

# Short-term trading of wind energy production using data-driven prescriptive optimization



*A. Stratigakos, S. Camal, T. Blondel, G. Kariniotakis*

Center PERSEE, MINES ParisTech, PSL University, Sophia Antipolis, France

*akylas.stratigakos@mines-paristech.fr*

May 25, 2021

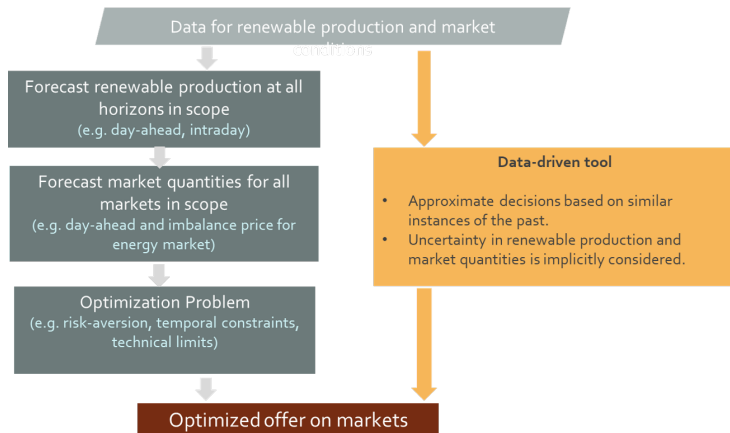
- 1 Introduction
- 2 Trading Wind Production
- 3 Methodology
- 4 Results
- 5 Conclusions

Several challenges arise when trading renewable production on short-term electricity markets:

- Decisions are affected by multiple sources of uncertainty (renewable production, market prices).
- Need to develop and deploy multiple analytics tools, which leads to a complex model chain.
- Impact of data on decision becomes unclear.

This motivates an alternative paradigm of tackling decision uncertainty, from a *prescriptive* analytics point of view.

- 1 Develop an alternative paradigm based on prescriptive methods, which simplifies the model chain.
- 2 Develop a framework to evaluate impact of data on decisions and relative optimization performance.



Renewable producer offering energy in a day-ahead (DA) market, subject to imbalance costs (*dual-price balancing market*).

**Objective:** Maximize revenue  $\rho$ , or equivalently minimize imbalance costs:

$$\rho = \pi^{DA} E^W - \underbrace{\left[ -\lambda^{\uparrow} (E^W - E^c)^- + \lambda^{\downarrow} (E^W - E^c)^+ \right]}_{\text{Imbalance Cost}},$$

where  $E^c$  the contracted energy (decision variable),  $E^W$  the stochastic production,  $\lambda^{\uparrow/\downarrow} \geq 0$  the marginal cost for upward/downward regulation [PCK07]. Standard newsvendor problem with analytical solution (*optimal quantile*).

Given a set of  $N$  sampled scenarios of uncertain production  $E^W$ , the stochastic optimization problem is formulated as:

$$\begin{aligned} & \underset{E^c, E_i^\uparrow, E_i^\downarrow}{\text{minimize}} && \frac{1}{N} \sum_{i=1}^N (-\bar{\lambda}^\uparrow E_i^\uparrow + \bar{\lambda}^\downarrow E_i^\downarrow) \\ & \text{subject to} && E^c + E_i^\uparrow + E_i^\downarrow = E_i^W, \quad i = 1, \dots, N, \\ & && E_i^\uparrow, -E_i^\downarrow \leq 0, \quad i = 1, \dots, N, \\ & && 0 \leq E^c \leq 1, \end{aligned}$$

where  $\bar{\lambda}^{\uparrow/\downarrow}$  the in-sample average of regulation costs and  $E_i^\uparrow, E_i^\downarrow$  the amount of energy shortage/ surplus at each  $i$ .

Consider a *conditional* stochastic optimization problem with uncertainty  $Y$  (e.g. renewable production) and feature data  $X$  (e.g. weather conditions):

$$\min_{z \in \mathcal{Z}} \mathbb{E}_{\mathbb{Q}}[c(z; Y) | X = x] = \min_{z \in \mathcal{Z}} \mathbb{E}_{Y \sim \mathbb{Q}_{\bar{x}}}[c(z; Y)],$$

with  $z \in \mathbb{R}^d$  the decision variables,  $\mathcal{Z}$  a convex set,  $c(\cdot)$  a cost function, and  $\mathbb{Q}_{\bar{x}}$  the true conditional marginal distribution of  $Y$ . Our goal is to approximate the problem using  $\mathcal{D}^{train} = \{(y_1, x_1), \dots, (y_N, x_N)\}$ .

## Sample Average Approximation (SAA)

Given training data on uncertainty  $y$ , the decision is approximated as

$$z^{SAA} = \operatorname{argmin}_{z \in \mathcal{Z}} \sum_{i \in N} \frac{1}{N} c(z; y_i). \quad (1)$$

Good theoretical properties, but ignores information encoded in  $x$  ("climatology decision").

## Forecast-Optimize (FO): Deterministic Forecasts

Estimate conditional expectation  $\hat{y}$ , solve deterministic optimization

$$\hat{z}^{EV} = \operatorname{argmin}_{z \in \mathcal{Z}} c(z; \hat{y}).$$

Ignores uncertainty in predictions.



## FO: Probabilistic Forecasts

Infer conditional distribution of  $Y$ , solve stochastic optimization

$$\hat{z}^{FO} = \operatorname{argmin}_{z \in \mathcal{Z}} \mathbb{E}_{y \sim \hat{\mathbb{Q}}_{\bar{x}}} [c(z; y)].$$

Requires modeling all variables and their dependencies.

## Predictions to Prescriptions (PP)

Apply a weighted SAA of the original problem [BK20]

$$\hat{z}(x) = \operatorname{argmin}_{z \in \mathcal{Z}} \sum_{i \in N} \omega_{N,i}(x) c(z; y_i),$$

with  $\omega_{N,i}(x)$  derived from local learning algorithms, e.g. Nearest Neighbors, Decision Trees. Considers uncertainty in decisions, asymptotically optimal decisions.

Applying the weighted SAA on the stochastic problem:

$$\begin{aligned} & \underset{E^c, E_i^\uparrow, E_i^\downarrow}{\text{minimize}} && \sum_{i=1}^N \omega_{N,i}(x) (-\bar{\lambda}^\uparrow E_i^\uparrow + \bar{\lambda}^\downarrow E_i^\downarrow) \\ & \text{subject to} && E^c + E_i^\uparrow + E_i^\downarrow = E_i^W, \quad i = 1, \dots, N, \\ & && E_i^\uparrow, -E_i^\downarrow \leq 0, \quad i = 1, \dots, N, \\ & && 0 \leq E^c \leq 1, \end{aligned}$$

where  $N$  is the number of historical observations in the training set.

Non-parametric local learning algorithms to estimate  $\omega_{N,i}(x)$ :

- *k Nearest Neighbors (kNN)*:  $\frac{1}{k} \mathbb{I}[x_i \in kNN]$ , for  $k$  neighbors.
- *Kernel Regression (KR)*:  $W\left(\frac{(x_i - x_t)^T (x_i - x_t)}{h(t)}\right)$ , where  $W(\cdot)$  kernel function and  $h(t)$  the kernel bandwidth. Here we consider tri-cube kernel and adaptive bandwidth.
- *Random Forest (RF)*:  $\frac{1}{B} \sum_{b \in B} \frac{\mathbb{I}[x_i \in R_l^b]}{\sum_{j \in N} \mathbb{I}[x_j \in R_l^b]}$ , where  $B$  the number of trees in the ensemble and  $R$  a tree leaf.

Workflow:

- Train local learning algorithms for prediction.
- For new query  $x$ , retrieve  $\omega_{N,i}(x)$  and solve the weighted SAA.

*Prescriptiveness*: Impact on efficacy of decisions (forecast value).

**Permutation Feature Importance**: Adapt the permutation importance [Bre01] to measure prescriptiveness.

- Estimate prescriptions, find out-of-sample expected cost (base score).
- Iterate through features, permute each one, and derive new prescriptions. Repeat process  $K$  times.
- Permutation importance measured as the *expected* cost increase.

**Coefficient of Prescriptiveness  $P$** : Unitless metric of relative optimization performance [BK20]. For a method  $i$  compare the revenue  $\hat{\rho}_i$  against the perfect foresight solution  $\rho^*$  and the SAA solution  $\hat{\rho}_{SAA}$ :

$$P_i = 1 - \frac{\hat{\rho}_i - \rho^*}{\hat{\rho}_{SAA} - \rho^*}.$$

The following approaches are compared:

- Forecast-Optimize with Random Forest (*FO-RF*): Probabilistic energy forecasting coupled with stochastic optimization (equivalent to quantile regression). Quantile Regression Forests as forecasting model.
- Predictive Prescriptions with  $\{kNN, KR, RF\}$  (*PP- $\{kNN, KR, RF\}$* ): Comparison of prescriptions derived from the different local learning algorithms.

For the coefficient of prescriptiveness  $P$  we also estimate the *SAA* and *Perfect Foresight* solution.

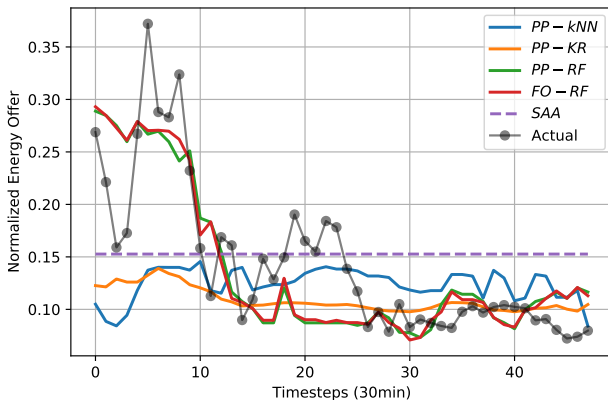
An aggregation of renewable plants participating in a DA market with a dual-price balancing price mechanism as price-taker.

- 49MW capacity, 3 Wind power plants (WP) + 1 PV power plant (SP) (16% PV share).
- Training period: from 01/2019 to 01/2020 (approximately  $N = 17000$  observations). Testing period: from 01/2020 to 05/2020.
- Feature data includes weather forecasts for the WP and SP locations.
- Market Data from French electricity market.

Predictive accuracy of the local learning algorithms (kNN, KR, RF) on forecasting renewable production. Models are tuned with cross validation. The RF algorithm shows the best overall performance.

Model	MAE (MW)	RMSE (MW)
kNN	5.09	6.70
KR	6.24	7.81
RF	<u>4.02</u>	<u>5.52</u>

Example of derived offers for a single day.



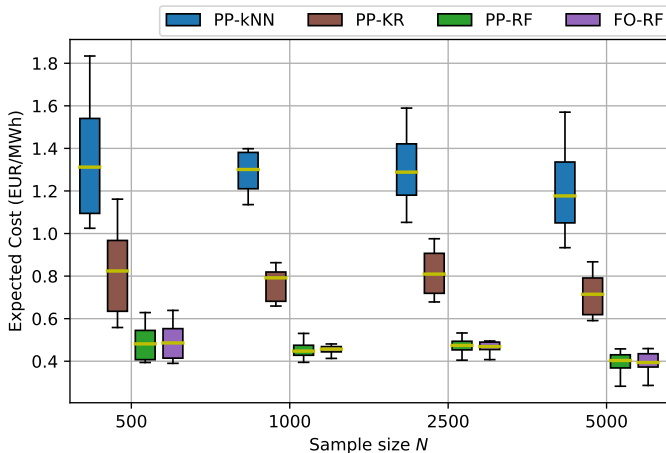


Predictive prescriptions with RF weights achieve similar trading performance as the standard FO approach. Also, while kNN outperforms KR for prediction, it shows worse prescriptive performance.

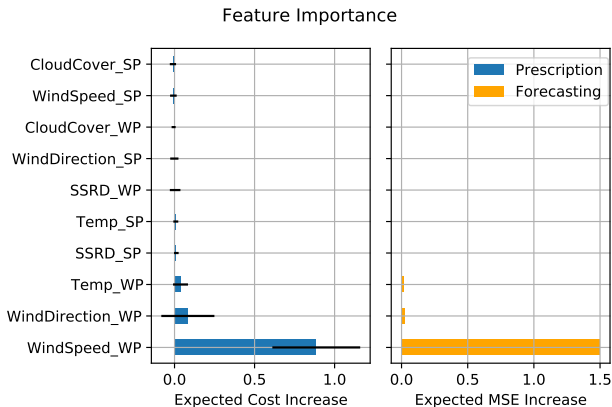
Model	Aggregated Cost (EUR)	Expected Cost (EUR/MWh)	Coefficient $P$
PP-kNN	187 219	1.32	0.04
PP-KR	108 270	0.76	0.45
PP-RF	<u>60 969</u>	<u>0.43</u>	<u>0.69</u>
FO-RF	61 067	<u>0.43</u>	<u>0.69</u>

# Results: Sensitivity to Sample Size

Testing the effect of sample size. Convergence for  $N = 5000$ .






Feature importance for prescription (left) and prediction (right). Wind Speed forecast at the WP site has the greatest impact on decisions with approximately 0.90 EUR/MWh increase in cost ( $> 200\%$  increase).



- This work examined a prescriptive analytics approach for short-term trading of wind production. We tested different local learning algorithms and examined the impact of data on the efficacy of decisions.
- The prescriptive approach with Random Forest weights led to similar results with the full stochastic solution, without the need to explicitly model distributional uncertainty.
- Future work will examine the joint impact of production and price uncertainty.

This research was carried in the frame of the European project Smart4RES (Grant No. 864337) supported by the Horizon 2020 Framework Program.



-  Dimitris Bertsimas and Nathan Kallus, *From predictive to prescriptive analytics*, Management Science **66** (2020), no. 3, 1025–1044.
-  Leo Breiman, *Random forests*, Machine learning **45** (2001), no. 1, 5–32.
-  Pierre Pinson, Christophe Chevallier, and George N Kariniotakis, *Trading wind generation from short-term probabilistic forecasts of wind power*, IEEE Transactions on Power Systems **22** (2007), no. 3, 1148–1156.