



**HAL**  
open science

# Resilient Feature-driven Trading of Renewable Energy with Missing Data

Matias Kühnau, Akylas Stratigakos, Simon Camal, Samuel Chevalier, Georges  
Kariniotakis

► **To cite this version:**

Matias Kühnau, Akylas Stratigakos, Simon Camal, Samuel Chevalier, Georges Kariniotakis. Resilient Feature-driven Trading of Renewable Energy with Missing Data. 2023. hal-04104548

**HAL Id: hal-04104548**

**<https://hal.science/hal-04104548>**

Preprint submitted on 24 May 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# Resilient Feature-driven Trading of Renewable Energy with Missing Data

Matias Kühnau  
DTU Wind  
Technical University of Denmark  
Lyngby, Denmark  
matias\_kuhnau@hotmail.com

Akylas Stratigakos  
Center PERSEE  
Mines Paris, PSL University  
06904 Sophia Antipolis, France  
akylas.stratigakos@minesparis.psl.eu

Simon Camal  
Center PERSEE  
Mines Paris, PSL University  
06904 Sophia Antipolis, France  
simon.camal@minesparis.psl.eu

Samuel Chevalier  
DTU Wind  
Technical University of Denmark  
Lyngby, Denmark  
schev@dtu.dk

George Kariniotakis  
Center PERSEE  
Mines Paris, PSL University  
06904 Sophia Antipolis, France  
georges.kariniotakis@minesparis.psl.eu

**Abstract**—Advanced data-driven methods can facilitate the participation of renewable energy sources in competitive electricity markets by leveraging available contextual information, such as weather and market conditions. However, the underpinning assumption is that data will always be available in an operational setting, which is not always the case in industrial applications. In this work, we present a feature-driven method that both directly forecasts the trading decisions of a renewable producer participating in a day-ahead market, and is resilient to missing data in an operational setting. Specifically, we leverage robust optimization to formulate a feature-driven method that minimizes the worst-case trading cost when a subset of features used during model training is missing at test time. The proposed approach is validated in numerical experiments against impute-then-regress benchmarks, with the results showcasing that it leads to improved trading performance when data are missing.

**Index Terms**—energy trading, data-driven optimization, missing data, renewable energy sources, robust optimization

## I. INTRODUCTION

The integration of renewable energy sources, such as wind and solar, in deregulated electricity markets necessitates dealing with their inherent variability and uncertainty. At the same time, as power systems are becoming increasingly data-centric [1], advanced data-driven methods, based on machine learning and optimization, enable improved decisions and help mitigate uncertainty.

The implicit assumption underpinning most applications of data-driven methods is that data will always be available when needed in an operational setting, i.e., at test time. Nonetheless, in real-world applications, data can go missing for a multitude of reasons. Network latency, equipment failures, or cyberattacks, can render input data unavailable at test time and compromise decision quality. For instance, an EC survey [2] on the timeliness of data published on the ENTSO-E

Transparency Platform<sup>1</sup> finds that “for every data domain, fewer than 40% of users reported that data were always there when needed.” To ensure that the outcome of data-driven decision-making processes remains consistent, it is pivotal to develop methods that are resilient to missing data.

### A. Related work

Leveraging data to optimize the participation of renewable energy sources in electricity markets has received a lot of attention over the last two decades. Typically, available data comes in the form of historical production data and associated contextual information (or *features*) related to weather and market conditions. Earlier works focus on leveraging contextual information to generate probabilistic forecasts of uncertain renewable production, which are subsequently used to derive optimal trading decisions—see, e.g., [3]. Recently, the focus is placed on data-driven optimization methods that directly map contextual information to decisions. We highlight [4], where meta-optimization is utilized to derive trading decisions, [5], where decision trees are trained to directly optimize for trading cost, and [6], where the trading decisions are directly forecast from contextual information.

Missing data at test time, i.e., the case when a subset of features used for model training becomes unavailable in an operational setting, poses a major challenge to model performance. This problem has been primarily examined w.r.t. forecasting applications. In [7], two approaches are proposed to handle missing data in wind power forecasting, namely retraining without the missing features and impute-then-regress. Retraining typically outperforms impute-then-regress methods, however, the number of additional models required might render it impractical. To address this challenge, [8] develops a robust optimization approach to enable model resilience while

This work was supported in part by the Smart4RES Project (Grant No 864337) funded under the Horizon 2020 Framework Program and in part by the Carnot M.I.N.E.S project Flexi4Value (Grant No 220000499) supported by ANR.

<sup>1</sup>Data on generation, transportation and consumption of electricity including information for the pan-European market

also maintaining practicality for energy forecasting applications. Nonetheless, missing data at test time for renewable trading applications have not been thoroughly examined. One exception is [6], where separate models are trained to deal with potential delays in data acquisition, effectively retraining without missing features.

### B. Aim and Contribution

In this work, inspired by [6] and [8], we develop a feature-driven model to directly forecast the trading decisions of a renewable producer participating in a day-ahead electricity market that is resilient to missing data. We leverage robust optimization to formulate a model that optimizes for the worst-case trading cost when a subset of features is missing at test time. The proposed approach is agnostic to the missingness mechanism (whereas [6] only addresses a specific missingness pattern due to data acquisition delays), maintains consistent performance, and remains practical. For validation, we examine an aggregation of wind power plants participating in a day-ahead market and compare the proposed approach against state-of-the-art methods coupled with imputation. The results illustrate that the proposed approach maintains consistent performance and outperforms impute-then-regress benchmarks.

The rest of this work is organized as follows. Section II formulates the problem of trading in a day-ahead market. Section III develops the proposed methodology. Section IV presents the experimental setup and discusses numerical results. Finally, Section V provides conclusions and perspective on future work.

## II. TRADING PROBLEM FORMULATION

In this section, we formulate the problem of trading renewable production in a day-ahead market.

### A. Trading in a Day-ahead Market

We assume a renewable producer, e.g., wind power producer, participating as a price-taker in a day-ahead market with a dual-price balancing mechanism. For each clearing period  $t$  of the day-ahead market, the producer submits an energy offer  $E_t^{da2}$ . As temporal constraints do not apply, we drop subscript  $t$  for simplicity. During real-time operation, the system operator activates balancing reserves to maintain the equilibrium between demand and supply. To recover the cost of reserve activation, all participants have to buy back (sell) the amount of energy shortage (surplus), thus balancing their individual position.

Let  $E$  be the uncertain renewable production,  $\lambda^{da}$  be the day-ahead clearing price, and  $\lambda^u, \lambda^d$  be the marginal cost of activating upward and downward reserves, respectively. We assume that if the total demand exceeds the total supply, i.e., the system requires upward regulation, then  $\lambda^u \geq \lambda^{da}$  and  $\lambda^d = \lambda^{da}$ . Conversely, if the total demand exceeds the total supply, i.e., the system requires downward regulation, then  $\lambda^d \leq \lambda^{da}$  and  $\lambda^u = \lambda^{da}$ . If  $E^{da} > E$ , i.e., the real-time production is less than the contracted amount, then

the producer buys back the difference at price  $\lambda^u$ , while if  $E^{da} < E$  sells the excess production at price  $\lambda^d$ . Hence, if the producer deviates from the contracted offer in the direction that helps decrease the amount of regulation required, no financial penalty is incurred. We further define  $\psi^u = \lambda^u - \lambda^{da}$  and  $\psi^d = \lambda^{da} - \lambda^d$  to be the non-negative upward and downward *unit regulation* costs.

The uncertain profit for a single period  $t$  can be concisely written as

$$\rho = \lambda^{da} E - [\psi^u (E^{da} - E)^+ + \psi^d (E - E^{da})^+], \quad (1)$$

where  $(t)^+ = \max(0, t)$ . Evidently, both the renewable production  $E$  and all the market quantities  $\lambda^{da}, \psi^u, \psi^d$  are unknown; thus, the trading profit  $\rho$  is also stochastic. The first term in (1) relates to the revenue from the day-ahead market, while the term in the bracket is the non-negative imbalance cost. A risk-neutral producer aims to maximize the expected trading profit. Since the first term in (1) does not depend on trading decisions, maximizing trading profit is equivalent to minimizing the imbalance cost. The problem of minimizing the imbalance cost is given by

$$\min_{E^{da}} \mathbb{E} [\max(\psi^u (E^{da} - E), \psi^d (E - E^{da}))] \quad (2a)$$

$$\text{s.t.} \quad 0 \leq E^{da} \leq \bar{E}, \quad (2b)$$

where the expectation is taken w.r.t. the joint distribution of uncertain parameters, constraint (2b) ensures that the trading decision is non-negative and respects the nominal capacity  $\bar{E}$ .

Problem (2) is an instance of the well-known newsvendor problem. Assuming we have access to the probability distribution of  $E$  and we know the true values of  $\psi^u, \psi^d$ , the optimal offer is derived analytically from

$$E^{da*} = F_E^{-1} \left( \frac{\psi^d}{\psi^d + \psi^u} \right), \quad (3)$$

where  $F_E^{-1}$  is the inverse of the cumulative distribution function of  $E$ —see [3] for a proof. Thus, the optimal offer equals the  $\tau$ -th quantile of the probability distribution, where  $\tau = \frac{\psi^d}{\psi^d + \psi^u}$ . Clearly, as both the probability distribution of  $E$  or the actual values of  $\psi^u, \psi^d$  are unknown and need to be forecast.

### B. Data-driven Approaches

We assume to have access to a training data set  $\mathcal{D} = \{(E_i, \psi_i^u, \psi_i^d, \mathbf{x}_i)\}_{i=1}^n$  of  $n$  observations of joint realizations of renewable production, unit regulation costs, and associated contextual information  $\mathbf{x}$ . Here,  $\mathbf{x}$  denotes a  $p$ -size feature vector, e.g., weather forecasts, historical production data, and market-related data, all of which are available to the producer prior to submitting the offer  $E^{da}$ . Data set  $\mathcal{D}$  can be used to approximate the solution to the original problem (2).

A first data-driven approximation is given by

$$\min_{E^{da}} \frac{1}{n} \sum_{i=1}^n \max(\psi_i^u (E^{da} - E_i), \psi_i^d (E_i - E^{da})) \quad (4a)$$

$$\text{s.t.} \quad 0 \leq E^{da} \leq \bar{E}, \quad i = 1, \dots, n, \quad (4b)$$

<sup>2</sup>We assume that the offer is always accepted at zero marginal cost.

which defines the so-called Sample Average Approximation (SAA) of (2). By linearizing the maxima terms in (4a), problem (4) can be reformulated into a linear programming (LP) problem. Effectively, the SAA finds the trading decision that minimizes the in-sample imbalance cost given data set  $\mathcal{D}$ .

While the SAA enjoys several nice theoretical properties, it does not leverage the available features  $\mathbf{x}$ . Typically, features  $\mathbf{x}$  can be utilized to train forecasting models that target the uncertain problem parameters. Conversely, we can construct a feature-driven policy that directly maps features to trading decisions and minimizes the imbalance cost, i.e., express  $E^{da}$  as a function of  $\mathbf{x}$ . Following [6], we consider an affine policy  $E^{da}(\mathbf{x}) = \mathbf{w}^\top \mathbf{x}$  parametrized by coefficients  $\mathbf{w}$ <sup>3</sup>. The problem of finding the optimal decision rules is given by

$$\min_{\mathbf{w}} \frac{1}{n} \sum_{i=1}^n \max(\psi_i^u(\mathbf{w}^\top \mathbf{x}_i - E_i), \psi_i^d(E_i - \mathbf{w}^\top \mathbf{x}_i)) \quad (5a)$$

$$\text{s.t. } 0 \leq \mathbf{w}^\top \mathbf{x}_i \leq \bar{E}, \quad i = 1, \dots, n, \quad (5b)$$

where we effectively replaced the decision variable in (4) with the feature-driven policy. For an out-of-sample realization of features, say  $\mathbf{x}_0$ , the optimal energy offer is derived directly from  $\mathbf{w}^\top \mathbf{x}_0$ , which bypasses the optimization solver. Evidently, it is possible for an out-of-sample decision to be infeasible; in this case, a projection to the feasible set  $[0, \bar{E}]$  is required.

Note that if the unit regulation costs  $\psi_i^u$  and  $\psi_i^d$  are replaced with a point forecast, i.e., their conditional expectation, then (5) is equivalent to quantile regression. In this case, we could also estimate the predictive density of  $E$  and find the optimal offer from the expected optimal quantile, as shown in (3).

### III. RESILIENT DATA DRIVEN ENERGY TRADING

An implicit assumption underpinning the feature-driven solution (5) is that the feature vector  $\mathbf{x}$  will always be available at test time. If a subset of features, however, is missing, then the performance of the feature-driven policy is compromised. In this section, we robustify (5) against missing features.

#### A. Modeling Uncertainty

To model uncertainty w.r.t. feature availability, we assume that  $\Gamma$  (integer) features, at most, could be missing from each observation. Following [8], we introduce binary variables  $\alpha_i \in \{0, 1\}^p$  that model feature availability for the  $i$ -th training observation. This is done as  $\mathbf{x}_i \odot (\mathbf{1} - \alpha_i)$  where  $\odot$  is the element-wise multiplication, where  $\alpha_{ij} = 1$  results in the deletion of the  $j$ -th feature.

In energy forecasting and trading applications, certain features may not make sense to be deleted, e.g., calendar variables. Moreover, a subset of features might be grouped, e.g., polynomial terms of the same variable or variables received from the same source, and, thus, would go missing together. To model these cases, a set of equality constraints are introduced as  $\mathbf{M}\alpha_i = \mathbf{0}$ , where  $\mathbf{M} \in \mathbb{R}^{m \cdot n}$  models the  $m$  additional constraints. For instance, if the first feature cannot be deleted,

then a row vector  $[1, \mathbf{0}]$  is appended to  $\mathbf{M}$ . Similarly, if  $\alpha_1 = \alpha_2$  then  $[-1, 1, \mathbf{0}]$  is appended to  $\mathbf{M}$ .

For the  $i$ -th observation, we model the uncertainty w.r.t. feature availability with the following discrete uncertainty set

$$\mathcal{U}_i = \{\alpha_i \mid \alpha_i \in \{0, 1\}^p, \sum_{j \in [p]} \alpha_{ij} = \Gamma, \mathbf{M}\alpha_i = \mathbf{0}\}. \quad (6)$$

As all vertices of (6) occur at integer values, we can replace (6) with its convex hull, given by (7)

$$\mathcal{A}_i = \{\alpha_i \mid \mathbf{0} \leq \alpha_i \leq \mathbf{1}, \sum_{j \in [p]} \alpha_{ij} = \Gamma, \mathbf{M}\alpha_i = \mathbf{0}\}, \quad (7)$$

which is a polyhedral uncertainty set. Therefore, using (7) instead of (6) results in equivalent solutions.

#### B. Resilient Feature-Driven Model

In this section, we develop the proposed resilient feature-driven model. First, the feature-driven problem (5) is equivalently written as

$$\min_{t, \mathbf{u}, \mathbf{w}} t, \quad (8a)$$

$$\text{s.t. } \frac{1}{n} \sum_{i=1}^n u_i \leq t, \quad (8b)$$

$$\psi_i^u(-E_i + \mathbf{w}^\top \mathbf{x}_i) \leq u_i, \quad i \in [n] \quad (8c)$$

$$\psi_i^d(E_i - \mathbf{w}^\top \mathbf{x}_i) \leq u_i, \quad i \in [n] \quad (8d)$$

$$0 \leq \mathbf{w}^\top \mathbf{x}_i \leq \bar{E}, \quad i \in [n] \quad (8e)$$

where we effectively linearized the inner maxima terms, with  $t, u_i$  denoting auxiliary variables,  $\mathbf{u}$  being an  $n$ -size vector that comprises of  $u_i$ , and  $[n]$  being shorthand for  $\{1, \dots, n\}$ .

Using (7), we robustify the feature-driven models (8) against uncertainty w.r.t. to feature availability. The robust version is given by

$$\min_{t, \mathbf{u}, \mathbf{w}} t, \quad (9a)$$

$$\text{s.t. } \frac{1}{n} \sum_{i=1}^n u_i \leq t, \quad (9b)$$

$$\psi_i^u(-E_i + \mathbf{w}(\mathbf{x}_i \odot (\mathbf{1} - \alpha_i))) \leq u_i, \quad i \in [n], \forall \alpha_i \in \mathcal{A}_i, \quad (9c)$$

$$\psi_i^d(E_i - \mathbf{w}(\mathbf{x}_i \odot (\mathbf{1} - \alpha_i))) \leq u_i, \quad i \in [n], \forall \alpha_i \in \mathcal{A}_i, \quad (9d)$$

$$0 \leq \mathbf{w}(\mathbf{x}_i \odot (\mathbf{1} - \alpha_i)) \leq \bar{E}, \quad i \in [n], \forall \alpha_i \in \mathcal{A}_i, \quad (9e)$$

which we term *resilient feature-driven* model. Effectively, (9) finds the set of linear decision rules that map features to decisions that minimize the worst-case trading cost when  $\Gamma$  features are missing. Constraints (9c)-(9e) are robust constraints. As we are dealing with an LP problem and a polyhedral uncertainty set, (9) can be reformulated into a tractable problem using techniques from robust optimization [9, Ch. 2].

For illustration, consider constraint (9c). As this constraint needs to be satisfied for all realizations of  $\alpha_i$ , it suffices to be

<sup>3</sup>To model the bias term, we assume a vector of ones is padded to  $\mathbf{x}$ .

satisfied for the worst-case scenario. For the  $i$ -th observation, assuming  $\psi_i^u > 0$ , (9c) is equivalently written as

$$-E_i + \mathbf{w}\mathbf{x}_i - \max_{\alpha_i \in \mathcal{A}_i} (\mathbf{w} \odot \mathbf{x}_i)^\top \alpha_i \leq u_i \frac{1}{\psi_i^u}. \quad (10)$$

The inner max problem is given by

$$\begin{aligned} \max_{\alpha_i} & (\mathbf{w} \odot \mathbf{x}_i)^\top \alpha_i, \\ \text{s.t.} & \quad \alpha_i \leq \mathbf{1} : \mu_i \geq \mathbf{0}, \\ & \quad \sum_{j \in [p]} \alpha_{ij} = \Gamma : \zeta_i, \\ & \quad \mathbf{M}\alpha_i = \mathbf{0} : \pi_i, \\ & \quad \alpha_i \geq \mathbf{0}, \end{aligned}$$

where  $\mu_i, \pi_i, \gamma_i$  are dual variables of appropriate size. As this problem is linear in  $\alpha_i$ , it can be replaced in (10) by its dual,

$$\min_{\mu_i, \gamma_i, \pi_i} \sum_{j \in [p]} \mu_{ij} + \gamma_i \Gamma, \quad (12a)$$

$$\mu_i + \gamma_i + \mathbf{M}\pi_i \geq \mathbf{x}_i \odot \mathbf{w}, \quad (12b)$$

$$\mu_i \geq \mathbf{0}, \quad (12c)$$

where the min operator becomes redundant. The rest of the constraints are reformulated similarly. Note that (9) ensures that in-sample trading decisions are always feasible. Nonetheless, it is possible for out-of-sample decisions to be infeasible. In this case, the inferred decision is projected onto the feasible set  $[0, \bar{E}]$ .

#### IV. NUMERICAL EXPERIMENTS

This section presents the experimental setup used to validate the proposed model and discusses our numerical results.

##### A. Experimental Setup and Input Data

First, we formulate different data-driven methods to derive trading decisions. Next, we examined their performance when features are missing. The following models are tested:

- *SAA*: the SAA solution (4) that does not account for contextual information and is not affected by missing features.
- *FD*: the feature-driven method (5).
- *EV*: this involves offering the expected production value, derived from standard forecasting. For simplicity, EV is the same as FD assuming  $\psi^u$  and  $\psi^d$  are equal.
- *RF*: a random forest model that maps data to decisions, as described in [5], using the same features as FD.
- *ResFD*: the proposed resilient (i.e., robust) version of FD.

Initial experiments showed that including the realizations of  $\psi_i^u, \psi_i^d$  in the feature-driven models led to worse performance. In fact, the best results were obtained using the in-sample means, e.g., replacing  $\psi_i^u$  with  $\frac{1}{n} \sum_{i \in [n]} \psi_i^u$ . This is also the case in [6] and effectively amounts to using the in-sample mean of unit regulation costs as the respective point forecast.

For models that cannot handle missing data, i.e., FD, EV, and RF, we follow the impute-then-regress approach with

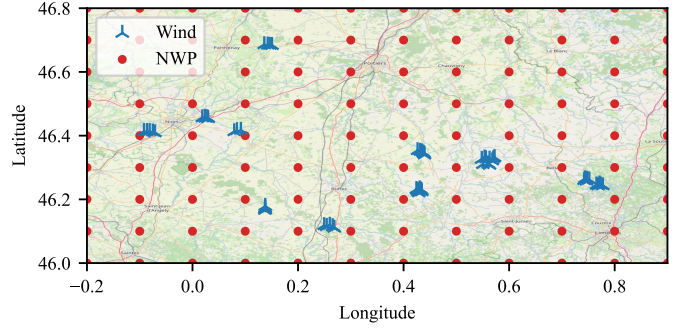


Fig. 1. Location of wind farms and NWP grid points.

mean imputation. For the proposed ResFD, missing data are set to zero and a different model is trained for each value of  $\Gamma$  and the respective model is used at test time. Also, for  $\Gamma = 0$ , ResFD is equivalent to FD.

Regarding input data, we use real-world production data, numerical weather predictions (NWP), and market data. All data sets span the same period, from October 2018 to September 2020, and are normalized prior to model training. The first year of data is reserved for model training, while the remaining data are used for testing.

The power production data comprise an aggregation of 13 wind farms located in mid-west France with a total nominal capacity of 120MW. The measured power production is effectively our target and represents the “perfect” day-ahead offer for each observation.

The NWP are used to create the feature vector  $\mathbf{x}$ . The NWP consist of weather forecasts in a spatial grid with a half-hourly resolution, namely wind speed, wind direction, solar irradiance, and cloud coverage. To model the nonlinear relationship between wind speed and wind power production, we further add a quadratic and cubic transformation of the wind speed variable. For each of the 13 wind farms, we find the NWP grid point that is closest in terms of Euclidean distance and append it to the feature vector, resulting in a total of 52 features. Fig. 1 illustrates the geographical distribution of wind farms and respective NWP. Each grid point determines a group of variables, i.e., all the variables go missing together (13 groups in total). Hence, the possible combinations of missing features are  $\sum_{k=0}^{13} \binom{13}{k}$ . In turn, this number renders the retraining method [7] impractical.

For market data, we consider data from the French electricity market. The market data are used to estimate the in-sample mean of the balancing penalties and for out-of-sample validation.

##### B. Results

First, we briefly compare baseline performance without missing features. Table I presents the mean imbalance cost in EUR/MWh. Overall, all feature-driven models significantly outperform SAA. Specifically, EV, FD, and ResFD are the best-performing models with an expected cost of 0.72 EUR/MWh, closely followed by RF.

TABLE I  
MEAN IMBALANCE COST (EUR/MWH) WITHOUT MISSING FEATURES.

SAA	EV	FD	RF	ResFD
2.32	0.72	0.72	0.77	0.72

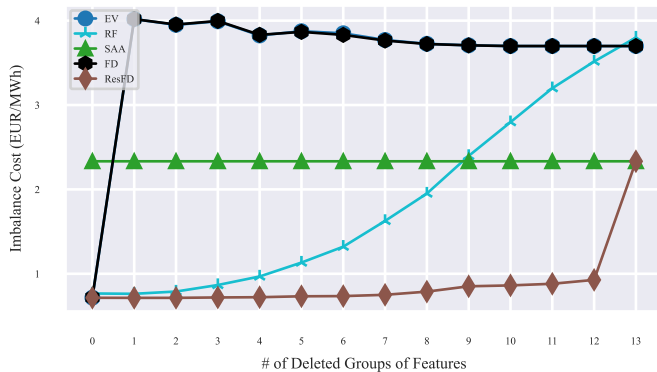


Fig. 2. Performance versus the number of missing feature groups.

Next, we evaluate performance when features are missing at test time. For each test observation, we randomly remove  $\Gamma$  NWP features, repeating the experiment 10 times.

Fig. 2 plots the mean imbalance cost over all the experiments versus the different values of missing features. Note that for zero missing features, all models perform according to Table I. As expected the overall trading performance decreases when data are missing at the time. The proposed ResFD is the best-performing model when data are missing and showcases fairly consistent performance with a relatively small increase in imbalance cost. The only exception is for the extreme case when  $\Gamma = 13$ , when SAA slightly outperforms ResFD. Regarding the impute-then-regress benchmarks, RF is fairly resilient for up to 3 missing feature groups, while both EV and FD are considerably worse for even a single missing feature group. Overall, ResFD leads to approximately 25% improved trading performance against the second-best model when data are missing.

We further conduct a sensitivity analysis w.r.t. the number of test observations that incur missing data. To this end, we sample a percentage of observations with missing data, and for each observation, we sample a number of missing feature groups, repeating the experiment 10 times.

Table II summarizes the results of the sensitivity analysis. Overall, ResFD outperforms the impute-then-regress methods, and the effect is more pronounced as the percentage of observations with missing data increases. Specifically, the average improvement over the second-best method, RF, is approximately 10%.

rier

## V. CONCLUSION

In this work, we developed a feature-driven model that directly forecasts the trading decisions for a renewable producer

TABLE II  
MEAN IMBALANCE COST (EUR/MWH) VERSUS PERCENTAGE OF OBSERVATIONS WITH MISSING FEATURES.

Percentage	SAA	RF	EV	FD	ResFD
5	2.33	0.82	0.86	0.86	<b>0.76</b>
10	2.33	0.88	1.01	1.01	<b>0.81</b>
25	2.33	1.03	1.44	1.44	<b>0.94</b>

participating in a day-ahead market and is resilient to missing features. Specifically, we formulated a model that minimizes the worst-case trading cost when a subset of features is missing at test time. To validate the proposed approach, we compared performance against a number of impute-then-regress benchmarks considering a day-ahead market with a dual-price balancing mechanism. Overall, the proposed approach led to approximately 25% reduction in trading costs compared to the best impute-then-regress method when data were missing from the test set. A sensitivity analysis examined performance when a subset of the test set had missing data and further validated the efficacy of the proposed method.

Future work could focus on trading applications in electricity markets with a single-price balancing mechanism. Further, the proposed methodology could also be applied to robustify the data-driven methods that aim at forecasting the solutions to constrained optimization problems given associated features.

## REFERENCES

- [1] E. Union, "Digitalisation of the energy system," [https://energy.ec.europa.eu/topics/energy-systems-integration/digitalisation-energy-system\\_en](https://energy.ec.europa.eu/topics/energy-systems-integration/digitalisation-energy-system_en), 2022, [Online; accessed 3-March-2023].
- [2] European Commission, "A review of the ENTSO-E transparency platform," 2017. [Online]. Available: [https://energy.ec.europa.eu/system/files/2018-05/review\\_of\\_the\\_entso\\_e\\_platform\\_0.pdf](https://energy.ec.europa.eu/system/files/2018-05/review_of_the_entso_e_platform_0.pdf)
- [3] P. Pinson, C. Chevallier, and G. N. Kariniotakis, "Trading wind generation from short-term probabilistic forecasts of wind power," *IEEE Transactions on Power Systems*, vol. 22, no. 3, pp. 1148–1156, 2007.
- [4] N. Mazzi and P. Pinson, "Purely data-driven approaches to trading of renewable energy generation," in *International Conference on the European Energy Market, EEM*, vol. 2016-July, 2016.
- [5] A. C. Stratigakos, S. Camal, A. Michiorri, and G. Kariniotakis, "Prescriptive trees for integrated forecasting and optimization applied in trading of renewable energy," *IEEE Transactions on Power Systems*, vol. 37, no. 6, pp. 4696 – 4708, Nov 2022.
- [6] M. A. Munoz, J. M. Morales, and S. Pineda, "Feature-driven improvement of renewable energy forecasting and trading," *IEEE Transactions on Power Systems*, vol. 35, no. 5, pp. 3753–3763, 2020.
- [7] R. Tawn, J. Browell, and I. Dinwoodie, "Missing data in wind farm time series: Properties and effect on forecasts," *Electric Power Systems Research*, vol. 189, p. 106640, 2020.
- [8] A. Stratigakos, P. Andrianesis, A. Michiorri, and G. Kariniotakis, "Towards resilient energy forecasting: A robust optimization approach," *IEEE Transactions on Smart Grid*, pp. 1–1, 2023.
- [9] D. Bertsimas and D. den Hertog, *Robust and adaptive optimization*. Dynamic Ideas LLC, 2020, vol. 958.