# Prediction Markets with Intermittent Contributions

Michael Vitali, Pierre Pinson
Dyson School of Design Engineering, Imperial College London, London, United Kingdom
{m.vitali24@imperial.ac.uk, p.pinson@imperial.ac.uk}

***Abstract*—Although both data availability and the demand for accurate forecasts are increasing, collaboration between stakeholders is often constrained by data ownership and competitive interests. In contrast to recent proposals within cooperative game-theoretical frameworks, we place ourselves in a more general framework, based on prediction markets. There, independent agents trade forecasts of uncertain future events in exchange for rewards. We introduce and analyse a prediction market that (i) accounts for the historical performance of the agents, (ii) adapts to time-varying conditions, while (iii) permitting agents to enter and exit the market at will. The proposed design employs robust regression models to learn the optimal forecasts' combination whilst handling missing submissions. Moreover, we introduce a pay-off allocation mechanism that considers both in-sample and out-of-sample performance while satisfying several desirable economic properties. Case-studies using simulated and real-world data allow demonstrating the effectiveness and adaptability of the proposed market design.**

***Index Terms*—Online learning, forecast combination, prediction markets, robust regression, renewable energy.**

## I. Introduction

Growing concerns about climate change and energy independence led to a rapid expansion of solar and wind energy production. Forecasting plays a crucial role in decision-making processes, requiring high-quality data and advanced models. The increase in data availability represents an opportunity to refine these techniques and increase forecast quality, and enable companies to increase profits or reduce costs. However, data ownership issues often hinder progress, as companies are reluctant to share information and to collaborate due to competitive interests and security concerns.

A solution to increase collaboration is to give incentives in exchange for these data. This can be done in different ways, depending on whether the companies are willing to share data. If this is the case, different methods have been proposed for renewable power forecasting in the literature exploiting spatio-temporal relations. For solar power forecasting, a model incorporating exogenous data from different sources was developed [1], while a similar approach has been applied to both wind and solar power forecasting [2]. In wind power forecasting, sparse spatio-temporal models were introduced in [3] and later extended in an online learning context [4]. Furthermore, to avoid the loss of privacy companies could face, privacy-preserving solutions using distributed learning methods have been proposed [5], with further advancements addressing online settings [6], [7]. However, all these solutions assume that the agents are willing to collaborate to improve forecasts and that they act rationally and truthfully. This is not always the case in practice. An alternative approach is to view the problem in a more general framework, through prediction markets, in which companies choose to share their individual forecasts and get rewarded for their contribution to the resulting aggregate forecast.

In this context, prediction markets have been increasingly studied in the last decade, gaining popularity across fields [8], with two types of market being proposed: *(i)* contribution-based and *(ii)* "winner takes it all". The latter was proposed by [9], where only the best solution is rewarded. However, this type of market ignores the fact that forecasts other than the best one can still provide additional information. Instead, a contribution-based market rewards all the participants based on the amount of information they brought into the final forecast (as for the example of [10]).

Nevertheless, existing solutions fail to consider some essential aspects of real-world applications such as: *(i)* real-time implementation, *(ii)* historical contributions of the participants, and *(iii)* the ability to accommodate intermittent participation, meaning that participants may join or exit the market at any time. The first two challenges can be seen as an online learning problem, whilst the third one as a missing data problem. For the latter, the literature has proposed robust variants of linear regression [11], [12] and online imputation strategies [13]. In addiction, a pay-off allocation mechanism must be designed to determine how rewards are distributed among sellers, reflecting the informational value they contribute to the final forecast. Also in this case, the allocation must reflect historical contributions rather than only instantaneous performance. In the literature, the most widely adopted approach is the Shapley-value-based allocation [14] (see, for example, [7]) and online versions of Shapley value calculation have also been employed for dynamic pay-off distribution [15]. However, these approaches primarily focus on in-sample pay-offs and fail to account for out-of-sample performance, overlooking the actual contribution of sellers to genuine forecast improvements. Finally, any allocation rule must also satisfy several desirable economic properties to ensure fairness and stability.

We introduce and analyse a prediction market[1] that aims at tackling the challenges described in the above. Building upon [10], our market design accommodates different clients and sellers that interact via a market operator. This results in the following contributions:

[1]Reproducibility code: https://github.com/MichaelVitali/prediction_markets

- The market operator optimally combine input forecasts, while allowing sellers to enter and exit the market at will. This is done through the use of a robust linear regression model that is able to predict in presence of missing data. Additionally, this approach is extended, for the first time, to operate in an online setting.
- A pay-off allocation is proposed accounting for both in-sample and out-of-sample rewards. For the latter, we use a scoring function that assesses the accuracy of each reported forecast and rewards accordingly. Instead, for the in-sample reward, we use a time-varying Shapley value. This is done to reward consistency and informational value provided by the seller over time.

The remainder of the manuscript is organized as follows: Section II introduces the main concepts of the prediction market design. Then, Section III describes its methodological components, from adaptive robust linear regression to pay-off allocation. It is followed by Section IV, showing different test cases and corresponding results. Finally, Section V concludes the paper and offers perspectives for future work.

## II. Preliminaries

### A. Prediction Markets

Market-based analytics can be broadly categorized into data markets and information markets, depending on whether the exchanged good is raw data or derived information. We focus on prediction markets, a specific subset of information markets. In prediction markets, participants trade forecasts about uncertain future events. These markets combine individual predictions to render a final forecast, communicated to the buyer. Contributors to prediction markets are rewarded in proportion to the value their individual forecasts add to the final forecast. This is done via a mechanism with formal mathematical guarantees concerning desirable economic properties such as budget balance, symmetry, etc.

### B. Market Setup

Our market design involves the interaction of multiple clients and sellers, through a central market operator. In its most general form, the market deals with nonparametric probabilistic forecasts, for a continuous variable of interest, based on a set of quantiles for various nominal levels. We define the participant roles as follows:

- client $c_i$: individual who requires a forecast for a variable of interest (i.e., wind power generation at a set of lead times). The client may or may not already possess their own forecast; in our setup, we focus on clients without an available forecast. We further assume that the client can evaluate the utility $U_t$ derived from a forecast, for instance in terms of profit gains or cost reductions, which is used as a basis for payment for forecast improvement.
- seller $s_i$: forecaster willing to provide forecasts for the variable of interest, and in the format required by the client. We assume that each seller is contracted and allowed to miss submitting predictions for a maximum proportion of time, e.g. 10% (which may or may not be at random). We denote the set of agents participating in the market as $\mathbf{S} = \{s_1, s_2, \dots, s_n\}$.
- market operator: central entity responsible for managing the market through a centralized platform. It enables clients to post forecasting tasks for a variable of interest, while allowing sellers to submit their predictions in response to these tasks. Beyond facilitating this exchange, the operator is also responsible for producing the final forecasts, collecting the payment of the clients, and re-distributing rewards among participating sellers following a well-defined allocation rule.

For clarity and simplicity, the following sections of this paper will focus on a setting with a single client interacting with multiple sellers.

### C. Overview of Market Operation

Fig. 1 provides an overview of the market and a high-level description of the sequential steps for the nominal level $\tau$. The process begins with the market opening a session for a specified task. This session remains open for a limited time, during which sellers may submit their forecasts. If the $i$-th forecast is not submitted, the model finds the missing information from other forecasters (see III-B for details). Once the session closes, the market operator aggregates the submitted forecasts into a combined prediction and delivers it to the client. After the event occurs, the client reports the realization $y_{t+1}$ to the market operator together with the generated utility $U_t$. At this stage, the market operator performs two parallel operations: *(1)* updating the model and propagating the new weights forward, and *(2)* computing pay-off allocations so that each seller receives their corresponding reward. In addition, the updated in-sample values are incorporated into the next model run.

### D. Forecast combination

We define the set of input forecasts as $\hat{\mathbf{X}}_t = \{\hat{\mathbf{X}}_{1,t}, \hat{\mathbf{X}}_{2,t}, \dots, \hat{\mathbf{X}}_{n,t}\} \in \mathbb{R}^{n\times k\times m}$, from the set of sellers $\mathbf{S}$, with $m$ the number of quantiles (with different nominal levels) and $k$ the lead time. The forecast provided by the $i$-th seller is $\hat{\mathbf{X}}_{i,t} = \{\hat{\mathbf{x}}_{i,t+1|t}, \dots, \hat{\mathbf{x}}_{i,t+k|t}\}$. The market operator generates the aggregated forecasts as a weighted average of input forecasts, for each nominal level $\tau$. The weight $w_i$ (for seller $s_i$) reflects their relative historical performance.

For simplicity, when describing methodological elements, we focus on the case $k = 1$ and a fixed nominal level $\tau$. The set of input forecasts then reduces to $\hat{\mathbf{x}}_t^{(\tau)} = \{\hat{x}_{1,t+1|t}^{(\tau)}, \dots, \hat{x}_{n,t+1|t}^{(\tau)}\}$. The combined forecast is given as a convex combination of the input forecasts, i.e.,

$$\hat{y}_{t+1|t}^{(\tau)} = \sum_{i=1}^{n} w_i \hat{x}_{i,t+1|t}^{(\tau)}, \quad \sum_{i=1}^{n} w_i = 1, \quad w_i \geq 0 \qquad (1)$$

In our setup, the weights $w_i$'s are non-negative and sum to one, to improve interpretability. In the general case, such constraints are not strictly necessary: weights could be negative

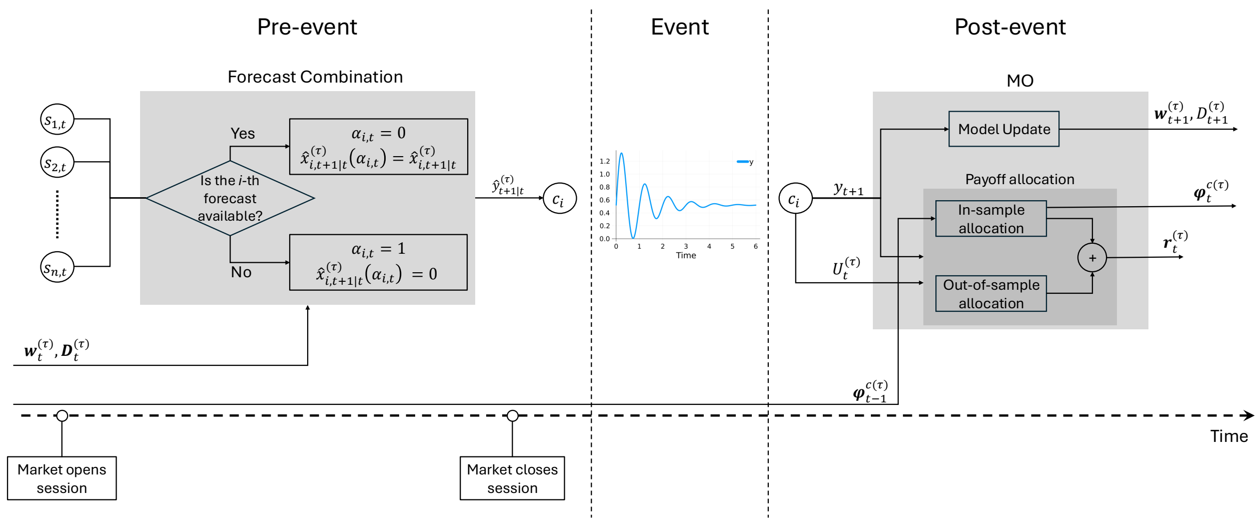


Fig. 1. Market design overview

and not sum to one. For a comprehensive overview of the state of the art with forecast combination, we refer the reader [16].

### E. Pay-off allocation

A pay-off function is central to the design of a market mechanism as it distributes the generated utility among the market players (sellers) according to their performances. For this reason, it is critical to design a pay-off function that encourages market participation, whilst reflecting seller's contributions. The pay-off function is characterized by several economic properties that can be mathematically proven, such as budget balance, zero element, etc. Furthermore, the pay-off function can be divided into in-sample and out-of-sample pay-off allocation. The former is used to reward forecast consistency and informational value provided by the sellers in time, whilst the out-of-sample is used to reward good forecasts.

## III. Methodology

### A. Forecast combination

In this work, the forecast combination is performed using a Linear Regression (LR) model, as introduced in (1). To adapt to potential changes in market dynamics over time, the method is extended to an online learning setting, allowing the model to continuously update and learn the optimal combination of forecasts (weights) in time. For any quantile $\tau$ of interest a different LR is learned enabling separate learning dynamics for each level.

The model's weights are updated online using online gradient descent, according to the following rule (the detailed derivation of the loss function is provided in Appendix B)

$$\mathbf{w}_{t+1}^{(\tau)} = \mathbf{w}_t^{(\tau)} - \eta \nabla \mathcal{L}(y_{t+1}, \hat{y}_{t+1|t}^{(\tau)}) \tag{2}$$

where $\eta$ is the learning rate, $\mathbf{w}_t^{(\tau)}$ is the weights vector assigned to the sellers at time $t$, $\hat{y}_{t+1|t}^{(\tau)}$ is the aggregate forecast, $y_{t+1}$ is the observation, and $\mathcal{L}$ is a convex loss function. Moreover, a project step is applied to satisfy the constraints in (1). In this work, the loss function is the quantile loss (or, pinball loss). It is defined as follows

$$\mathcal{L}^{(\tau)}(y, \hat{y}) = \begin{cases} (y - \hat{y})\tau & \text{if } y \geq \hat{y} \\ (\hat{y} - y)(1 - \tau) & \text{if } \hat{y} > y \end{cases} \tag{3}$$

with $\hat{y}$ being the aggregated forecast. Note that since the quantile loss is not differentiable in zero, we used the sub-gradient recently studied in [17].

### B. Adaptive Robust Linear Regression

The forecast combination proposed using linear regression assumes that the sellers submits their prediction at every market session, assumption that is often unrealistic in real-world scenarios. Since we want to allow sellers to participate at will, we address this limitation adopting a robust variant of the LR model capable of handling missing forecasts. This solution was first introduced in [11] and later applied in [12]. The core idea of this method is to learn a linear correction matrix $\mathbf{D}^{(\tau)}$ among input forecast, and use this correction to modify the combination weights when some input forecasts are unavailable. In essence, the model compensates for the missing information by extracting additional one from the remaining available forecasts. In this work, we extend this method to the online learning setting.

To model the availability of forecasts, we introduce a binary variable $\alpha_{i,t}$ which takes the value 1 if the $i$-th forecast is unavailable at time $t$, and 0 otherwise. Using this, we redefine the forecast vector as

$$\hat{x}_{i,t+1|t}^{(\tau)}(\alpha_{i,t}) = \begin{cases} \hat{x}_{i,t+1|t}^{(\tau)} & \text{if } \alpha_{i,t} = 0 \\ 0 & \text{if } \alpha_{i,t} = 1 \end{cases} \tag{4}$$

With this in mind, we can define the forecast combination in vector form as

$$\begin{aligned} \hat{y}_{t+1|t}^{(\tau)} &= [\boldsymbol{\theta}^{(\tau)}(\boldsymbol{\alpha}_t)]^\top \hat{\mathbf{x}}_t^{(\tau)}(\boldsymbol{\alpha}_t) \\ &= (\mathbf{w}^{(\tau)} + \mathbf{D}^{(\tau)}\boldsymbol{\alpha}_t)^\top \hat{\mathbf{x}}_t^{(\tau)}(\boldsymbol{\alpha}_t) \end{aligned} \tag{5}$$

From this formula, we can see that $\mathbf{w}^{(\tau)}$ represents the model with all the features available and $D_{i,j}^{(\tau)}$ is the linear correction applied to $w_i^{(\tau)}$ when the $j$-th forecast in missing.

We now extend the robust approach to an online learning framework. The updates are defined as follows (the detailed derivation of the loss function is provided in Appendix B).

$$\mathbf{w}_{t+1}^{(\tau)} = \mathbf{w}_t^{(\tau)} - \eta \nabla \mathcal{L}(y_{t+1}, \hat{y}_{t+1|t}^{(\tau)}) \tag{6}$$

$$\mathbf{D}_{t+1}^{(\tau)} = \mathbf{D}_t^{(\tau)} - \eta \nabla \mathcal{L}(y_{t+1}, \hat{y}_{t+1|t}^{(\tau)}) \tag{7}$$

with $\eta$ learning rate. Also in this case, a projection step is applied to $\mathbf{w}_{t+1}^{(\tau)}$ to satisfy the constraints in (1). From (5)-(7), it is evident that the updates depend on the forecast availability vector $\boldsymbol{\alpha}_t$. Specifically, when computing the gradient of the loss function, we observe that the weights corresponding to missing forecasts are not updated, whilst the linear correction matrix is updated only when at least one forecast is missing.

### C. Pay-off Allocation

When the client provides the true realization $\mathbf{y}_t$, the total available reward $U_t$ has to be split accordingly to both the in-sample and out-of-sample allocations. The amount of allocation allocated to the two is defined by $\delta$. Let's assume that the total reward is divide equally for each quantile level ($U_t^{(\tau)}$). We have that the reward for the $i$-th sellers at time $t$ for the quantile level $\tau$ is

$$r_{i,t}^{(\tau)} = U_t^{(\tau)}[\delta r_{i,t}^{\text{is}(\tau)} + (1-\delta) r_{i,t}^{\text{oos}(\tau)}] \tag{8}$$

where $r_{i,t}^{\text{is}(\tau)}$ is the in-sample allocation and $r_{i,t}^{\text{oos}(\tau)}$ the out-of-sample one. Finally, the total reward for each seller is defined as $r_{i,t} = \sum_{\tau=1}^{m} r_{i,t}^{(\tau)}$.

*1) In-sample allocation:* The primary objective of the in-sample allocation is to reward sellers who consistently contribute valuable information to the market. This problem is seen as a cooperative game, where the allocation of the total reward is determined using Shapley values (the definition of which is available at, e.g., [14]). The marginal contribution of seller $s_i$ at time $t$ is defined as

$$\varphi_{i,t}^{\text{s}(\tau)} = \begin{cases} \text{SHAP}_{i,t}^{(\tau)}(y_{t+1}, \hat{\mathbf{x}}_t^{(\tau)}, \boldsymbol{\theta}_t^{(\tau)}) & \text{if } \alpha_{i,t} = 0 \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

If a seller is unavailable at time $t$, their Shapley value is set to zero. For the remaining sellers, the Shapley values are computed using the corrected weights. In doing so, the method acknowledges that sellers with strong performance helps mitigate the negative impact of missing forecasts on the final aggregated prediction giving them higher value.

However, our goal is to reward sellers not only for their current contributions, but also for the historical informational value they have provided. To achieve this, we employ an online variant of the Shapley value, which is updated over time as

$$\varphi_{i,t}^{\text{c}(\tau)} = \lambda \varphi_{i,t-1}^{\text{c}(\tau)} + (1-\lambda) \varphi_{i,t}^{\text{s}(\tau)} \tag{10}$$

with $\lambda$ forgetting factor.

Finally, the in-sample reward is calculated as

$$r_{i,t}^{\text{is}(\tau)} = \begin{cases} \frac{\max(0, \varphi_{i,t}^{\text{c}(\tau)})}{\sum_j \max(0, \varphi_{j,t}^{\text{c}(\tau)}) \mathbf{1}_{\{\alpha_{j,t}=0\}}}, & \text{if } \alpha_{i,t} = 0 \\ 0, & \text{otherwise} \end{cases} \tag{11}$$

From the above formula, if the recursive value is negative, it is set to zero. Otherwise, the reward is scaled back considering only the participating sellers so that $\sum_j r_{j,t}^{\text{is}(\tau)} = 1$.

*2) Out-of-sample allocation:* Differently from the previous allocation, the out-of-sample allocation is used to reward sellers for their instantaneous performance. This is performed using a scoring function

$$sc_{i,t}^{(\tau)} = \begin{cases} 1 - \frac{\mathcal{L}(y_{t+1}, \hat{x}_{i,t+1|t}^{(\tau)})}{\sum_j \mathcal{L}(y_{t+1}, \hat{x}_{j,t+1|t}^{(\tau)}) \mathbf{1}_{\{\alpha_{j,t}=0\}}}, & \text{if } \alpha_{i,t} = 0 \\ 0, & \text{otherwise} \end{cases} \tag{12}$$

where $\mathcal{L}(y_{t+1}, \hat{x}_{i,t+1|t}^{(\tau)})$ is the loss for the $i$-th forecast. In our framework, the quantile loss is used to evaluate the forecasting accuracy. Similarly to the in-sample allocation, the score for any missing seller is set to zero, while the scores for present sellers are computed exclusively based on the subset of available forecasts.

We have that the out-of-sample reward is defined as

$$r_{i,t}^{\text{oos}(\tau)} = \frac{sc_{i,t}^{(\tau)}}{\sum_j sc_{j,t}^{(\tau)} \mathbf{1}_{\{\alpha_{j,t}=0\}}} \tag{13}$$

*3) Proprieties:* In our setting, the pay-off allocation function must satisfy key economic properties to incentivize participation, encourage truthful forecasts, and ensure consistent rewards. The straightforward properties are **budget balance** and **symmetry**. The former guarantees that the market operator redistributes all utility among the sellers. The latter, instead, ensures that two sellers who provide identical forecasts receive identical rewards. Moreover, the **zero-element** property is satisfied for missing forecasts. In this case, the corresponding seller should receive no reward. This is easy to verify, as missing forecasts are always assigned a reward of zero by construction. Another important property is **individual rationality**, which guarantees that no seller is penalized for participating (their pay-off is always non-negative). From III-C, we can see that negative rewards cannot be possible. Finally, we consider **truthfulness**, which is the most difficult property to establish. Truthfulness requires that sellers maximize their expected reward only by reporting their true forecasts. In the proposed allocation mechanism, this condition also holds. In summary, our pay-off allocation mechanism satisfies budget balance, symmetry, zero-element, individual rationality, and truthfulness. Proofs are gathered in Appendix A.

## IV. Application and Case-studies

To demonstrate the proposed market and methods, we begin by evaluating the forecasting combination algorithms and pay-off allocation across several examples, starting with two synthetic test cases and concluding with a real-world forecasting scenario. These case studies are, of course, simplified versions of what would be implemented in real-world applications.

### A. Synthetic Test Cases

To test the methods proposed in the previous sections, we generated two different environments. First, we implement a time-invariant process that shows the efficacy of the methods proposed and their convergence. Then, a time-varying process is proposed to showcase the ability of the methods to adapt to changes in the dynamics of the environment.

In both scenarios, we consider a single buyer and three sellers. Each forecaster is modeled using a Normal distribution, and the realizations $Y_t$ are generated as a combination of the sellers' distributions. We refer to the standard linear regression model using quantile regression as *QR*, and with *RQR* to the robust implementation. To evaluate performance, we performed a Monte Carlo simulation consisting of 200 independent experiments, with $T = 20000$.

*1) Time-invariant case:* The primary goal of this scenario is to verify that the proposed methods works correctly and converge to optimal weights over time. Let $\mu_{i,t}$ and $\sigma_{i,t}$ denote the mean and standard deviation of the $i$-th seller at time $t$. The seller's distribution is defined as follows

$$f_{i,t} \sim \mathcal{N}(\mu_{i,t}; \sigma_{i,t}) \tag{14}$$

where $\mu_{i,t} = C_i + \alpha \epsilon_{i,t}$, with $C_i$ constant, $\epsilon_{i,t} \sim \mathcal{N}(0, 1)$. In our setup, we consider three sellers having $C_1 = 0$, $C_2 = 1$ and $C_3 = 2$, $\alpha = 0.5$, and $\sigma_{1,t} = \sigma_{2,t} = \sigma_{3,t} = 1$.

Finally, the realizations are generated as follows

$$Y_t \sim \mathcal{N}(\mu_t; \sigma_t) \tag{15}$$

where $\mu_t = \sum_{i=1}^{n} w_i \mu_{i,t}$, $\sigma_t = \sum_{i=1}^{n} w_i \sigma_{i,t}$, and $\mathbf{w}$ is the vector of weights to be learned. In our case, the weights are set to $\mathbf{w} = [0.1, 0.6, 0.3]$.

Figure 2 illustrates the convergence behavior of both proposed methods (QR top, RQR bottom) with time horizon $k = 1$ and quantile level $\tau = 0.5$. In the RQR case, forecasters are randomly missing with probability 5%. As expected, both algorithms converge over time toward the optimal weights combination.

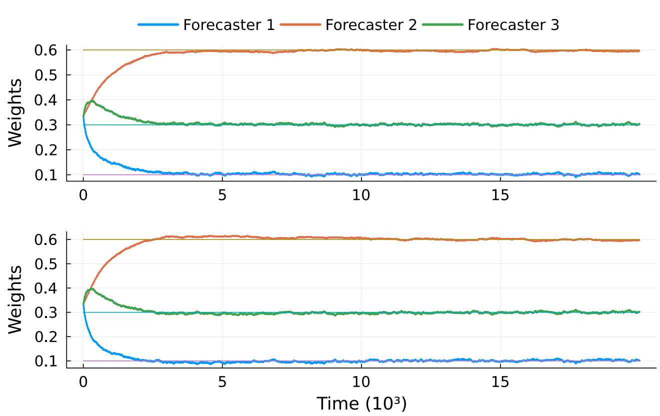


Fig. 2. Convergence of estimated weights for QR (top) and RQR (bottom) with $k = 1$ and $\tau = 0.5$.

Figure 3 shows the pay-off allocation for both methods. We consider three quantile levels, $\mathbf{m} = [0.1, 0.5, 0.9]$, with a total reward of £100 at each time $t$ equally distributed across levels. The final reward of the $i$-th seller is obtained by summing the rewards for each quantile. As the model weights converge, the reward trajectories stabilize. In the RQR case (bottom plot), the total rewards fluctuate more due to missing forecasts.

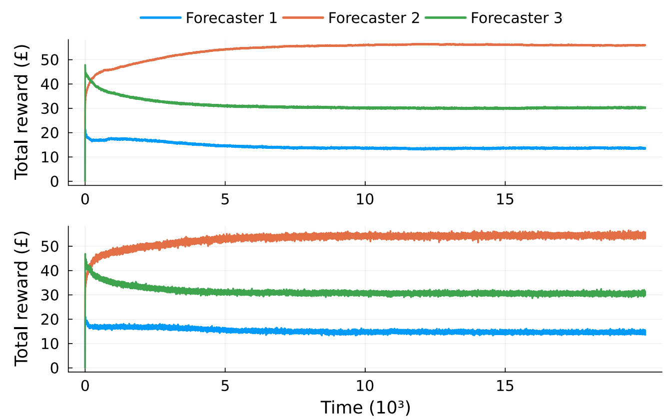


Fig. 3. pay-off allocation for QR (top) and RQR (bottom) with three quantile levels $m = 0.1, 0.5, 0.9$ and a total step reward of £100.

*2) Time-varying case:* In this setting, we simulate a dynamic environment in which the combination of weights evolves over time. The goal is to demonstrate that the proposed methods can detect these changes and adapt accordingly. This property is crucial, as real-world applications are inherently dynamic and subject to temporal variation. To do so, we introduce a periodic coefficient defined as

$$\beta_t = \tfrac{1}{2}\left(1 + \sin\left(\tfrac{2\pi t}{T}\right)\right), \quad \beta_t \in [0, 1] \tag{16}$$

Using this coefficient, we define the target weight vector as

$$\mathbf{w}_t^{\text{target}} = (1 - \beta_t)\mathbf{w}^{(1)} + \beta_t \mathbf{w}^{(2)} \tag{17}$$

where $\mathbf{w}^{(1)}$ and $\mathbf{w}^{(2)}$ are two different weight combinations. The actual weights used at time $t$ are then updated recursively

$$\mathbf{w}_t = \lambda \mathbf{w}_{t-1} + (1 - \lambda)\mathbf{w}_t^{\text{target}} \tag{18}$$

Finally, the realizations are generated according to (15).

The results shown in Fig. 4 reports the evolution of the estimated weights in a dynamic scenario. Both methods are able to adapt to the evolving environment, with the estimated weights following the underlying periodic pattern. As expected, convergence is not exact, but the overall dynamics are well captured. Also in this case the quantile level considered is $\tau = 0.5$, and for RQR (bottom) the missing rate is 5%.

### B. Performance with Varying Missingness

In this section, we compare the adaptive QR methods against two benchmark imputation strategies: mean imputation and last-value imputation. The evaluation focuses on the model's tracking performance and how it is affected by different rates of missingness over time. All results are presented for the quantile level $\tau = 0.1$.

The performance of the proposed RQR algorithm was evaluated against two benchmark models, last-impute and

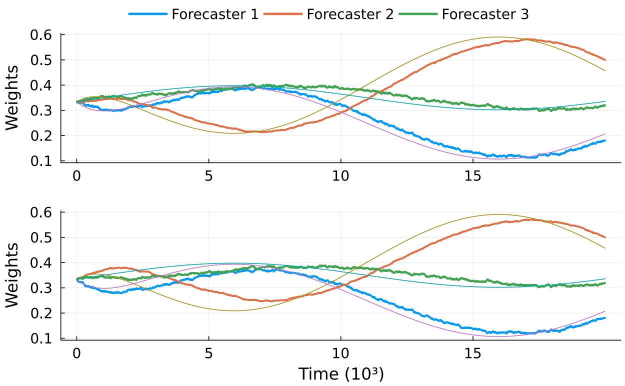


Fig. 4. Convergence of estimated weights for QR (top) and RQR (bottom) with $k = 1$ and $\tau = 0.5$ in a time-varying scenario.

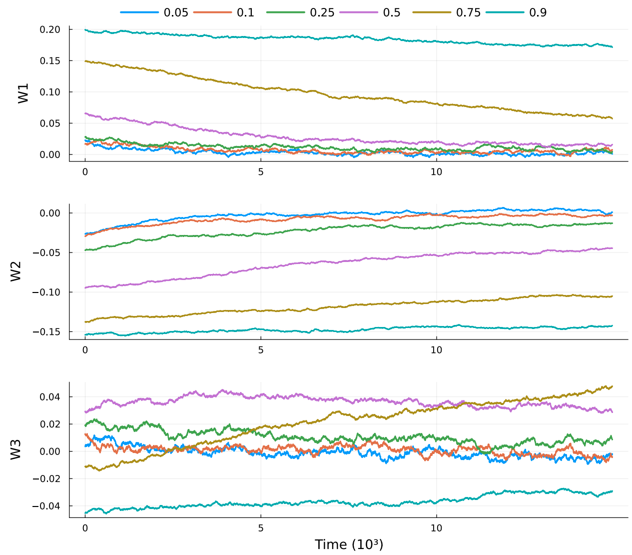


Fig. 5. Bias with varying missingness rate ranging from 5% to 90%.

mean-impute. Three different sellers were considered with weights $w_1, w_2, w_3$. The evaluation was based on two key metrics: bias and variance. The results presented in Table I, calculated excluding the first 5000 steps (to consider the results after convergence), show that RQR model has lower variance ($\sim 1.7$) compared to the imputation ones ($\sim 2.5$). This lower variance indicates that RQR's predictions are more stable and exhibit greater consistency across different samples of data. In terms of bias, RQR generally achieved superior or comparable performance. The most significant difference was observed for $w_2$ and $w_3$, where RQR's bias was much lower than that of last-impute, whilst having similar results to mean-impute.

TABLE I
BIAS AND VARIANCE FOR THE PROPOSED RQR AND THE TWO BENCHMARK MODELS (ALL VALUES ARE IN UNITS OF $10^{-3}$)

| | $w_1$ | | $w_2$ | | $w_3$ | |
|---|---|---|---|---|---|---|
| | bias | var | bias | var | bias | var |
| RQR | $4.8 \pm 3.9$ | $1.7 \pm 0.14$ | $-5 \pm 6.5$ | $1.7 \pm 0.2$ | $0.3 \pm 4$ | $1.7 \pm 0.16$ |
| Last Impute | $5 \pm 4.6$ | $2.5 \pm 0.2$ | $-35 \pm 7$ | $2.5 \pm 0.3$ | $-18 \pm 6$ | $2.5 \pm 0.2$ |
| Mean Impute | $7 \pm 5$ | $2.5 \pm 0.2$ | $-8 \pm 9.4$ | $2.5 \pm 0.4$ | $-0.3 \pm 0.6$ | $2.5 \pm 0.2$ |

Fig. 5 illustrates the bias of the RQR method under varying levels of missingness. The missingness rate ranges from 5% to 90%, with the constraint that at least one seller is always present. The results indicate that, at low missingness rates, the model exhibits an average bias close to zero. As the missingness rate increases, the bias also rises, which is expected since the model becomes less capable of learning the optimal combination and compensating through the linear correction. In terms of variance, however, the differences are minimal, indicating that the model maintains consistent performance even under high missingness conditions.

### C. Real-world Forecasting Problem

We use a wind energy forecasting case study to demonstrate the application of the proposed market framework. The scenario considers two sellers, both employing the same forecasting model but using different weather forecasts as input. The goal is to show that the market can identify the most effective combination of forecasts, while also showcasing how rewards are redistributed among participants.

Specifically, we assume that a client requests a day-ahead forecast of offshore wind energy production in Belgium. The sellers respond by submitting probabilistic forecasts in the form of quantiles with a 15-minute resolution. The actual production data used for evaluation are obtained from the open dataset of the Belgian Electricity System Operator (Elia) [18]. For the sellers' predictions, a multi-layer perceptron (MLP) model is employed. Both sellers rely on the same model architecture, with the difference lying in the input weather data. Seller $s_1$ uses forecasts from the European Centre for Medium-Range Weather Forecasts (ECMWF) [19], whereas seller $s_2$ relies on forecasts provided by the National Oceanic and Atmospheric Administration (NOAA), specifically the GFS forecasting model [20]. The datasets span from January 2025 to July 2025. The models are trained and validated using data up to June, while July is reserved for out-of-sample testing. The weather forecasts used by both seller include wind speed and wind direction at both 10 and 100 meters. In Fig. 6, we present an example of the combined forecast, illustrated for the median (quantile level $\tau = 0.5$). The figure shows that the combined forecast provides a better forecast for the actual production compared to the individual forecasts.

A quantitative comparison is reported in Table II, where the prediction losses are summarized considering varying levels of missingness. For the median quantile ($\tau = 0.5$), performance is evaluated using the MAE, while for the lower and upper quantiles ($\tau = 0.1$ and $\tau = 0.9$), the quantile loss is employed. The results show that the combined forecast consistently outperforms the individual forecasts, with the QR method delivering the strongest overall performance, as expected. Moreover, for $\tau = 0.1$ and $\tau = 0.5$, the loss associated with the RQR method increases as missingness

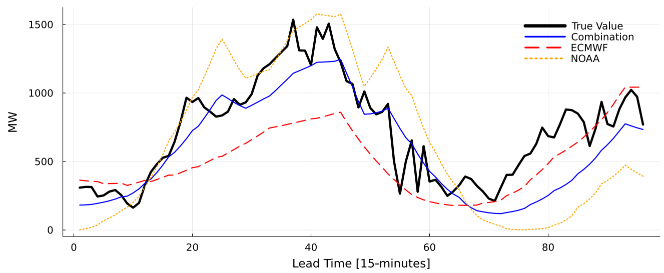


Fig. 6. Wind energy power generation reported by $s_1$ (ECMWF) and $s_2$ (NOAA), compared with the combined forecasting and the real production.

grows, whereas this effect is not observed for $\tau = 0.9$.

TABLE II
QUANTILE LOSS FOR DIFFERENT NOMINAL LEVELS AND METHODS.

| | Loss (MW) | | | | | |
|---|---|---|---|---|---|---|
| | ECMWF | NOAA | QR | RQR - 5% | RQR - 10% | RQR - 20% |
| 0.1 | 185.5 | 169.6 | 113.2 | 125.6 | 135.6 | 157.5 |
| 0.5 | 87.8 | 86.2 | 72.9 | 77.6 | 82.2 | 88.8 |
| 0.9 | 211.9 | 275.4 | 161.6 | 152.5 | 141.4 | 123.2 |

Finally, Table III reports the reward distribution for both sellers across the proposed methods for the different missingness levels. Note that RQR with a $0\%$ missing rate is equivalent to QR. The results show that the NOAA-based model provides more consistent performance over time, whereas the ECMWF-based model achieves higher accuracy on individual days. Moreover, ECMWF exhibits a slight increase in rewards, likely driven by the redistribution of rewards during periods when submissions from the second seller are unavailable. In contrast, the NOAA-based method shows a modest decrease in rewards, reflecting the impact of missing days.

TABLE III
TOTAL REWARDS (£) FOR EACH SELLER WITH VARYING MISSINGNESS

| | | Rewards (£) | | |
|---|---|---|---|---|
| | | In-sample | Out-of-sample | Total |
| 0% | ECMWF | 674.5 | 473.2 | 1147.7 |
| | NOAA | 1132.2 | 417.7 | 1549.9 |
| 5% | ECMWF | 709 | 475.5 | 1184.5 |
| | NOAA | 1109.3 | 424.5 | 1533.8 |
| 10% | ECMWF | 742.1 | 478.3 | 1220.4 |
| | NOAA | 1061.3 | 421.7 | 1483.0 |
| 20% | ECMWF | 792.1 | 465.2 | 1257.3 |
| | NOAA | 1066.0 | 428.8 | 1496.8 |

## V. Conclusions

The growing availability of data offers unprecedented opportunities to advance forecasting models and improve the integration of renewable energy generation. Yet, competitive interests and concerns over data privacy often prevent stakeholders from sharing information. There, we proposed a new prediction market platform that enables communication among stakeholders while incentivizing participation through rewards. Our market design considers many real-world application challenges and demonstrates how those can be solved. Several challenges remain open. E.g., the market can be extended to a fully dynamic setting, where the market allows a dynamic set of participants over time. And, alternative pay-off allocation mechanisms beyond Shapley value should be explored to reduce computational complexity and enhance scalability.

## References


[1] R. Bessa, A. Trindade, C. S. Silva, and V. Miranda, "Probabilistic solar power forecasting in smart grids using distributed information," *International Journal of Electrical Power & Energy Systems*, vol. 72, pp. 16–23, 2015.

[2] J. R. Andrade and R. J. Bessa, "Improving renewable energy forecasting with a grid of numerical weather predictions," *IEEE Transactions on Sustainable Energy*, vol. 8, no. 4, pp. 1571–1580, 2017.

[3] J. Dowell and P. Pinson, "Very-short-term probabilistic wind power forecasts by sparse vector autoregression," *IEEE Transactions on Smart Grid*, vol. 7, no. 2, pp. 763–770, 2016.

[4] J. W. Messner and P. Pinson, "Online adaptive lasso estimation in vector autoregressive models for high dimensional wind power forecasting," *Int. Journal of Forecasting*, vol. 35, no. 4, pp. 1485–1498, 2019.

[5] L. Cavalcante, R. J. Bessa, M. Reis, and J. Browell, "Lasso vector autoregression structures for very short-term wind power forecasting," *Wind Energy*, vol. 20, no. 4, pp. 657–675, 2017.

[6] B. Sommer, P. Pinson, J. W. Messner, and D. Obst, "Online distributed learning in wind power forecasting," *International Journal of Forecasting*, vol. 37, no. 1, pp. 205–223, 2021.

[7] C. Gonçalves, R. J. Bessa, and P. Pinson, "Privacy-preserving distributed learning for renewable energy forecasting," *IEEE Transactions on Sustainable Energy*, vol. 12, no. 3, pp. 1777–1787, 2021.

[8] C. Horn, M. Ohneberg, B. Ivens, and A. Brem, "Prediction markets: A literature review 2014," *Journal of Prediction Markets*, vol. 8, pp. 89–126, 2014.

[9] J. Witkowski, R. Freeman, J. W. Vaughan, D. M. Pennock, and A. Krause, "Incentive-compatible forecasting competitions," *Management Science*, vol. 69, no. 3, p. 1354–1374, 2022.

[10] A. A. Raja, P. Pinson, J. Kazempour, and S. Grammatico, "A market for trading forecasts: A wagering mechanism," *International Journal of Forecasting*, vol. 40, no. 1, pp. 142–159, 2024.

[11] D. Bertsimas, A. Delarue, and J. Pauphilet, "Adaptive optimization for prediction with missing data," *Machine Learning*, vol. 114, 2025, art. no. 124.

[12] A. Stratigakos and P. Andrianesis, "Learning data-driven uncertainty set partitions for robust and adaptive energy forecasting with missing data," 2025. [Online]. Available: https://arxiv.org/abs/2503.20410

[13] C. F. Grønvald, "Online regression markets with varying feature space," Master's thesis, Technical University of Denmark, 2024.

[14] L. S. Shapley *et al.*, "A value for n-person games," *Contributions Theory Games*, vol. 2, no. 28, p. 307–317, 1953.

[15] A. A. Raja and S. Grammatico, "Online coalitional games for real-time payoff distribution with applications to energy markets," *IEEE Transactions on Energy Markets, Policy and Regulation*, vol. 1, no. 2, pp. 97–106, 2023.

[16] X. Wang, R. J. Hyndman, F. Li, and Y. Kang, "Forecast combinations: An over 50-year review," *International Journal of Forecasting*, vol. 39, no. 4, pp. 1518–1547, 2023.

[17] Y. Shen, D. Xia, and W.-X. Zhou, "Online quantile regression," 2025. [Online]. Available: https://arxiv.org/abs/2402.04602

[18] Elia, "Wind power production estimation and forecast on belgian grid." [Online]. Available: https://opendata.elia.be/explore/dataset/ods031/table/?sort=datetime&refine.offshoreonshore=Offshore

[19] European Centre for Medium-Range Weather Forecasts, "Ecmwf ifs high-resolution operational forecasts," 2016. [Online]. Available: https://rda.ucar.edu/datasets/d113001/

[20] National Centers for Environmental Prediction, National Weather Service, NOAA, U.S. Department of Commerce, "Ncep gfs 0.25 degree global forecast grids historical archive," 2015. [Online]. Available: https://rda.ucar.edu/datasets/d084001/

[21] T. Falconer, J. Kazempour, and P. Pinson, "Bayesian regression markets," *Journal of Machine Learning Research*, vol. 25, no. 180, p. 1–38, 2024.

## Appendix A
## Analysis of Market Properties

We gather here the proofs for the various market properties mentioned and discussed in Section III-C3.

### A. Budget balance

For any forecasts vector $\hat{\mathbf{x}}_t^{(\tau)}$ and realization $y_{t+1}$, we have

$$\begin{aligned}\sum_i r_{i,t}^{(\tau)} &= \sum_i U_t^{(\tau)} \left[\delta r_{i,t}^{\text{is}(\tau)} + (1-\delta) r_{i,t}^{\text{oos}(\tau)}\right] \\ &= U_t^{(\tau)} \left[\delta \sum_i r_{i,t}^{\text{is}(\tau)} + (1-\delta) r_{i,t}^{\text{oos}(\tau)}\right] \\ &= U_t^{(\tau)}\end{aligned} \tag{19}$$

It means that the the sum of revenues is equal to the sum of payments (which is the definition of budget balance).

### B. Symmetry

Suppose two sellers, $s_i$ and $s_j$, always provide identical forecasts. By the Shapley symmetry property, this implies

$$\varphi_{i,t}^{\text{c}(\tau)} = \varphi_{j,t}^{\text{c}(\tau)}$$

The same holds for the out-of-sample reward, since

$$\mathcal{L}(y_{t+1}, \hat{x}_{i,t+1|t}^{(\tau)}) = \mathcal{L}(y_{t+1}, \hat{x}_{j,t+1|t}^{(\tau)})$$

Consequently, the final reward assigned to the two sellers will also be identical.

### C. Individual Rationality

From (11)-(13), we observe that the minimum possible reward for the $i$-th available seller's forecast is zero, whilst it is explicitly set to zero for unavailable ones. Hence, the total reward always satisfies $r_{i,t} \geq 0$.

### D. Truthfulness

Following the original work on robust linear regression used for our setting [11], the method can be interpreted as a linear regression model with $d + d^2$ features. Under this interpretation, we can directly apply the truthfulness proof of [21], which shows that altering a feature leads to a strictly higher loss compared to leaving it unaltered. Moreover, it has been established that for linear models, Shapley values preserve truthfulness [21]. Consequently, in our case, the in-sample reward is maximized when the forecast is reported truthfully. Since any alteration also increases the out-of-sample loss, the corresponding reward is lower than that obtained under truthful reporting. Therefore, the overall reward is maximized when the true forecast is provided.

### E. Zero Element

If the $i$-th seller does not submit a forecast at time $t$, we set $\alpha_{i,t} = 1$. From (11) and (12), it follows directly that the corresponding reward is assigned a value of zero whenever a forecast is missing.

## Appendix B
## Loss Function Derivations

We show here the derivations for the loss gradient used in III-A and III-B.

### A. Forecast Combination

The combined forecast $\hat{y}_{t+1|t}$ is defined by the linear combination given in (1). Applying the chain rule to the Quantile Loss with respect to the individual weight $w_{i,t}$ yields the following sub-gradient

$$\frac{\mathrm{d}\mathcal{L}_{i,t}(y_{t+1}, \hat{y}_{t+1|t})}{\mathrm{d}w_{i,t}} = \begin{cases} -\tau \hat{x}_{i,t+1|t} & \text{if } y_{t+1} > \hat{y}_{t+1|t} \\ (1-\tau)\hat{x}_{i,t+1|t} & \text{if } \hat{y}_{t+1|t} > y_{t+1} \end{cases} \tag{20}$$

Here, $y_{t+1}$ is the observation, $\hat{y}_{t+1|t}$ the aggregated forecast, $\tau$ the quantile level and $\hat{x}_{i,t+1|t}$ the sellers' submitted forecast.

With this we can define the weight update for the $i$-th seller

$$w_{i,t+1} = w_{i,t} - \eta \frac{\mathrm{d}\mathcal{L}_{i,t}(y_{t+1}, \hat{y}_{t+1|t})}{\mathrm{d}w_{i,t}} \tag{21}$$

### B. Adaptive Robust Linear Regression

In the context of adaptive robust linear regression, the aggregated forecast $\hat{y}_{t+1|t}$ is defined by the linear combination depending on the forecast availability vector $\boldsymbol{\alpha}_t$, as shown in (5). The derivative of the loss with respect to the individual aggregation weight remains the same as established in (20). Instead, different is the case for the linear correction matrix $\mathbf{D}^{(\tau)}$. This matrix is crucial for handling missing forecasts, with $D_{i,j;t}$ representing the correction applied to seller $i$ when seller $j$ is unavailable. The resulting sub-gradient is

$$\frac{\mathrm{d}\mathcal{L}_{i,t}(y_{t+1}, \hat{y}_{t+1|t})}{\mathrm{d}D_{i,j;t}} = \begin{cases} -\tau \hat{x}_{i,t+1|t}\alpha_{j,t} & \text{if } y_{t+1} > \hat{y}_{t+1|t} \\ (1-\tau)\hat{x}_{i,t+1|t}\alpha_{j,t} & \text{if } \hat{y}_{t+1|t} > y_{t+1} \end{cases} \tag{22}$$

where $D_{i,j;t}$ is the linear correction applied to seller $i$ when the $j$-th seller is missing, $y_{t+1}$ is the observation, $\hat{y}_{t+1|t}$ the aggregated forecast, $\tau$ the quantile level, $\hat{x}_{i,t+1|t}$ the sellers' forecast, and $\alpha_{i,t}$ the forecast availability.

We can now define the iterative update rules for both sets of parameters as

$$w_{i,t+1} = w_{i,t} - \eta \frac{\mathrm{d}\mathcal{L}_{i,t}(y_{t+1}, \hat{y}_{t+1|t})}{\mathrm{d}w_{i,t}} \tag{23}$$

$$w_{i,t+1} = w_{i,t} - \eta \frac{\mathrm{d}\mathcal{L}_{i,t}(y_{t+1}, \hat{y}_{t+1|t})}{\mathrm{d}D_{i,j;t}} \tag{24}$$

with $\eta$ learning rate.

We can see that the weights $w_{i,t}$ are updated when the $i$-th seller's forecast is present, while the correction matrix element $D_{i,j;t}$ is updated when the $j$-th input forecast is missing.