Penalising Forecast Accuracy? Data Valuation Challenges in Power Distribution Systems

Andrey Churkin[1], Pierre Pinson[1], David Pozo[2] and Janusz Bialek[3]

[1] Imperial College London, Dyson School of Design Engineering, London, United Kingdom
[2] European Commission, Joint Research Centre (JRC), Petten, Netherlands
[3] Imperial College London, Department of Electrical and Electronic Engineering, London, United Kingdom

***Abstract*—Operation of modern power systems depends on high-quality data and forecasts. As the provision of valuable forecasts shifts to independent providers, such as renewable and distributed generators, there is a growing need for forecast-sharing markets. However, existing forecast valuation mechanisms, often inspired by game theory, overlook physical network constraints and the dispersed locations of forecast providers in realistic power systems. This work shows that widely used valuation methods, such as the Shapley value, can yield counterintuitive results when applied to load forecasting in distribution networks. Specifically, due to the location of providers and voltage constraints, less accurate forecasts may appear more valuable than precise ones, while some providers may even be penalised. Moreover, value-oriented mechanisms lack incentive compatibility, as certain providers may systematically misreport load forecasts to receive higher payments. These findings expose the limitations of existing forecast valuation approaches and suggest that incorporating accuracy-based scoring may enhance value-oriented mechanisms.**

***Index Terms*—Data valuation, distribution system operator, electric vehicle aggregator, electricity market, load forecasting.**

I. Introduction

Reliable data and accurate forecasts are critical for the efficient operation of modern power systems. The provision of valuable forecasts is shifting from system operators to independent providers, such as renewable and distributed generators. This decentralisation creates the need for data marketplaces, where the value of shared data can be quantified algorithmically [1]. Numerous data valuation mechanisms have been developed in the literature, particularly within machine learning, where game-theoretic concepts such as the Shapley value are used to quantify the usefulness of datasets [2]–[5]. In power systems, similar valuation mechanisms have been applied to address a broad range of problems, including markets for renewable energy forecasting [6], [7], markets for load forecasting [8]–[11], valuation of flexibility in distribution networks [12], asset sharing in energy communities [13], and data sharing between power and traffic networks [14].

However, existing valuation mechanisms overlook physical network constraints and the dispersed locations of forecast providers in realistic power systems. These factors introduce additional complexity for forecast-sharing markets and can lead to counterintuitive outcomes. This study investigates data valuation challenges in distribution networks, with a focus on electric vehicle aggregator (EVA) load forecasting. A two-stage forecast-sharing market is implemented to examine how forecast accuracy, location, and network constraints affect the value of EVA forecasts. Several established value-oriented game-theoretic mechanisms are applied to remunerate forecast providers, including the Shapley value, the Nucleolus, and the leave-one-out (LOO) method. The results show that network constraints, such as voltage limits, have a significant impact on the value of forecasts. Binding constraints partition system operation into distinct supply regimes, where different generators must be dispatched to maintain feasibility. Transitions between these regimes are costly, as they often require expensive balancing actions. Consequently, the value of a given forecast depends on how it influences regime transitions, which is largely determined by the location of the provider.

Specifically, this work exposes the following key challenges:

- **Challenge #1**: Additional or more accurate forecasts do not necessarily improve system dispatch optimisation. Under certain network and market conditions, the cooperative game has an empty Core, indicating that no coalitionally rational allocation exists.
- **Challenge #2**: Identical forecasts submitted at different network locations can receive unequal payments. Because of network constraints, less accurate forecasts may appear more valuable than precise ones, and even perfect (error-free) forecasts may be penalised.
- **Challenge #3**: Value-oriented mechanisms are not incentive compatible. Certain providers can strategically misreport their forecasts to receive higher payments.

These findings highlight the limitations of existing forecast valuation approaches and raise concerns about the fairness and incentive compatibility of forecast-sharing markets under physical network constraints. To address these limitations, this work introduces the *accuracy-adjusted Shapley value*, which combines value-oriented and accuracy-based mechanisms.

The forecast valuation model, case study data, and simulation results are published as an open-source tool, `ForecastVal-E` [15]. The remainder of the paper is organised as follows. Section II presents the modelling framework

Submitted to the 24th Power Systems Computation Conference (PSCC 2026).
Corresponding author: Andrey Churkin https://andreychurkin.ru/

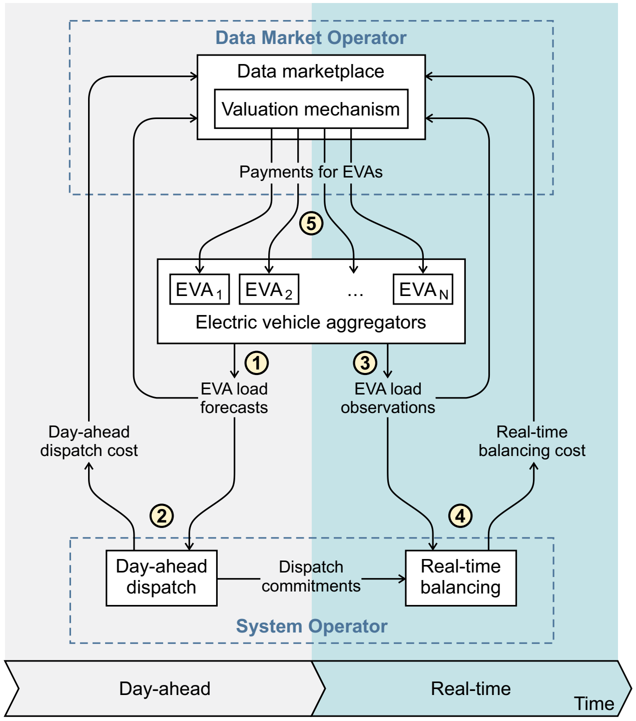

Fig. 1. Forecast valuation framework for a two-settlement electricity market. The circled numbers indicate the chronological order of valuation steps.

consisting of day-ahead and real-time dispatch optimisation models. Section III introduces game-theoretic and accuracy-based mechanisms for forecast valuation. In Section IV, extensive simulations are performed for a 33-bus distribution system with three EVAs participating in a forecast-sharing market. Section V discusses properties and scalability of the proposed framework. Finally, Section VI concludes the paper.

II. Modelling Framework: Load Forecasting in a Two-Settlement Electricity Market

This section introduces dispatch optimisation models incorporating EVA load forecasts in the context of a two-settlement electricity market [9]. The proposed forecast valuation framework consists of two time stages, corresponding to the day-ahead dispatch and real-time operation. An overview of the framework is presented in Fig. 1, where the circled numbers indicate the steps required to perform forecast valuation. First, ①, in the day-ahead electricity market, each EVA submits its forecast of expected EV power consumption. This information is used by the system operator to determine the optimal day-ahead dispatch ②, including dispatch costs and commitments for the balancing market. These results, together with the EVA forecasts, are provided to the data market operator for valuation. Then, during real-time operation, actual EV loads are observed ③, and the system operator estimates the cost of balancing actions resulting from forecast errors ④. Finally, this information is submitted to the data marketplace, where the valuation mechanism determines payments for EVAs ⑤.

To obtain optimal dispatch solutions for both day-ahead and balancing stages, an optimal power flow (OPF) model is required. In this work, the DistFlow power flow formulation is implemented, which captures the operation of radial distribution networks, including active and reactive power flows, voltage limits, and line capacity constraints [16]. This formulation yields a continuous nonlinear programming (NLP) problem, namely a quadratically constrained program (QCP). The constraints for each time stage are defined in (1a)-(1i). Here, $\mathcal{N}$, $\mathcal{G}$, $\mathcal{L}$ denote the sets of buses, generators, and lines, respectively, while $\mathcal{A}$ is the set of forecast providers. In (1a), the total power of each generator is expressed as the sum of its initial setpoint and upward or downward regulations. Constraint (1b) imposes generator capacity limits. Equations (1c)-(1d) define power flows between buses n and m in a radial grid, where $\Re(Z_{(n,m)})$ and $\Im(Z_{(n,m)})$ denote the real and imaginary parts of the branch impedance, and $p^t_{m,d}$ and $q^t_{m,d}$ represent nodal loads. The voltage relation between buses n and m is defined by (1e). Equation (1f) is the relation between branch power flows and currents. Line capacity constraint is imposed in (1g), where $\overline{S}$ is apparent power rating (thermal limit). Voltage limits are set in (1h). Equation (1i) specifies the stage-dependent EV load, equal to the forecast in the day-ahead stage and the observation in the real-time stage.

Based on the DistFlow formulation, the day-ahead dispatch model is expressed as Model 1 (2a)-(2c). This model corresponds to a day-ahead electricity market clearing problem with the objective of minimising the total cost of power production across all generators. Constraint (2b) sets the initial dispatch of all generators to zero, reflecting the first optimisation stage with no prior commitments. Constraint (2c) specifies the EV load forecast as the improved forecast provided by an EVA i, or the baseline forecast (typically conservative) available to the system operator if EVA i does not share its forecast. The parameter χ is a binary indicator of forecast availability.

The real-time dispatch problem is expressed as Model 2 (3a)-(3b), with the objective of minimising the total cost of balancing actions from generators. In (3b), the initial dispatch of generators is fixed to the optimal commitments obtained from Model 1, indicated by the star symbol $\star$. It follows that the solution to the real-time balancing model depends on the outcome of the day-ahead model: the volume and cost of balancing actions increase as the real-time dispatch deviates further from the day-ahead dispatch due to forecast errors.

The total system cost is equal to the sum of dispatch costs in the day-ahead and real-time stages, defined as

$$\mathcal{C}(\boldsymbol{\chi}) = \mathrm{val}\big(\textbf{Model 1} \,\big|\, \hat{\mathbf{L}}(\boldsymbol{\chi})\big) + \mathrm{val}\big(\textbf{Model 2}\big) \tag{4}$$

To estimate the contributions of forecasts, this work uses the *avoided cost* as a valuation metric. It is calculated as the difference between dispatch solutions obtained without EVA forecasts and solutions obtained when all EVAs provide load forecasts to the system operator, as

$$\Delta\mathcal{C}(\boldsymbol{\chi}) = \mathcal{C}(\mathbf{0}) - \mathcal{C}(\boldsymbol{\chi}) \tag{5}$$

This metric captures the monetary value of improved load forecasting compared to conservative predictions available to

DistFlow-based formulation for distribution network dispatch optimisation [continuous NLP, QCP]

Variables: (for day-ahead and real-time stages $t \in \mathcal{T} = \{\mathrm{DA}, \mathrm{RT}\}$)

$p_{n,g}^{t,\uparrow}, p_{n,g}^{t,\downarrow}, q_{n,g}^{t,\uparrow}, q_{n,g}^{t,\downarrow}$		active and reactive power regulations by generators	$n \in \mathcal{N},\ g \in \mathcal{G}$
$p_{n,g}^{t}, q_{n,g}^{t}$		total active and reactive power outputs of generators	$n \in \mathcal{N},\ g \in \mathcal{G}$
$p_{(n,m)}^{t}, q_{(n,m)}^{t}$		active and reactive power flows	$(n,m) \in \mathcal{L}$
V_n^t	$({V_n^t}^2 = w_n^t)$	bus voltages	$n \in \mathcal{N}$
$I_{(n,m)}^t$	$({I_{(n,m)}^t}^2 = l_{(n,m)}^t)$	branch currents	$(n,m) \in \mathcal{L}$

Constraints: (for day-ahead and real-time stages $t \in \mathcal{T} = \{\mathrm{DA}, \mathrm{RT}\}$)

$$p_{n,g}^t = p_{n,g}^{t,\circ} + p_{n,g}^{t,\uparrow} - p_{n,g}^{t,\downarrow}, \qquad q_{n,g}^t = q_{n,g}^{t,\circ} + q_{n,g}^{t,\uparrow} - q_{n,g}^{t,\downarrow} \qquad \forall n \in \mathcal{N},\ g \in \mathcal{G} \tag{1a}$$

$$\underline{p}_{n,g} \le p_{n,g}^t \le \overline{p}_{n,g}, \qquad \underline{q}_{n,g} \le q_{n,g}^t \le \overline{q}_{n,g} \qquad \forall n \in \mathcal{N},\ g \in \mathcal{G} \tag{1b}$$

$$p_{(n,m)}^t = \sum_{d \in \mathcal{D}} p_{m,d}^t - \sum_{g \in \mathcal{G}} p_{m,g}^t + \Re(Z_{(n,m)}) l_{(n,m)}^t + \sum_{(m,k) \in \mathcal{L}} p_{(m,k)}^t + \sum_{i \in \mathcal{A}} L_{n,i}^t \qquad \forall (n,m) \in \mathcal{L} \tag{1c}$$

$$q_{(n,m)}^t = \sum_{d \in \mathcal{D}} q_{m,d}^t - \sum_{g \in \mathcal{G}} q_{m,g}^t + \Im(Z_{(n,m)}) l_{(n,m)}^t + \sum_{(m,k) \in \mathcal{L}} q_{(m,k)}^t \qquad \forall (n,m) \in \mathcal{L} \tag{1d}$$

$$w_m^t = w_n^t + \left| Z_{(n,m)} \right|^2 l_{(n,m)}^t - 2\big(\Re(Z_{(n,m)}) p_{(n,m)}^t + \Im(Z_{(n,m)}) q_{(n,m)}^t\big) \qquad \forall (n,m) \in \mathcal{L} \tag{1e}$$

$${p_{(n,m)}^t}^2 + {q_{(n,m)}^t}^2 = l_{(n,m)}^t w_n^t \qquad \forall (n,m) \in \mathcal{L} \tag{1f}$$

$${p_{(n,m)}^t}^2 + {q_{(n,m)}^t}^2 \le {\overline{S}_{(n,m)}}^2 \qquad \forall (n,m) \in \mathcal{L} \tag{1g}$$

$${\underline{V}_n}^2 \le w_n^t \le {\overline{V}_n}^2 \qquad \forall n \in \mathcal{N} \tag{1h}$$

$$L_{n,i}^t = \begin{cases} \hat{L}_{n,i}, & t = \mathrm{DA} \quad \text{(EVA load forecast)} \\ L_{n,i}, & t = \mathrm{RT} \quad \text{(EVA load observation)} \end{cases} \qquad \forall n \in \mathcal{N},\ i \in \mathcal{A} \tag{1i}$$

MODEL 1 Day-ahead electricity market [NLP, QCP]

$$\min \sum_{n \in \mathcal{N}} \sum_{g \in \mathcal{G}} C_g^{\mathrm{DA}}\big(p_{n,g}^{\mathrm{DA}}, q_{n,g}^{\mathrm{DA}}\big) \tag{2a}$$

s.t.

Constraints (1a)-(1i) for $t = \mathrm{DA}$

$$p_{n,g}^{\mathrm{DA},\circ} = 0, \quad q_{n,g}^{\mathrm{DA},\circ} = 0 \qquad \forall n \in \mathcal{N},\ g \in \mathcal{G} \tag{2b}$$

$$\hat{L}_{n,i} = \begin{cases} \hat{L}_{n,i}^{\text{improved}}, & \chi_i = 1 \\ \hat{L}_{n,i}^{\text{baseline}}, & \chi_i = 0 \end{cases} \qquad \forall n \in \mathcal{N},\ i \in \mathcal{A} \tag{2c}$$

MODEL 2 Real-time balancing dispatch [NLP, QCP]

$$\min \sum_{n \in \mathcal{N}} \sum_{g \in \mathcal{G}} C_g^{\mathrm{RT}}\big(p_{n,g}^{\mathrm{RT},\uparrow}, p_{n,g}^{\mathrm{RT},\downarrow}, q_{n,g}^{\mathrm{RT},\uparrow}, q_{n,g}^{\mathrm{RT},\downarrow}\big) \tag{3a}$$

s.t.

Constraints (1a)-(1i) for $t = \mathrm{RT}$

$\big(p^{\mathrm{DA},\star},\ q^{\mathrm{DA},\star}\big) \in \mathrm{Opt}\big(\textbf{MODEL 1}\ \text{(2a)-(2c)}\big)$

$$p_{n,g}^{\mathrm{RT},\circ} = p_{n,g}^{\mathrm{DA},\star}, \quad q_{n,g}^{\mathrm{RT},\circ} = q_{n,g}^{\mathrm{DA},\star} \qquad \forall n \in \mathcal{N},\ g \in \mathcal{G} \tag{3b}$$

the system operator. The next section introduces the mechanisms for allocating these savings as payments to EVAs.

III. Forecast Valuation Mechanisms

Building on the dispatch optimisation models defined earlier, this section introduces the mechanisms for quantifying the value of EVA load forecasts and determining payments to EVAs. Two main approaches are considered: (i) game-theoretic mechanisms designed to allocate the total avoided cost among EVAs according to their contributions to cost reductions, and (ii) an accuracy-based mechanism that links payments to forecast accuracy.

A. Game-Theoretic Forecast Valuation

Widely used approaches to data valuation originate from cooperative game theory, in which interactions among players are modelled as transferable utility (TU) games. In this context, the value of a player's dataset is determined by its contributions to the collective outcome of all players. Formally, a cooperative game with TU is defined as a pair (N, v), where:

- N is a finite set of players. A subset of N is called a coalition, denoted as S. The largest possible coalition containing all players is called the grand coalition. The collection of all coalitions is denoted by 2^N.
- $v : 2^N \to \mathbb{R}$ is a characteristic function associating each coalition S with a real number $v(S)$, the value of coalition S.

The connection of the cooperative game formulation (N, v) to the dispatch optimisation models is straightforward. The set of players is given by the EVAs participating in the forecast-sharing market, $N = \mathcal{A}$. Coalitions correspond to combinations of load forecasts, where the value of each coalition S is the cost reduction achieved through forecast sharing, i.e.

$$v(S) = \Delta\mathcal{C}\big(\boldsymbol{\chi}(S)\big), \quad v(\varnothing) = 0 \tag{6}$$

The empty coalition with no players, denoted by $\varnothing$, has zero value since no EVA forecasts are provided. The indicator vector of forecast availability for coalition S is defined as

$$\big[\boldsymbol{\chi}(S)\big]_i = \mathbb{1}\{\, i \in S \,\}, \quad \boldsymbol{\chi}(\varnothing) = \mathbf{0},\ \boldsymbol{\chi}(N) = \mathbf{1} \tag{7}$$

By characterising all coalitions using the optimal solutions to the dispatch optimisation models, the cooperative game formulation captures the combinatorial nature of the forecast valuation problem and provides a framework for estimating payments to forecast providers. Below, the game-theoretic solution concepts are introduced to allocate the avoided cost of the grand coalition, $v(N) = \Delta\mathcal{C}\big(\boldsymbol{\chi}(N)\big)$, among players.

Many solution concepts revolve around measuring the marginal contribution of player i to coalition S, which is defined as the difference between the value of the coalition with and without player i:

$$\mathrm{MC}_i(S) = v(S \cup \{i\}) - v(S) \quad \forall i \in S \quad \forall S \subseteq N \tag{8}$$

The Shapley value [17] is defined for each player i as the weighted average of its marginal contributions across all possible coalitions, where $|N|$ denotes the total number of players, and $|S|$ is the number of players in coalition S:

$$\mathrm{Sh}_i(v) = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(|N| - |S| - 1)!}{|N|!} \mathrm{MC}(S)_i \tag{9}$$

The Shapley value is widely recognised as a useful concept due to its desirable properties, such as *efficiency*, *symmetry*, and the *null player* property [18]. However, its applicability is generally limited to small games with a few players, as the number of possible coalitions, 2^N, grows exponentially. In practice, applications typically involve games with 10 or fewer players, which requires evaluating up to 1024 coalitions. Nevertheless, recent advances in approximation techniques have significantly improved scalability, with some studies reporting solvable cases involving up to 200 players [13]. The scalability of the forecast valuation framework is further discussed in Section V.

Another important solution concept is the Nucleolus. Instead of averaging the marginal contributions of players, it considers the excess of coalitions, defined as the difference between the value of coalition S and the sum of allocations to the players in this coalition, i.e.

$$e_S(x) = v(S) - \sum_{i \in S} x_i \tag{10}$$

Specifically, the Nucleolus is the allocation that minimises the maximum excess across all coalitions lexicographically, as

$$\mathrm{Nuc}(v) = \operatorname*{arg\,min}_{x \in \mathbb{R}^N,\, \sum_i x_i = v(N)} \mathrm{lex}\big(e(x)\big) \tag{11}$$

This allocation balances dissatisfaction with payments across coalitions, while satisfying the efficiency constraint $\sum_i x_i = v(N)$. Finding the Nucleolus involves solving a series of linear programming problems, for which useful guidance is given in [19].

Note that using the Shapley value or the Nucleolus requires evaluating all possible coalitions of players, 2^N, which is computationally challenging for large games. A simpler approach is the leave-one-out (LOO) method, which considers only the contributions of players to the grand coalition, defined as

$$\mathrm{LOO}_i(v) = \mathrm{MC}_i(N) + \frac{v(N) - \sum_{j \in N} \mathrm{MC}_j(N)}{|N|} \tag{12}$$

This intuitive method measures the impact of removing one forecast at a time. In the game-theoretic literature, the solution obtained by the LOO method is also known as the equal allocation of non-separable costs [20], [21].

To analyse the applicability of valuation mechanisms, the concept of the Core is often used. The Core is the set of all coalitionally rational allocations, where no coalition receives less than its value, defined as

$$\begin{aligned} \mathrm{Core}(v) = \Big\{\, x \in \mathbb{R}^{|N|} \;\Big|\; & \sum_{i \in N} x_i = v(N), \\ & \sum_{i \in S} x_i \geq v(S) \quad \forall S \subseteq N \,\Big\} \end{aligned} \tag{13}$$

Due to this property, allocation solutions within the Core are considered stable, since every player and every coalition is guaranteed at least its value and therefore has no incentive to leave the grand coalition.

The above game-theoretic concepts are used to allocate the value of forecasts among providers. However, these value-oriented approaches do not account for the accuracy of forecasts, i.e., forecast errors. To address this limitation, the next subsection introduces an accuracy-based valuation approach.

B. Accuracy-Based Forecast Valuation

The accuracy of EVA load forecasts is characterised by the forecast error, defined as the difference between the forecast and the observation:

$$\varepsilon_i = \hat{L}_i - L_i \tag{14}$$

In this notation, a positive error indicates load overestimation, while a negative error corresponds to underestimation. Based on the forecast error, each forecast is assigned an accuracy score, $\eta_i \in [0, 1]$, according to the following rule:

$$\eta_i = 1 - \left(\frac{|\varepsilon_i|}{\tau}\right)^{\rho} \tag{15}$$

Here, $\tau > 0$ is a scaling parameter that sets the error threshold (in MW) at which the accuracy score becomes zero. Parameter $\rho > 0$ controls the shape (smoothness) of the error penalisation function, that is, how fast the score decreases for less accurate forecasts. Under this rule, more accurate forecasts receive higher scores, which can then be used to determine payments to forecast providers. For example, payments can be allocated proportionally to the obtained accuracy scores. A comprehensive analysis of the accuracy-scoring and error-penalisation functions is left for future research.

In this work, the accuracy scores are used to propose an allocation mechanism called the *accuracy-adjusted Shapley value*. In this mechanism, parameter $\alpha \in [0,1]$ determines the share of the grand coalition value that is allocated based on forecast accuracy, while the remaining part $(1-\alpha)$ is distributed according to the Shapley value, i.e.

$$\mathrm{Sh}_i^{\mathrm{AA}}(v;\alpha) = (1-\alpha)\,\mathrm{Sh}_i(v) + \alpha\,\frac{\eta_i}{\sum_{j\in N}\eta_j}\,v(N), \qquad \alpha\in[0,1] \tag{16}$$

For $\alpha = 1$, valuation is performed solely on the basis of accuracy, while for $\alpha = 0$, the mechanism reduces to the standard Shapley value. Note that the accuracy-adjusted Shapley value does not preserve all properties of the standard Shapley value. Axiomatic properties of this mechanism are further discussed in Section V.

IV. Results and Discussion

This section presents the case study and simulation results. A distribution network with several geographically dispersed EVAs is considered, for which the day-ahead and real-time dispatch optimisation models are solved. Then, to estimate the value of EVA load forecasts, multiple mechanisms are implemented, including the Shapley value, the Nucleolus, the LOO method, and the accuracy-adjusted Shapley value. Finally, the challenges of forecast valuation under network constraints are discussed, raising concerns over the applicability of existing game-theoretic value-oriented mechanisms.

The forecast valuation framework is implemented in Julia (v1.10.4), with dispatch optimisation models formulated in JuMP (v1.22.2) and solved using Gurobi solver (v11.0.2). All models, code, and case studies are available at [15].

A. Case Study and Modelling Assumptions

To demonstrate the challenges of data valuation in distribution networks, the canonical IEEE 33-bus test system is selected, which represents a 12.66 kV radial distribution network with a total demand of 3.7 MW and 2.3 MVAr [16]. The system is visualised in Fig. 2 as a graph where the size of the nodes is proportional to the nodal apparent power demand. Three EVAs are placed in the network, connected to buses 22, 32, and 14. The choice of only three EVAs in this work makes it possible to formulate and solve the forecast-sharing TU cooperative game without computational difficulties. The small number of forecast providers also facilitates the visualisation and interpretation of the resulting allocation solutions (payments to providers). For illustrative purposes, all EVAs are assumed to be identical, with the same load observation at the real-time stage. Nevertheless, due to the differences in their locations, these EVAs are exposed to different operational constraints, which affects the value of their forecasts. For example, as highlighted in red in Fig. 2, two feeders operate at voltages below 0.95 p.u. due to voltage drops along

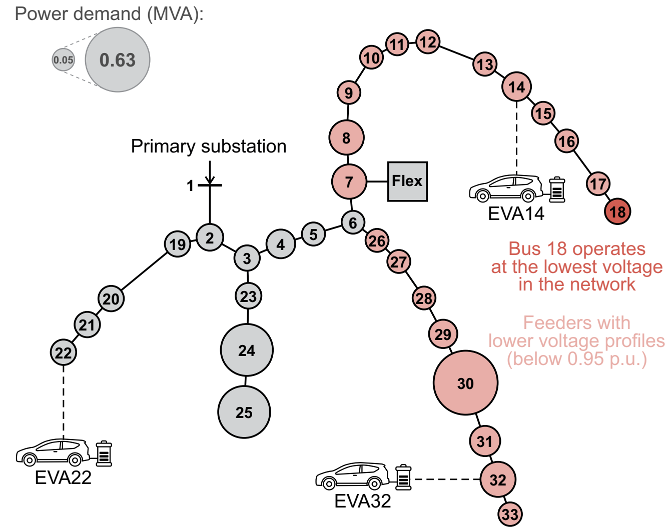

Fig. 2. Case study: 33-bus radial distribution network with 3 EVAs. The size of the nodes corresponds to the nodal apparent power demand. Nodes in red indicate feeders that operate under lower voltage profiles (below 0.95 p.u.). Bus 18, shown in dark red, operates at the lowest voltage in the network and may lead to binding voltage constraints during dispatch.

the lines. This imposes additional voltage constraints on the operation of EVA32 and EVA14, where power consumption may be curtailed once the minimum voltage limit of 0.90 p.u. is reached. To avoid voltage violations and ensure feasible network operation, a flexible generator is connected to bus 7. This generator provides additional support when demand is high, supplying part of the load and alleviating voltage drops.

The cost for day-ahead generation, C_g^{DA}, and the costs of upward and downward regulations in the balancing market, $C_g^{\mathrm{RT},\uparrow}$ and $C_g^{\mathrm{RT},\downarrow}$, are specified as follows. Power supplied from the main grid (primary substation) is priced at 100 EUR/MWh in the day-ahead market, while balancing costs are 300 EUR/MWh for upward regulation and 200 EUR/MWh for downward regulation. The flexible generator at bus 7 is assumed to be more expensive to operate, with costs of 200 EUR/MWh in the day-ahead market and 400 EUR/MWh for both upward and downward balancing. Note that reactive power is not explicitly priced in this study and is therefore excluded from the dispatch objective functions. Although a nonlinear OPF formulation with both active and reactive power flows is employed, reactive power consumption by EVs is neglected. The flexible generator is also restricted to active power provision, without reactive power capability.

This case is constructed in a way that voltage limits can become binding and alter the optimal dispatch, requiring the activation of expensive flexibility. These voltage-driven redispatch actions create divergent valuations of load forecasts. Nevertheless, the same valuation framework applies to cases with other binding constraints, e.g., line congestion.

B. Single-Instance Valuation of Improved Forecasts

To illustrate the value of EVA load forecasts, a simple instance of the forecast-sharing market is analysed first. Each

TABLE I
COALITIONAL ANALYSIS FOR A SINGLE FORECAST-SHARING MARKET INSTANCE

Coalition S (combination of forecasts)	Day-ahead cost, EUR/h	Balancing cost, EUR/h	Total cost, EUR/h	Value of coalition $v(S)$ as its avoided cost, EUR/h	Marginal contributions $\mathrm{MC}(S)$, EUR/h		
					by EVA22	by EVA32	by EVA14
{EVA22}	434.92	143.04	577.96	20.37	20.37	–	–
{EVA22, EVA32}	446.42	145.21	591.63	6.70	-16.28	-13.67	–
{EVA22, EVA32, EVA14}	467.75	102.26	570.01	28.32	-15.99	17.62	21.62
{EVA22, EVA14}	447.33	140.30	587.63	10.71	-12.48	–	-9.66
{EVA32}	436.23	139.13	575.36	22.98	–	22.98	–
{EVA32, EVA14}	457.22	96.80	554.02	44.31	–	21.13	21.34
{EVA14}	436.79	138.35	575.15	23.19	–	–	23.19
∅	424.74	173.59	598.33	0.00	–	–	–

EVA is assumed to provide an improved forecast with lower error than the baseline forecast available to the system operator. Specifically, the improved forecast, $\hat{L}_{n,i}^{\text{improved}}$, is 0.20 MW, while the load observation in real-time, $L_{n,i}$, is 0.25 MW. In the absence of forecast sharing, the system operator relies on its own baseline prediction, $\hat{L}_{n,i}^{\text{baseline}}$, which is assumed to be 0.10 MW per EVA. These assumptions define the input data for the cooperative game, where the system costs are evaluated for all possible coalitions of EVA load forecasts. The resulting costs and marginal contributions are reported in Table I.

The presented coalitional analysis reveals the challenges of data valuation in distribution systems. Specifically, the grand coalition may not yield the highest value for the system. Coalition {EVA32, EVA14} achieves an avoided cost of 44.31 EUR/h, whereas the grand coalition {EVA22, EVA32, EVA14} achieves only 28.32 EUR/h. As a result, some marginal contributions are negative and the Core of the game is empty, which means that no allocation satisfies the coalitional rationality condition (13). Such cases are problematic for market design, as the provision of improved forecasts by all EVAs does not necessarily guarantee the maximum system benefit.

This outcome arises from the impact of network constraints. EVA14 is located near bus 18, which operates at the lowest voltage in the network and where the minimum voltage limit becomes binding. In case of this voltage limit violation, the flexible generator must be activated to support the network. When only EVAs 32 and 14 provide improved forecasts, the day-ahead dispatch anticipates this violation and schedules power support from the flexible generator. However, when EVA22 also provides an improved forecast, the day-ahead dispatch allocates more power supply from the primary substation. This surplus must later be rebalanced at a higher cost once the flexible generator is activated, which reduces the overall value of the grand coalition.

Different allocation mechanisms are compared for this market instance, including the Shapley value, the Nucleolus, the LOO method, and the accuracy-adjusted Shapley value. Table II reports the corresponding allocations of the total avoided cost, expressed as payments to the EVAs. Two main observations can be made from these allocation solutions. First, all mechanisms assign very different values to the EVAs, even though the three forecasts shared are identical in accuracy and capacity. This confirms that the forecast value strongly depends on the provider's location and network constraints. Second, some allocations are negative, which implies that certain EVAs would be penalised for providing improved forecasts. This raises concerns for the market design, as it appears counterintuitive or unfair that sharing accurate forecasts may result in a financial loss for the provider.

TABLE II
COMPARISON OF FORECAST VALUATION MECHANISMS
FOR A SINGLE FORECAST-SHARING MARKET INSTANCE

Allocation mechanism	Net payments, EUR/h		
	EVA22	EVA32	EVA14
Shapley value	-3.33 (-11.77%)	14.77 (52.16%)	16.88 (59.61%)
Nucleolus	2.19 (7.73%)	12.96 (45.76%)	13.17 (46.51%)
Leave-one-out method	-19.48 (-68.77%)	21.46 (75.77%)	26.34 (93.00%)
Accuracy-adjusted Shapley value, $\alpha = 0.3$	0.50 (1.76%)	13.17 (46.51%)	14.65 (51.72%)
Accuracy-adjusted Shapley value, $\alpha = 1.0$	9.44 (33.33%)	9.44 (33.33%)	9.44 (33.33%)

In this market instance, the Nucleolus allocates positive payments to all players and does not penalise EVA22. However, this outcome is specific to this case. As shown in the subsequent simulations, the Nucleolus does not guarantee non-negative allocations and, similar to the Shapley value and the LOO method, can penalise certain providers. The accuracy-adjusted Shapley value also does not penalise EVA22. Since all EVA forecasts have identical errors, the accuracy-based component adds a positive adjustment to EVA22, making its total net payment positive. Nevertheless, ensuring non-negativity in general requires a careful selection of the weighting parameter α and the accuracy scoring rule. When $\alpha = 1.0$, allocations are determined solely by forecast accuracy, and, since all forecasts are identical in this case, each EVA receives an equal payment corresponding to 33.33% of the total avoided cost.

This analysis highlights the crucial role of network constraints in forecast valuation. Binding constraints partition the system operation into distinct supply regimes, where different generators must be dispatched to maintain feasibility. Transitions between these regimes are costly, as they often require expensive balancing actions. Consequently, the value of given

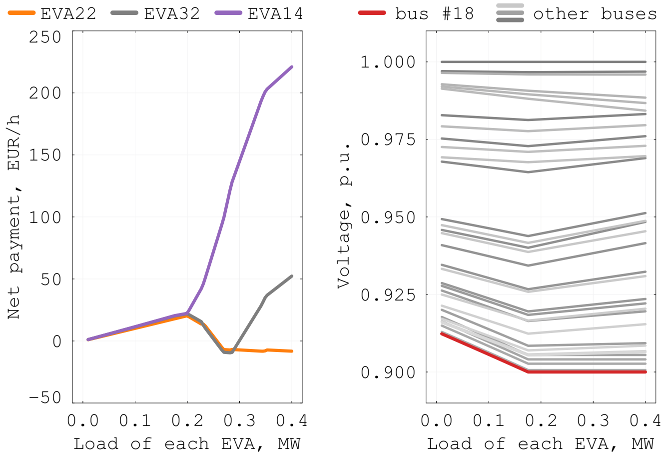

Fig. 3. Iterative analysis of error-free EVA load forecasts for different loads: (left) payments to EVA for forecast sharing calculated using the Shapley value; (right) nodal voltages at real-time operation.

forecasts depends on how they influence regime transitions. Even an improved forecast can lead to a suboptimal system dispatch and have a negative marginal contribution.

C. The Value of Perfect Forecasts

The previous analysis showed that improved forecasts still require balancing actions arising from forecast errors, which may lead to uneven values of forecasts and negative payments to providers. To further test the forecast-sharing market, this subsection examines the value of perfect error-free forecasts, where each EVA is assumed to know and share its true load observation. Thus, the system operator acquires perfect information and takes the optimal dispatch decisions.

The value of perfect forecasts is estimated iteratively using the Shapley value for 100 EVA load observations in the range [0.01, 0.4] MW, as shown in Fig. 3. This analysis illustrates how allocations change with different load levels under perfect information. At low EVA loads, the system is unconstrained and all voltages remain within limits. In such cases, EVAs make similar contributions by sharing perfect forecasts, and the Shapley value allocates nearly identical payments to all providers. However, at higher EVA loads, voltage constraints become binding and the value of forecasts diverges: EVA14 receives much higher payments, while EVA22 is penalised. This effect arises from the need to dispatch the flexible generator, which introduces additional balancing costs for certain coalitions. As a result, the forecast from EVA22, although useful for the grand coalition, makes negative contributions in some coalitions. The Shapley value reflects this effect as a negative allocation. Thus, even with perfect forecasts, network constraints can lead to unequal or negative payments.

Note that only the Shapley value allocations are visualised in Fig. 3. The Nucleolus and the LOO method produce very similar allocation patterns for perfect forecasts, while the accuracy-based valuation splits the avoided cost equally among the EVAs, since all forecasts are error-free.

D. Monte Carlo Analysis with Forecast Errors

Previous subsections considered idealised cases where all three EVAs provide identical forecasts. To investigate a realistic forecast-sharing market under uncertainty, this subsection employs Monte Carlo analysis incorporating forecast errors. A total of 1000 scenarios are generated, where the load forecast of each EVA deviates from the observation by ±25%. Specifically, the load observations, $L_{n,i}$, are assumed to be 0.20 MW per EVA, load forecasts, $\hat{L}_{n,i}^{\text{improved}}$, are uniformly distributed within the range [0.15, 0.25] MW, and the baseline forecast, $\hat{L}_{n,i}^{\text{baseline}}$, is fixed at 0.10 MW per EVA. Under these assumptions, all generated market instances result in a positive total avoided cost, $v(N)$. That is, EVA forecasts collectively reduce the total system costs compared to the baseline forecasts. Scenarios with inaccurate or malicious forecasts that increase system costs are not considered. The resulting allocations are visualised in Fig. 4 for three allocation mechanisms: the Shapley value, the Nucleolus, and the LOO method. The figure illustrates the relationship between forecast errors and EVA payments, with negative values corresponding to penalties. The density plots on the right show the distributions of payments, from which the expected values can be derived. Note that since each EVA has a load observation of exactly 0.20 MW across all simulations, the resulting payments are presented without rescaling by power consumption.

Several observations can be made from these allocation solutions. First, the shapes of the allocation-accuracy relationships differ markedly across the three EVAs. EVA14, located in the most constrained part of the network, has a distribution of payments distinct from those of EVA22 and EVA32. Also, EVA14 receives higher payments on average. Second, in contrast to EVA14, both EVA22 and EVA32 may receive penalties even for accurate forecasts with low errors. This confirms the earlier finding that the location of forecast providers has a strong influence on forecast valuation. Third, the forecast-sharing market is not incentive compatible. For instance, the payment profile of EVA14 indicates a clear incentive to overestimate its load, that is, to report higher values. By doing so, EVA14 could systematically secure higher payments and avoid the region of load underestimation, where its payments are close to zero. Taken together, these observations indicate that the practical application of forecast-sharing markets with value-oriented allocation mechanisms is at risk in the presence of physical network constraints.

To address the issues of negative allocations and incentive incompatibility, the accuracy-adjusted Shapley value (16) is applied to the same 1000 Monte Carlo scenarios. The scoring rule (15) with $\tau = 0.05$ MW and $\rho = 2.0$ is used to calculate the accuracy-based payments for EVA forecasts. The results are presented in Fig. 5 for different values of the parameter α, which determines the share of payments calculated based on forecast accuracy. When α is 0.3, payments are still largely shaped by the Shapley value allocations, reflecting the impact of network constraints and supply regimes. In contrast, when α is 1.0, payments are determined solely by forecast accuracy,

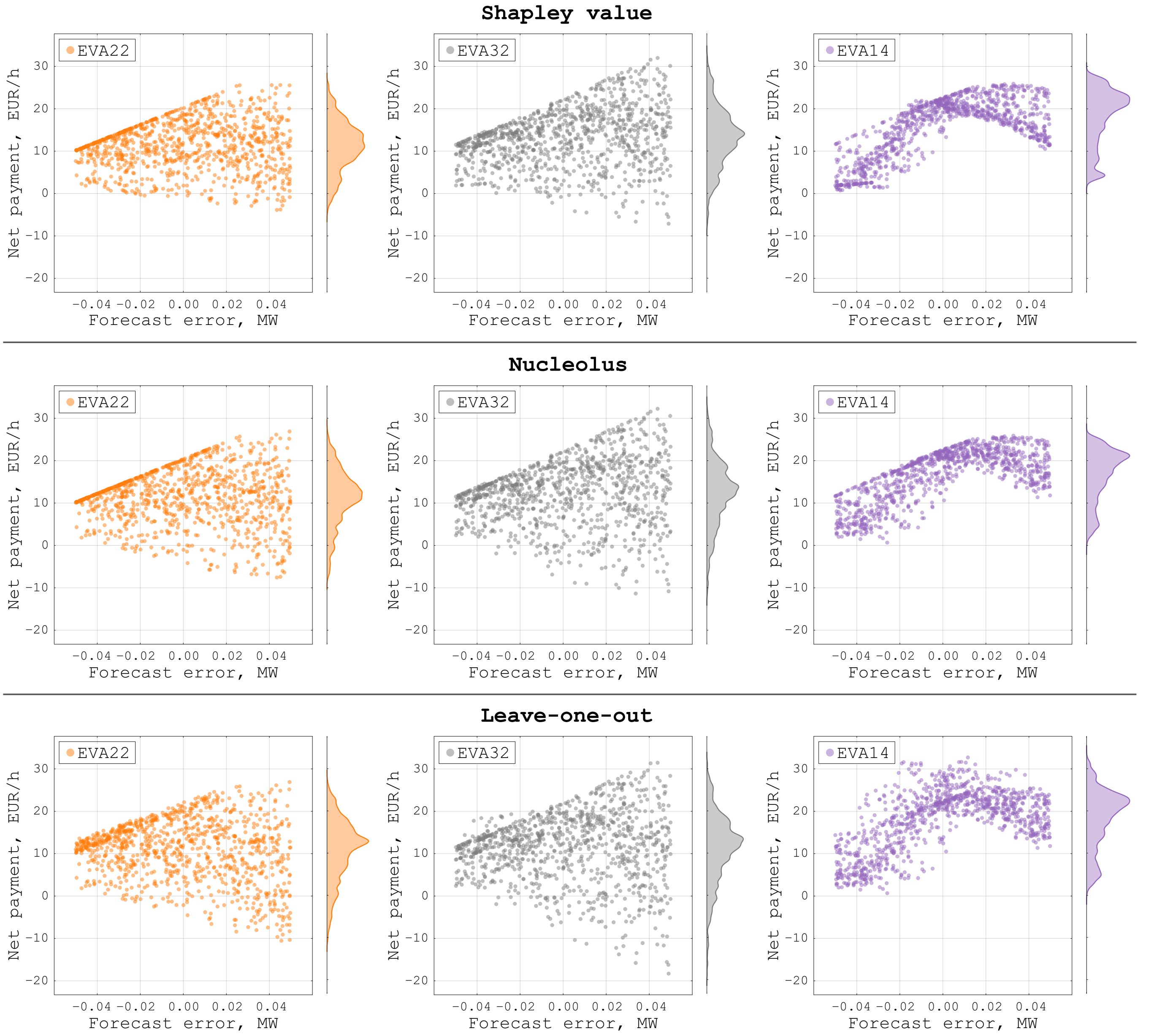

Fig. 4. Monte Carlo analysis of 1000 EVA load forecasts using different forecast valuation mechanisms: (top) the Shapley value; (middle) the Nucleolus; (bottom) the LOO method. Net payments indicate the allocation of the total avoided cost due to forecast sharing, which is positive in all simulated scenarios. Negative payment values correspond to penalties assigned to some EVAs. Forecast errors are calculated as the difference between the EVA load forecast and load observations. Right-hand density plots show the distribution of payments for each EVA, with wider regions indicating more frequent outcomes.

resulting in a fully non-negative and symmetric allocation rule for all EVAs. This illustrates how introducing an accuracy-based scoring can mitigate penalties for accurate forecasts and reduce sensitivity to network-induced asymmetries. However, a purely accuracy-based valuation does not reflect the true monetary impact of EVA forecasts and the value of their locations. Therefore, a balance between accuracy-based and value-oriented approaches is required for practical implementation.

V. Applicability and Scalability

This section discusses practical aspects of implementing forecast-sharing markets, including market design properties, axiomatic properties of the accuracy-adjusted Shapley value, and computational scalability.

A. Market Design Properties

The proposed forecast valuation framework is defined based on the outcomes of the day-ahead and real-time balancing markets. Although the forecast-sharing market is operated separately from electricity market clearing, it introduces monetary transfers between the system operator and forecast providers. Therefore, to apply this framework in practice, it is necessary to clarify its market design properties such as *cost recovery*, *revenue adequacy*, and *market efficiency* [22]–[24].

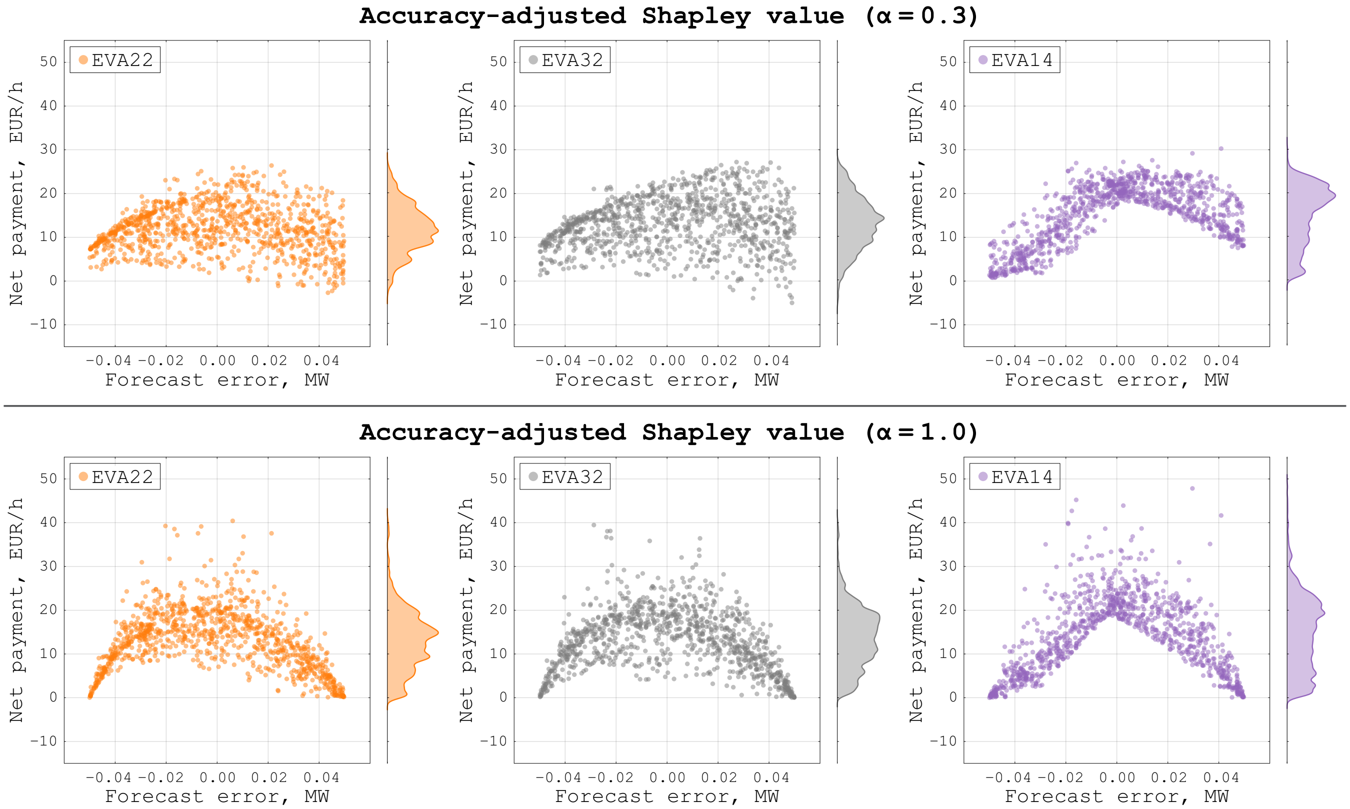

Fig. 5. Monte Carlo analysis of 1000 EVA load forecasts using the accuracy-adjusted Shapley value mechanism: (top) with $\alpha = 0.3$; (bottom) with $\alpha = 1.0$. Parameter α determines the accuracy-based share of payments. When $\alpha = 1.0$, the payments are calculated entirely on the basis of forecast accuracy, without using the Shapley value. Points show net payment to each EVA (EUR/h) versus its forecast error (load forecast minus load observation, in MW). Right-hand density plots show the distribution of payments for each EVA, with wider regions indicating more frequent outcomes.

An overview of these properties in the context of a two-stage forecast-sharing market design is provided below:

- Cost recovery property requires that the compensation paid to each forecast provider should equal or exceed the cost of acquiring or producing data. Even though producing load forecasts may incur non-negligible costs, studies on forecasting and regression markets typically assume that providers are willing to accept any non-negative payment for their data [25], [26]. However, as demonstrated in Section IV, providers may receive negative payments (penalties) under value-oriented mechanisms. Thus, this property cannot be guaranteed in general, unless additional constraints or rules are introduced.
- Revenue adequacy refers to the ability of the market operator to meet settlement obligations, that is, to make payments to all forecast providers without incurring a financial deficit. In the proposed framework, both game-theoretic and accuracy-based mechanisms are budget-balanced with respect to the value of the grand coalition: the payments to providers allocate exactly the total avoided cost. However, in reality, savings due to load forecasts can be volatile or even negative. Therefore, practical deployment of forecast-sharing markets would require additional rules, such as a fixed procurement budget and caps and floors for payments.
- Market efficiency refers to the ability of the forecast-sharing mechanism to support operational decisions that minimise the total system cost given the available information. If forecasts are reported truthfully, the system operator can use them to improve dispatch and reduce expected operating costs. It follows that market efficiency can be guaranteed through the optimality of the system dispatch. However, as demonstrated in Section IV, purely value-based payments can create incentives for strategic misreporting, where providers may obtain higher payments by submitting distorted forecasts. Such behaviour reduces the quality of information available to the system operator and can lead to an efficiency loss. This motivates combining value-based valuation mechanisms with accuracy-based methods that explicitly incentivise truthful reporting.

These properties indicate that, while forecast-sharing markets can, in principle, unlock operational benefits, a purely value-oriented mechanism is not sufficient for practical deployment. Physical network constraints and the strategic behaviour of providers can lead to volatile and potentially infeasible money transfers. Therefore, practical market designs will require additional mechanism layers, e.g., explicit accuracy incentives, as well as adjustments to the forecast procurement budget and payments.

TABLE III
AXIOMATIC PROPERTIES OF THE ACCURACY-ADJUSTED SHAPLEY VALUE

Property	Shapley value	Accuracy-adjusted Shapley value
Efficiency	The total value created by the grand coalition is fully distributed among players. In the forecast sharing setting, the sum of allocations to all forecast providers is exactly equal to the avoided cost of the grand coalition, with no value left undistributed. The Shapley value guarantees this property by design. ✓	As defined in (16), the accuracy-adjusted mechanism allocates the value of the grand coalition using a weighted sum of the Shapley value (9) and the accuracy scoring rule (15). Therefore, by design, the sum of payments to all providers remains equal to the total avoided cost. This property holds. ✓
Symmetry	Players who contribute in the same way are treated equally. If two players have identical marginal contributions across all possible coalitions, they should be assigned equal payments. The Shapley value guarantees this property by design. ✓	According to (16), any two providers who submit identical forecasts (with the same accuracy scores) and have identical marginal contributions across all coalitions will receive equal payments. Moreover, the outcome is independent of player labelling and order, which preserves anonymity. This property holds. ✓
Null player	A player who never changes the value of any coalition receives zero payment. In the forecast sharing setting, if using a forecast from a provider does not change the system costs, that provider has no economic contribution and is not paid. The Shapley value guarantees this property by design. ✓	Under the accuracy-adjusted mechanism, a provider may still receive a non-zero payment due to the accuracy-scoring component (15), even if their marginal contribution to all possible coalitions is zero. Note that a provider can have a zero payment for accuracy if their forecast error reaches the threshold τ. But, such an error is expected to make marginal contributions at least in some coalitions, leading to a non-zero allocation. This property is not guaranteed in general, unless additional rules are introduced for null players. ✗
Additivity (linearity)	The allocation preserves linear addition of games and value functions: if the total value is the sum of two value components, each player's total payment is the sum of their payments for each component. For example, in the forecast sharing setting, the value function is derived from the sum of dispatch costs in the day-ahead and real-time stages, as defined in (4). Additivity implies that the total payment to a provider equals the sum of its allocations for the day-ahead and real-time components. The Shapley value guarantees this property by design. ✓	The accuracy-based payment introduces a nonlinear transformation of errors into payments, as defined in (15), (16). Therefore, the resulting allocation rule is not additive in general: the allocation for a combined value does not necessarily decomposes into the sum of component-wise allocations. This property is not guaranteed in general. ✗

B. Properties of the Accuracy-adjusted Shapley Value

The desirable axiomatic properties of the Shapley value guarantee a transparent and fair distribution of coalition value among participants. However, the proposed accuracy-adjusted Shapley value modifies the allocation rule by introducing an accuracy-scoring component, as defined in (16). Therefore, the modified mechanism does not inherit all Shapley properties, as summarised in Table III. Specifically, the proposed adjustment preserves efficiency and symmetry, whereas the null player and additivity properties are not guaranteed in general. Note that the accuracy-adjusted Shapley value is designed to prioritise incentive compatibility over preserving all axiomatic properties. A more comprehensive axiomatic analysis of this mechanism is left for future research.

C. Scalability

Computational complexity of the forecast valuation framework stems from the need to solve the day-ahead and real-time dispatch optimisation problems, (2a)-(2c) and (3a)-(3b), for different load forecast inputs. The simplest approach is the accuracy-based valuation, which requires the dispatch optimisation problems to be solved twice: using baseline and improved forecasts. Then, the avoided cost due to improved forecasting can be calculated via (5) and allocated using the accuracy-scoring rule (15). In contrast, game-theoretic concepts such as the Shapley value and the Nucleolus require evaluating the value of many coalitions, i.e., solving the dispatch optimisation problems for different combinations of improved forecasts. As the number of possible coalitions, 2^N, grows exponentially with the number of players, finding optimal dispatch solutions for all coalitions becomes computationally prohibitive for markets with many providers.

Nevertheless, several approximation approaches have been proposed to estimate the Shapley value while minimising computational burden. These approaches include random sampling techniques [1], [6], [13], [27], [28], compressed sensing [29], [30], and player clustering algorithms [13]. For example, as demonstrated in [13], the combination of such approximations can make Shapley-based allocation mechanisms tractable in markets with up to 200 participants. Therefore, similar approaches are expected to enable practical forecast-sharing markets with dozens or hundreds of providers. A comprehensive scalability assessment of the proposed forecast valuation framework is left for future research.

VI. CONCLUSION

This paper investigates data valuation challenges in distribution networks, focusing on EVA load forecasting. Several value-oriented game-theoretic mechanisms, such as the

Shapley value, are applied to remunerate forecast providers. It is shown that due to network constraints, identical forecasts provided at different locations can receive unequal payments. In extreme cases, even perfect forecasts can be penalised. Furthermore, value-oriented mechanisms lack incentive compatibility, as certain providers may misreport their loads to receive higher payments. These findings highlight the limitations of existing forecast valuation approaches and raise concerns about the fairness and incentive compatibility of forecast-sharing markets in the presence of physical network constraints. Future research should explore hybrid valuation mechanisms that combine value-oriented and accuracy-based methods to ensure operational efficiency in distribution systems and fair compensation for forecast providers.

Acknowledgements

This work was supported by the Electric Power Innovation for a Carbon-free Society (EPICS) Centre, grant no. EP/Y025946/1. The authors also acknowledge support from the ViPES2X project under EUDP grant 640222-496237.

References

[1] A. Agarwal, M. Dahleh, and T. Sarkar, "A marketplace for data: An algorithmic solution," in *Proceedings of the 2019 ACM Conference on Economics and Computation*, 2019.

[2] A. Ghorbani and J. Zou, "Data Shapley: Equitable valuation of data for machine learning," in *Proceedings of the 36th International Conference on Machine Learning, ICML 2019*, 2019.

[3] R. Jia, D. Dao, B. Wang, F. A. Hubis, N. Hynes, N. M. Gurel, B. Li, C. Zhang, D. Song, and C. Spanos, "Towards efficient data valuation based on the Shapley value," in *Proceedings of the 22nd International Conference on Artificial Intelligence and Statistics*, 2019.

[4] J. Yoon, S. Arik, and T. Pfister, "Data valuation using reinforcement learning," in *Proceedings of the 37th International Conference on Machine Learning*, vol. 119. PMLR, 2020.

[5] R. H. L. Sim, X. Xu, and B. K. H. Low, "Data Valuation in Machine Learning: Ingredients, Strategies, and Open Challenges," in *Proceedings of the 31st International Joint Conference on Artificial Intelligence, IJCAI-22*, 2022.

[6] C. Goncalves, P. Pinson, and R. J. Bessa, "Towards data markets in renewable energy forecasting," *IEEE Transactions on Sustainable Energy*, vol. 12, no. 1, 2021.

[7] R. Mieth, J. M. Morales, and H. V. Poor, "Data valuation from data-driven optimization," *IEEE Transactions on Control of Network Systems*, vol. 12, no. 1, 2025.

[8] L. Han, J. Kazempour, and P. Pinson, "Monetizing Customer Load Data for an Energy Retailer: A Cooperative Game Approach," in *2021 IEEE Madrid PowerTech - Conference Proceedings*, 2021.

[9] B. Wang, Q. Guo, T. Yang, L. Xu, and H. Sun, "Data valuation for decision-making with uncertainty in energy transactions: A case of the two-settlement market system," *Applied Energy*, vol. 288, 2021.

[10] Z. Sun, L. Von Krannichfeldt, and Y. Wang, "Trading and Valuation of Day-Ahead Load Forecasts in an Ensemble Model," *IEEE Transactions on Industry Applications*, vol. 59, no. 3, 2023.

[11] Y. Zhou, Q. Wen, J. Song, X. Cui, and Y. Wang, "Load data valuation in multi-energy systems: An end-to-end approach," *IEEE Transactions on Smart Grid*, vol. 15, no. 5, 2024.

[12] A. Churkin, W. Kong, J. N. Melchor Gutierrez, E. A. Martínez Ceseña, and P. Mancarella, "Tracing, Ranking and Valuation of Aggregated DER Flexibility in Active Distribution Networks," *IEEE Transactions on Smart Grid*, vol. 15, no. 2, 2024.

[13] S. Cremers, V. Robu, P. Zhang, M. Andoni, S. Norbu, and D. Flynn, "Efficient methods for approximating the Shapley value for asset sharing in energy communities," *Applied Energy*, vol. 331, 2023.

[14] S. Lv, S. Chen, T. Zhang, C. Chen, J. Xu, and Z. Wei, "Promote data sharing in integrated power-traffic networks: A coalition game approach," *IEEE Transactions on Sustainable Energy*, vol. 16, no. 3, 2025.

[15] A. Churkin, "ForecastVal-E: Forecast Valuation for Energy Data Markets," https://github.com/AndreyChurkin/ForecastVal-E, 2025.

[16] M. E. Baran and F. F. Wu, "Network reconfiguration in distribution systems for loss reduction and load balancing," *IEEE Transactions on Power Delivery*, vol. 4, no. 2, 1989.

[17] L. S. Shapley, "A value for n-person games," in *Contributions to the Theory of Games, Volume II*. Princeton University Press, 1953.

[18] A. Churkin, J. Bialek, D. Pozo, E. Sauma, and N. Korgin, "Review of Cooperative Game Theory applications in power system expansion planning," *Renewable and Sustainable Energy Reviews*, vol. 145, 2021.

[19] M. Guajardo and K. Jörnsten, "Common mistakes in computing the nucleolus," *European Journal of Operational Research*, vol. 241, no. 3, 2015.

[20] H. Moulin, "The separability axiom and equal-sharing methods," *Journal of Economic Theory*, vol. 36, no. 1, 1985.

[21] T. S. Driessen and Y. Funaki, "Coincidence of and collinearity between game theoretic solutions," *OR Spektrum*, vol. 13, no. 1, 1991.

[22] J. M. Morales, M. Zugno, S. Pineda, and P. Pinson, "Electricity market clearing with improved scheduling of stochastic production," *European Journal of Operational Research*, vol. 235, no. 3, 2014.

[23] J. Kazempour, P. Pinson, and B. F. Hobbs, "A stochastic market design with revenue adequacy and cost recovery by scenario: Benefits and costs," *IEEE Transactions on Power Systems*, vol. 33, no. 4, 2018.

[24] G. Zakeri, G. Pritchard, M. Bjorndal, and E. Bjorndal, "Pricing wind: a revenue adequate, cost recovering uniform price auction for electricity markets with intermittent generation," *INFORMS Journal on Optimization*, vol. 1, no. 1, 2019.

[25] P. Pinson, L. Han, and J. Kazempour, "Regression markets and application to energy forecasting," *Transactions in Operations Research*, vol. 30, no. 3, 2022.

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

[27] J. Castro, D. Gómez, and J. Tejada, "Polynomial calculation of the Shapley value based on sampling," *Computers & Operations Research*, vol. 36, no. 5, 2009.

[28] J. Castro, D. Gómez, E. Molina, and J. Tejada, "Improving polynomial estimation of the Shapley value by stratified random sampling with optimum allocation," *Computers and Operations Research*, vol. 82, 2017.

[29] R. Jia, D. Dao, B. Wang, F. A. Hubis, N. Hynes, N. M. Gürel, B. Li, C. Zhang, D. Song, and C. J. Spanos, "Towards efficient data valuation based on the Shapley value," in *The 22nd International Conference on Artificial Intelligence and Statistics*, 2019.

[30] V. Erofeeva and S. Parsegov, "A novel sparsity-based deterministic method for Shapley value approximation, with applications," *Information Sciences*, vol. 704, 2025.