Prescriptive Trees for Integrated Forecasting and Optimization Applied in Trading of Renewable Energy

Akylas Stratigakos, Simon Camal, Andrea Michiorri, George Kariniotakis
Center PERSEE, MINES Paris, PSL University
akylas.stratigakos@minesparis.psl.eu

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 864337

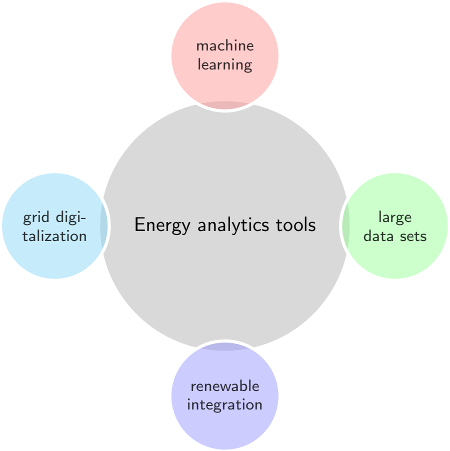
Context and Motivation

Data-driven decisions in two steps:

- *Forecast*: estimation of uncertain parameters (renewable production).
- *Optimize*: Derive an optimal set of actions (prescriptions).

Issues:

- Forecast accuracy $\neq$ forecast *value*
- Each parameter requires a separate forecasting model
- Impact of data on decisions is obscure



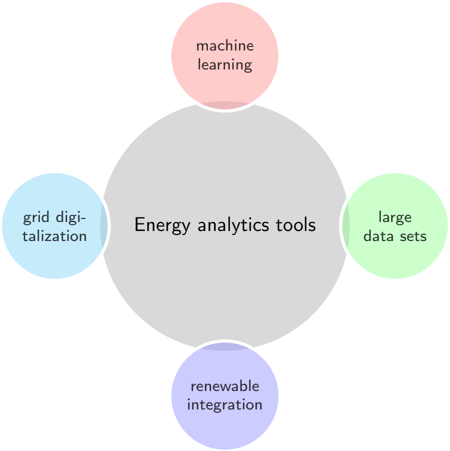
Context and Motivation

Data-driven decisions in two steps:

- *Forecast*: estimation of uncertain parameters (renewable production).
- *Optimize*: Derive an optimal set of actions (prescriptions).

Issues:

- Forecast accuracy $\neq$ forecast *value*
- Each parameter requires a separate forecasting model
- Impact of data on decisions is obscure



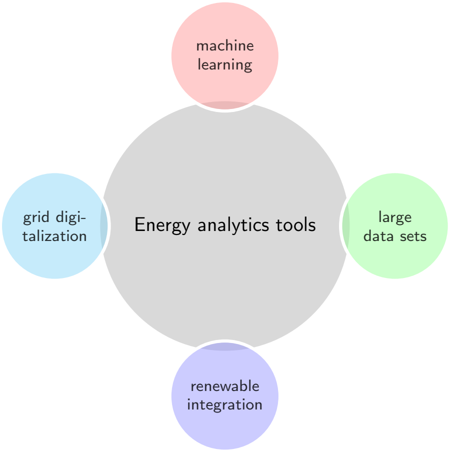
Smart4RES

Data-driven decisions in two steps:

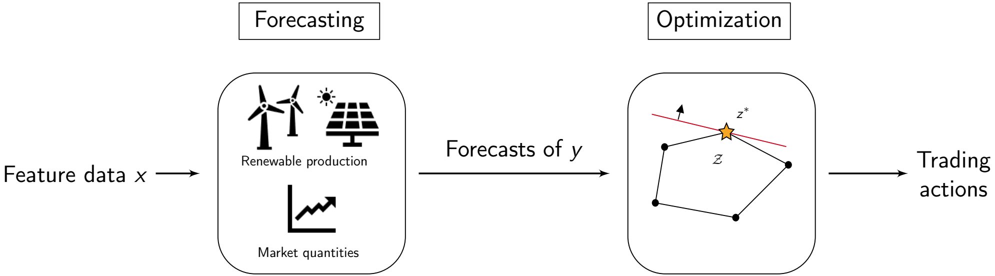
- *Forecast*: estimation of uncertain parameters (renewable production).
- *Optimize*: Derive an optimal set of actions (prescriptions).

Issues:

- Forecast accuracy $\neq$ forecast *value*
- Each parameter requires a separate forecasting model
- Impact of data on decisions is obscure

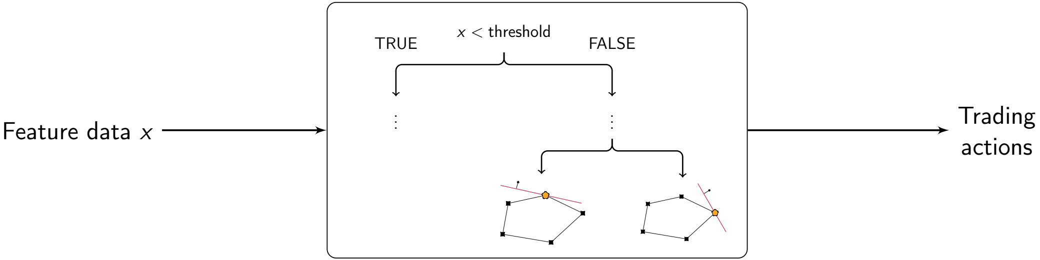


Closing the loop between forecasting and optimization to improve prescriptive performance in renewable trading applications.



Closing the loop between forecasting and optimization to improve prescriptive performance in renewable trading applications.

Integrated Forecasting-Optimization



Outline

Stochastic optimization problems with contextual information (or *prescriptive analytics problem*):

- Uncertain parameters Y: renewable production, market quantities
- Associated features X: weather forecasts, historical market data.
- Vector of decisions z: energy offers.

The goal is to minimize:

$$v = \min_{z \in \mathcal{Z}} \mathbb{E}_{\mathbb{Q}}[c(z;Y)|X = \overline{x}] = \min_{z \in \mathcal{Z}} \mathbb{E}_{y \sim \mathbb{Q}_{\overline{x}}}[c(z;Y)] \quad (1)$$

where $\mathcal{Z}$ the feasible set, $c(\cdot)$ the cost function, $\mathbb{Q}$ the joint distribution of (X, Y), $\overline{x}$ a new observation of X, and $\mathbb{Q}_{\overline{x}}$ predictive density of Y conditioned on $\overline{x}$.
Types of problems: constrained, multi-temporal, *single-stage*.

How to solve this problem

Training data set: $\{(y_i, x_i)\}_{i=1}^{n}$

Forecast, then optimize (FO): train forecasting model $f : \mathcal{X} \to \mathcal{Y}$, infer predictive densities (conditional expectation), solve stochastic (deterministic) problem.

Predictive Prescriptions: find similar observations, solve a weighted Sample Average Approximation (SAA) conditioned on $\overline{x}$ [HPB10, BK20]

$$\hat{z}(\overline{x}) = \arg\min_{z \in \mathcal{Z}} \sum_{i=1}^{n} \omega_{n,i}(\overline{x}) c(z; y_i), \tag{2}$$

$\omega_{n,i}(\overline{x})$: weights from local learning algorithms, e.g., kNN and decision trees.

- If $\omega_{n,i}(\overline{x})$ are derived by training for prediction: equivalent to **FO** with probabilistic forecasts.
- Proposed: derive $\omega_{n,i}(\overline{x})$ by directly **minimizing downstream costs**

Training data set: $\{(y_i, x_i)\}_{i=1}^n$
Forecast, then optimize (FO): train forecasting model $f : \mathcal{X} \to \mathcal{Y}$, infer predictive densities (conditional expectation), solve stochastic (deterministic) problem.
Predictive Prescriptions: find similar observations, solve a weighted Sample Average Approximation (SAA) conditioned on $\overline{x}$ [HPB10, BK20]

$$\hat{z}(\overline{x}) = \arg\min_{z \in \mathcal{Z}} \sum_{i=1}^{n} \omega_{n,i}(\overline{x}) c(z; y_i), \quad (2)$$

$\omega_{n,i}(\overline{x})$: weights from local learning algorithms, e.g., kNN and decision trees.

- If $\omega_{n,i}(\overline{x})$ are derived by training for prediction: equivalent to **FO** with probabilistic forecasts.
- Proposed: derive $\omega_{n,i}(\overline{x})$ by directly **minimizing downstream costs**

Search over functions $f : \mathcal{X} \to \mathcal{Y}$ to minimize in-sample decision costs using a weighted SAA of the original problem. Formally:

$$\min_{f \in \mathcal{F}, z^f(x_i) \in \mathcal{Z}} \quad \sum_{i \in [n]} c(z^f(x_i); y_i) \tag{3a}$$

$$\text{s.t.}$$

$$z^f(x_i) = \arg\min_{z \in \mathcal{Z}} \sum_{j \in [n]} \omega^f_{n,j}(x_i) c(z; y_j) \quad \forall i \in [n], \tag{3b}$$

where $[n] := \{1, \dots, n\}$.

Smart4RES

Prescriptive trees: trees that output prescriptions rather than predictions. Following CART [BFSO84], apply recursive binary splits:

$$\min_{j,s}\left[\min_{z_1\in\mathcal{Z}}\sum_{i\in R_1} c(z_1;y_i)+\min_{z_2\in\mathcal{Z}}\sum_{i\in R_2} c(z_2;y_i)\right]. \tag{4}$$

Inner min problems in (4) correspond to the SAA solution of each partition.

No analytical solution for constrained problems → Training is too costly!

Train ensemble with random splits [GEW06] to reduce costs:

1. At each node of each tree, sample a subset of K features from X.
2. For each feature, sample a candidate split point.
3. Solve (4) for each candidate split, apply recursively.

Smart4RES

Prescriptive trees: trees that output prescriptions rather than predictions. Following CART [BFSO84], apply recursive binary splits:

$$\min_{j,s} \left[\min_{z_1 \in \mathcal{Z}} \sum_{i \in R_1} c(z_1; y_i) + \min_{z_2 \in \mathcal{Z}} \sum_{i \in R_2} c(z_2; y_i) \right]. \tag{4}$$

Inner min problems in (4) correspond to the SAA solution of each partition. No analytical solution for constrained problems → Training is too costly!

Train ensemble with random splits [GEW06] to reduce costs:

1. At each node of each tree, sample a subset of K features from X.
2. For each feature, sample a candidate split point.
3. Solve (4) for each candidate split, apply recursively.

Prescriptive trees: trees that output prescriptions rather than predictions. Following CART [BFSO84], apply recursive binary splits:

$$\min_{j,s}\left[\min_{z_1\in\mathcal{Z}}\sum_{i\in R_1} c(z_1;y_i)+\min_{z_2\in\mathcal{Z}}\sum_{i\in R_2} c(z_2;y_i)\right]. \quad (4)$$

Inner min problems in (4) correspond to the SAA solution of each partition.
No analytical solution for constrained problems → Training is too costly!
Train ensemble with random splits [GEW06] to reduce costs:

1. At each node of each tree, sample a subset of K features from X.
2. For each feature, sample a candidate split point.
3. Solve (4) for each candidate split, apply recursively.

Explainability: Measuring prescriptiveness

Q: What is the impact of features on prescriptive performance?

A: Feature importance for predictive accuracy $\rightarrow$ adapt to measure *prescriptiveness*.

- *Mean Decrease Impurity* (MDI): For each feature, measure the expected cost reduction when it is used at node splits (in-sample, no computational cost).
- *Permutation Importance*: Shuffle feature observations, derive prescriptions, find expected cost increase (out-of-sample, high computational cost).

Other ideas: Shapley value, LIME (for single prescription), etc.

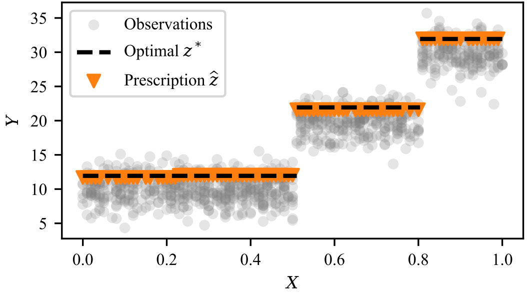
Q: What is the impact of features on prescriptive performance?

A: Feature importance for predictive accuracy → adapt to measure *prescriptiveness*.

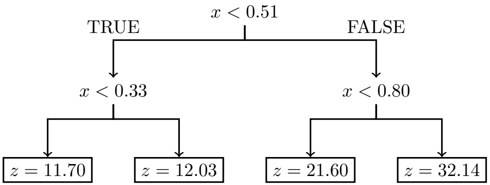
- *Mean Decrease Impurity* (MDI): For each feature, measure the expected cost reduction when it is used at node splits (in-sample, no computational cost).
- *Permutation Importance*: Shuffle feature observations, derive prescriptions, find expected cost increase (out-of-sample, high computational cost).

Other ideas: Shapley value, LIME (for single prescription), etc.

Toy newsvendor problem: cost function $c(z; Y) = 2(Y - z)^{-} + 10(z - Y)^{+}$, uncertain demand $Y = g(X) + \epsilon$, X single feature, ϵ noise.



In-sample fit



Tree structure

Outline

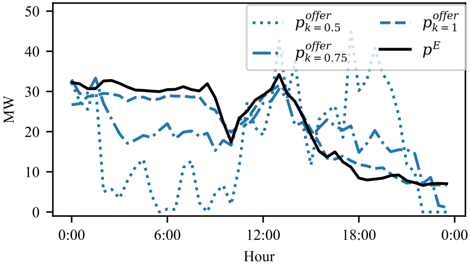
- **RES trading**: trading as a price-taker in a DA market, under single-/dual-pricing balancing mechanism.
- **RES trading with storage**: extend the above to include storage, co-optimize DA offers *and* storage control policy [SCMK22].
- **Clearing DA market**: stochastic market clearing with network constraints, test on IEEE-24 system

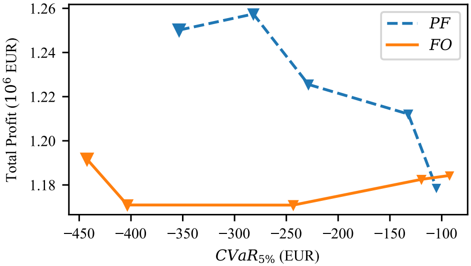
Objective: Balance trading performance (*prescriptive*) and forecast accuracy (*predictive*), jointly consider uncertainty in *both* energy and regulation penalties.

$$\min_{p^{offer}} \quad \mathbb{E}\left[\underbrace{(1-k)(-\rho^{\text{single/dual}})}_{\text{prescriptive}} + \underbrace{k\left\|p^{E}-p^{offer}\right\|_2^2}_{\text{predictive}}\right]$$
$$\text{s.t.} \quad p^{min} \leq p^{offer} \leq p^{max},$$

where $\rho^{\text{single/dual}}$ the profit function, k a design parameter controls the trade-off. For $k=0$ retrieve "0-1" or newsvendor (depends on market design) loss, for $k=1$ retrieve standard regression loss.

Illustrative results for a market with a single-price balancing mechanism:





Energy offers become riskier as k decreases.

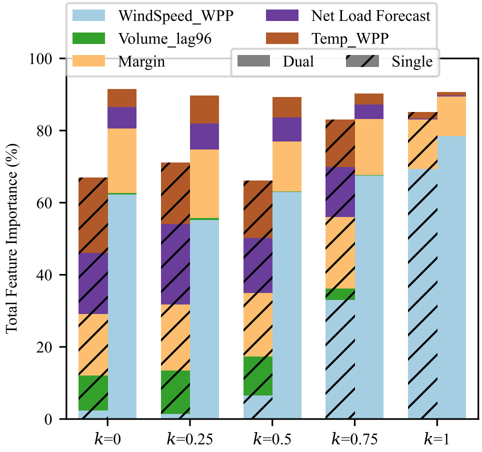
Risk-reward trade-off against the standard FO.

For $k = 0$, derive optimal trading offer:

- *Single*-price: offer is either 0 or 1, *only regulation costs* matter.
- *Dual*-price: offer is the optimal energy *quantile* given expected regulation costs.

For $k = 1$:

- Standard regression, offer expected production.



Norm. prescriptive feature importance.

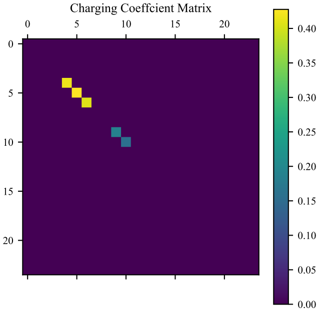
Objective: Trade-off between DA arbitrage actions and minimizing imbalance volume during real-time operation.
Linear decision rules for recourse actions: $\widetilde{p} = \widehat{p} + D\xi$, where $\widehat{p}$ the scheduled DA decisions, D lower-triangular coefficient matrix, ξ uncertainty (forecast error).

$$\min_{\mathcal{P}} \quad \mathbb{E}\left[\sum_{t=1}^{T} \underbrace{-(1-k)\pi_t^{da} p_t^{offer}}_{\text{DA arbitrage}} + \underbrace{k \left\| p_t^{output} - p_t^{offer} \right\|_2^2}_{\text{real-time control}}\right] \quad (5a)$$

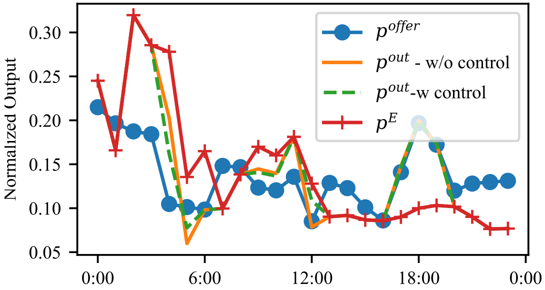
$$\text{s.t.} \quad \text{Offer limits, state transition function, storage technical constraints,} \quad (5b)$$

$$0 \leq \widetilde{p} \leq \overline{p} \quad \forall \xi \in \Xi \quad (5c)$$

Uncertainty set Ξ changes dynamically based on forecasts/ weight samples.



Coefficent matrix for charging



Offer and control actions.

Outline

Summary:

- Integrated forecasting-optimization to improve prescriptive performance in renewable trading applications.
- Explainability of decisions: impact of feature data on decision costs.
- Outperforms the "forecast, then optimize" modeling approach without requiring multiple forecast models.
- Market price forecasting relatively more important for single- than dual-price balancing mechanism.

Future work:

- Moving from batch to online learning setting.
- Interpretable model for data-driven decisions (optimal trees).

Summary:

- Integrated forecasting-optimization to improve prescriptive performance in renewable trading applications.
- Explainability of decisions: impact of feature data on decision costs.
- Outperforms the "forecast, then optimize" modeling approach without requiring multiple forecast models.
- Market price forecasting relatively more important for single- than dual-price balancing mechanism.

Future work:

- Moving from batch to online learning setting.
- Interpretable model for data-driven decisions (optimal trees).

Outline

[BFSO84] Leo Breiman, Jerome Friedman, Charles J Stone, and Richard A Olshen.
Classification and regression trees.
CRC press, 1984.

[BK20] Dimitris Bertsimas and Nathan Kallus.
From predictive to prescriptive analytics.
Management Science, 66(3):1025–1044, 2020.

[GEW06] Pierre Geurts, Damien Ernst, and Louis Wehenkel.
Extremely randomized trees.
Machine learning, 63(1):3–42, 2006.

[HPB10] Lauren Hannah, Warren Powell, and David Blei.
Nonparametric density estimation for stochastic optimization with an observable state variable.
Advances in Neural Information Processing Systems, 23:820–828, 2010.

[SCMK22] Akylas Stratigakos, Simon Camal, Andrea Michiorri, and Georges Kariniotakis.
Prescriptive trees for integrated forecasting and optimization applied in trading of renewable energy.
IEEE Transactions on Power Systems, 2022.

Smart4RES

THANK YOU !

Visit the project website at

www.smart4res.eu

@Smart4RES Project @Smart4RES