at time period t and they are more spread out at $t+1$, σ_t is small and σ_{t+1} is large, so we let the size increase at a larger rate. If the standard deviation of the forecasts is similar at both time periods, then we want the sizes to be similar. Values for $d_{i,t}$ higher than 0.2 were clipped. In Fig. 5, we compare the constant, time-dependent and slowly-varying time-dependent uncertainty set sizes. Evidently, the constant size is too conservative, The time-dependent sizes fluctuate significantly, whereas the slowly-varying time-dependent sizes — obtained by including constraints (5) — are smoother and tighter than the constant sizes. Finally, we note that we did not observe any significant changes in the out of sample MAE, when adding the slowly-varying constraints. We expect that the implementation of such constraints will provide the same probabilistic guarantees with less conservative uncertainty sets.

VI. Conclusion

In summary, we proposed a standardized framework for constructing uncertainty sets for wind power generation by ensembling forecasting models. We applied our method to real-world data from a large US ISO and showed empirically that, using our ensemble model, which is on par with the best-performing individual forecasting models, we can construct individual uncertainty sets with relatively tight bounds and good probabilistic guarantees. Moreover, when we consider the aggregate wind generation, the bounds become tighter with better guarantees.

References

[1] D. Bertsimas, E. Litvinov, X. A. Sun, J. Zhao, and T. Zheng, "Adaptive robust optimization for the security constrained unit commitment problem," *IEEE Trans. Power Syst.*, vol. 28, no. 1, pp. 52–63, 2012.

[2] R. Jiang, M. Zhang, G. Li, and Y. Guan, "Two-stage network constrained robust unit commitment problem," *Eur. J. Oper. Res.*, vol. 234, no. 3, pp. 751–762, 2014.

[3] Y. An and B. Zeng, "Exploring the modeling capacity of two-stage robust optimization: Variants of robust unit commitment model," *IEEE Trans. Power Syst.*, vol. 30, no. 1, pp. 109–122, 2014.

[4] A. Lorca and X. A. Sun, "Multistage robust unit commitment with dynamic uncertainty sets and energy storage," *IEEE Trans. Power Syst.*, vol. 32, no. 3, pp. 1678–1688, 2016.

[5] M. Zhang, J. Fang, X. Ai, B. Zhou, W. Yao, Q. Wu, and J. Wen, "Partition-combine uncertainty set for robust unit commitment," *IEEE Trans. Power Syst.*, vol. 35, no. 4, pp. 3266–3269, 2020.

[6] C. Li, J. Zhao, T. Zheng, and E. Litvinov, "Data-driven uncertainty sets: Robust optimization with temporally and spatially correlated data," in *IEEE Power Energy Soc. Gen. Meeting*, pp. 1–5, 2016.

[7] H. Jin, H. Sun, Q. Guo, and J. Wu, "Robust unit commitment considering reserve from grid-scale energy storage," in *IEEE 8th Int. Power Electron. Motion Control Conf. (IPEMC-ECCE Asia)*, pp. 246–251, 2016.

[8] D. Bertsimas and D. den Hertog, *Robust and adaptive optimization*. Dynamic Ideas, 2022.

[9] Y. Guan and J. Wang, "Uncertainty sets for robust unit commitment," *IEEE Trans. Power Syst.*, vol. 29, no. 3, pp. 1439–1440, 2013.

[10] A. Wasilkoff, P. Andrianesis, and M. Caramanis, "Day-ahead estimation of renewable generation uncertainty set for more efficient market clearing," in *IEEE Power Energy Soc. Gen. Meeting*, pp. 1–5, 2023.

[11] D. Bertsimas and L. Boussioux, "Ensemble modeling for time series forecasting: an adaptive robust optimization approach," *arXiv preprint arXiv:2304.04308*, 2023.

[12] D. Bertsimas, V. Gupta, and N. Kallus, "Data-driven robust optimization," *Mathematical Programming*, vol. 167, pp. 235–292, 2018.

[13] P. L. Bartlett and S. Mendelson, "Rademacher and gaussian complexities: Risk bounds and structural results," *J. Machine Learning Res.*, vol. 3, no. Nov, pp. 463–482, 2002.