Promoting Load Forecast Sharing: A Cost-Oriented Coalition Game Approach

Lingfeng Tang, Haipeng Xie*, Dongrui Quan, Zhaohong Bie
School of Electrical Engineering
Xi'an Jiaotong University
Xi'an, China
{lftang, qdr2002}@stu.xjtu.edu.cn, {haipengxie, zhbie}@xjtu.edu.cn

***Abstract*—Data sharing presents a promising approach to enhancing load forecasting accuracy, however, its potential is severely limited by the absence of methods that can fully harness the data value and provide effective economic incentives. To address this gap, this paper proposes a cost-oriented coalition game approach to promote the load forecast sharing among multiple participants. Specifically, for a user requiring day-ahead load forecasts for power scheduling, a coalition game-based framework for sharing forecasts among the user and other participants is established, where the load forecast serves as the shared object, rather than raw data or forecasting models that pose privacy risks. Within the proposed framework, an ex-ante cost-oriented load forecast aggregation method is developed to combine multiple forecasts into an ensemble model, where the downstream operational cost of day-ahead scheduling and intra-day balancing, instead of the intermediate forecasting error, is directly leveraged to learn the ensemble model. On the basis, an ex-post profit allocation approach based on Shapley value is utilized to offer effective economic incentives for each participant. Case study results using real-world load data demonstrate that the proposed cost-oriented load forecast sharing approach outperforms both the user itself and accuracy-oriented methods, while maintaining coalition stability.**

***Index Terms*—data sharing, load forecasting, coalition game, ensemble learning**

I. Introduction

In the digital era, data plays a central and indispensable role in improving load forecasting accuracy, which is crucial for electricity consumers to reduce operational costs. Traditionally, only the user itself leverages its own data to learn a load forecasting model. However, with the development of distributed collaborative learning paradigms, such as federated learning [1] and ensemble learning [2], data sharing, including the exchange of forecasting models or forecasts among multiple participants, has been shown to significantly enhance load forecasting accuracy and reduce operational costs.

Regarding the collaborative learning among multiple participants for load forecasting, its practical implementation is severely hindered by the lack of motivation for data sharing. First, as the shared digital asset, a key challenge lies in how to effectively leverage data in order to fully harness its value for the operation of electricity consumers. Secondly, providing fair economic incentives for each participant according to their respective contributions is also indispensable for motivating the consistent participation in data sharing.

To address these challenges, several studies have been conducted, but the relevant exploration is still in its infancy. A data transaction framework is proposed in [3] to promote data sharing between the electricity retailer and data suppliers for uncertainty reduction. Then, a non-regret data auction mechanism is developed in [4], where the central agent aggregates raw data from sellers into a linear quantile regression model and allocates the resulting profit to each seller. On the basis, a general regression market is systematically introduced in [5]. A least absolute shrinkage and selection operator (LASSO)-based regression market is further developed in [6] to determine personalized data payment thresholds. Beyond raw data sharing, [7] proposes a novel forecasting model trading approach that enables the bilateral transaction and purchased model adaptation. Moreover, a day-ahead forecast output trading framework is established in [8] to improve the forecasting accuracy through ensemble learning.

However, the studies mentioned above still face three unresolved issues. i) Most data-sharing methods involve the exchange of raw data among participants [3]–[6], which poses significant privacy risks. While forecasting model trading, as proposed in [7], offers an alternative approach to data sharing, it suffers from high communication burdens and the risk of model abuse. ii) Previous data-sharing methods primarily focus on enhancing forecasting accuracy [8], but they fail to directly link forecasting errors with operational costs, limiting their ability to fully capitalize on the value of shared data. iii) Whether profit allocation can sustain participants' consistent engagement in data sharing has been rarely explored.

To fill the gap, this paper proposes a cost-oriented coalition game approach to facilitate load forecast sharing. The main contributions lie in three folds.

1) Establish a coalition game-based load forecast sharing framework that facilitates collaboration among multiple participants to reduce operational costs of the electricity user.

2) Propose a cost-oriented load forecast aggregation method that directly leverages the downstream operational costs of the forecasting task to learn the ensemble model.

3) Develop a profit allocation approach for load forecast sharing that guarantees coalition stability, which is validated through theoretical analysis and numerical tests.

The work is supported by the National Natural Science Foundation of China (U24B6010)

979-8-3315-8138-1

II. Cost-Oriented Coalition Game-based Load Forecast Sharing Framework

A. Problem Statement

This paper considers an electricity user, such as an industrial consumer, that requires day-ahead load forecasts $\hat{\boldsymbol{p}}^{\text{load}} = \{\hat{p}_t^{\text{load}}\}_{t=1}^{T}$ to optimally schedule its energy storage system (ESS) and procure power over the time horizon T. The objective of the day-ahead scheduling stage is to minimize operational costs C_1 while ensuring load satisfaction, which can be formulated as:

$$\min C_1 = \sum_{t=1}^{T}[\lambda_t^{\text{g}} p_t^{\text{g}} + \beta_1 (p_t^{\text{ess, dis}})^2 + \beta_2 p_t^{\text{ess, dis}}] \tag{1a}$$

$$\text{s.t.} \quad p_t^{\text{g}} + p_t^{\text{ess, dis}} = \hat{p}_t^{\text{load}} + p_t^{\text{ess, ch}} \tag{1b}$$

$$e_{t+1}^{\text{ess}} = e_t^{\text{ess}} + (\eta^{\text{ch}} p_t^{\text{ess, ch}} - p_t^{\text{ess, dis}}/\eta^{\text{dis}})\Delta t \tag{1c}$$

$$0 \le p_t^{\text{ess, ch(dis)}} \le \overline{p}^{\text{ess, ch(dis)}} \tag{1d}$$

$$0 \le e_t^{\text{ess}} \le \overline{e}^{\text{ess}} \tag{1e}$$

where p_t^{g} represents the purchased power at the price λ_t^{g}. $p_t^{\text{ess, ch(dis)}}$ and e_t^{ess} denote the charging/discharging power and the stored energy of the ESS, limited by $\overline{p}^{\text{ess, ch(dis)}}$ and $\overline{e}^{\text{ess}}$, respectively. η^{ch} and η^{dis} represent the charging and discharging efficiencies, respectively. The second and third terms in the objective represent the degradation costs of the ESS, with the charging power component excluded due to the symmetry of each cycle [9]. β_1 and β_2 are the corresponding coefficients.

In the intra-day balancing stage, due to imperfect load forecasting, over-estimation results in the procurement of excess day-ahead power, incurring a corresponding penalty with the coefficient λ^{pe}. Conversely, under-estimation necessitates the purchase of real-time power balancing services at the price λ_t^{rt}. Hence, the cost C_2 of the intra-day balancing stage can be written as:

$$C_2 = \sum_{t=1}^{T}(\lambda^{\text{pe}}[\hat{p}_t^{\text{load}} - p_t^{\text{load}}]^+ + \lambda_t^{\text{rt}}[\hat{p}_t^{\text{load}} - p_t^{\text{load}}]^-) \tag{2}$$

where $[x]^+$ and $[x]^-$ denote the functions $\max(x,0)$ and $\max(-x,0)$, respectively.

However, the user may face challenges in producing accurate forecasts due to limited expertise or insufficient focus on forecasting tasks. A practical solution for the user is to outsource the forecasting task, providing the task goal and historical load dataset, while offering incentives to participants to submit reference forecasts. The user can then combine its own forecast with those from participants to generate a final forecast for day-ahead scheduling.

B. Coalition Game Formulation of Load Forecast Sharing

To promote load forecast sharing, a cost-oriented coalition game is utilized to coordinate the user and multiple participants in generating a day-ahead load profile. The cost reductions in day-ahead scheduling and intra-day balancing stages, resulting from improved forecasting, are treated as the profit, which is subsequently allocated among the forecast-sharing participants based on their individual contributions.

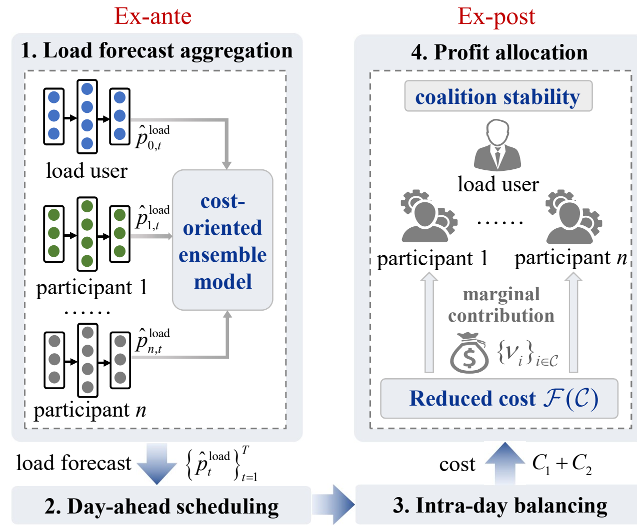

Fig. 1. Proposed load forecast sharing framework

Let the coalition game be populated by the non-empty player set $\mathcal{N} = \{L, P_1, P_2, \dots, P_n\}$, consisting of the load user and n forecast-sharing participants. A coalition is defined as a subset $\mathcal{C} \subseteq \mathcal{N}$, with the grand coalition being the set $\mathcal{N}$, which includes all players. Therefore, a characteristic function $\mathcal{F}(\mathcal{C})$, which maps each coalition $\mathcal{C}$ to its reduced operational cost, can be defined as:

$$\mathcal{F}(\mathcal{C}) = C\left[\hat{\boldsymbol{p}}^{\text{load}}(\boldsymbol{x}^{\mathcal{C}})\right] - C\left[\hat{\boldsymbol{p}}^{\text{load}}(\boldsymbol{x}^{\mathcal{C}_0})\right] \tag{3}$$

where $C\left[\hat{\boldsymbol{p}}^{\text{load}}(\boldsymbol{x}^{\mathcal{C}})\right]$ denotes the sum of day-ahead and intra-day operational costs under the load forecasting $\hat{\boldsymbol{p}}^{\text{load}}(\boldsymbol{x}^{\mathcal{C}})$, $\boldsymbol{x}^{\mathcal{C}}$ is the collection of forecasts from the coalition $\mathcal{C}$. $\mathcal{C}_0$ represents the coalition only containing the user. Note that the value of any coalition excluding the user is zero.

To guarantee the coalition stability, a profit allocation set denoted as $\{\nu_i\}_{i\in\mathcal{C}}$ that satisfies the efficiency condition, the individual rationality condition and the coalition rationality condition should be formulated. The above coalition stability conditions can be written as:

$$\sum_{i\in\mathcal{N}} \nu_i = \mathcal{F}(\mathcal{N}) \tag{4a}$$

$$\nu_i \ge \mathcal{F}(i), i \in \mathcal{N} \tag{4b}$$

$$\sum_{i\in\mathcal{C}} \nu_i \ge \mathcal{F}(\mathcal{C}), \mathcal{C} \in \mathcal{N} \tag{4c}$$

Remark 1 (Challenges for Load Forecast Sharing): Within the coalition game framework, two key challenges need to be addressed: i) How can the data value of shared forecasts be fully leveraged to reduce operational costs? ii) How should the incremental profit be allocated to ensure coalition stability?

C. Framework Description

To address the above two challenges, a load forecast sharing framework is proposed as depicted in Fig. 1:

1) Ex-ante cost-oriented load forecast aggregation refers to the process where the load user combines its own load

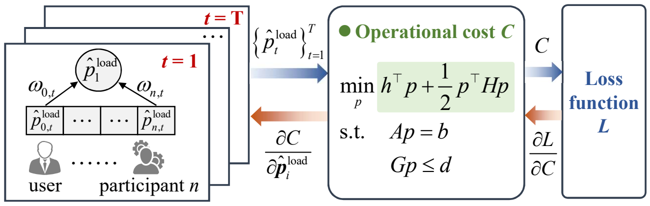

Fig. 2. End-to-end gradient derivation for ensemble model learning

forecasts with those of n forecast-sharing participants into an ensemble model. The combination weights are determined based on the reduced operational costs, rather than the improved forecasting accuracy, as described in Section III.

2) Ex-post profit allocation refers to the allocation of the reduced costs among the load user and n forecast-sharing participants based on the Shapley value, which ensures the coalition stability as illustrated in Section IV.

III. Ex-Ante Cost-Oriented Load Forecast Aggregation in an Ensemble Model

A. Cost-Oriented Ensemble Model

To combine load forecasts from the user itself and n forecast sharing participants, the weighted averaging ensemble model is employed to obtain the final forecast values:

$$\hat{p}_t^{\text{load}} = \sum_{i \in \mathcal{N}} \omega_{i,t} \hat{p}_{i,t}^{\text{load}} + b_t \tag{5}$$

where $\omega_{i,t}$ denotes the forecast weight for i-th forecast at time step t, and b_t represents the intercept at time step t. All parameters in the ensemble model is collected and denoted as θ, where the total parameter number is $T(n+2)$.

In contrast to previous ensemble models that solely focus on the forecasting accuracy minimization, a cost-oriented load forecasting ensemble model is proposed to directly leverage the downstream reduced operational costs of the forecasting task to optimize θ. Hence, the loss function L of the cost-oriented load forecasting ensemble model can be defined as:

$$L = \mathbb{E}_{(\boldsymbol{x}_i^{\mathcal{C}}, \boldsymbol{p}_i^{\text{load}}) \sim D} [C(\hat{\boldsymbol{p}}_i^{\text{load}}(\boldsymbol{x}_i^{\mathcal{C}})) - C(\boldsymbol{p}_i^{\text{load}})] \tag{6}$$

where D denotes the dataset consisting of samples $(\boldsymbol{x}_i^{\mathcal{C}}, \boldsymbol{p}_i^{\text{load}})$.

B. End-to-end Gradient Derivation for Model Learning

To minimize the loss function, an end-to-end gradient of L with respect to θ can be derived through the ensemble model, as well as the day-ahead scheduling and intra-day balancing process, based on the chain rule as depicted in Fig. 2:

$$g = \frac{\partial L}{\partial C} \frac{\partial C}{\partial \hat{\boldsymbol{p}}_i^{\text{load}}} \frac{\partial \hat{\boldsymbol{p}}_i^{\text{load}}}{\partial \theta} \tag{7}$$

where the first and the third fractions both correspond to the gradients of linear functions. Hence, the remainder of this section will focus on calculating the derivative of C with respect to $\hat{\boldsymbol{p}}_i^{\text{load}}$ through the optimization-based day-ahead scheduling and intra-day balancing process.

As shown in Eq. (2), the intra-day balancing cost C_2 is independent from the day-ahead scheduling result, thus it can be incorporated into the objective function in Eq. (1) to constitute the total cost C. The resulting optimization model can be written as a compact convex quadratic optimization form:

$$\min_p \quad h^\top p + \frac{1}{2} p^\top H p \tag{8a}$$

$$\text{s.t.} \quad Ap = b \tag{8b}$$

$$Gp \leq d \tag{8c}$$

where p represents all decision variables p_t^{g}, $p_t^{\text{ess, ch}}$ and $p_t^{\text{ess, dis}}$ over the time horizon T. The collection of all matrices and coefficient vectors $\{h, H, A, b, G, d\}$ is denoted as Θ. The Karush-Kuhn-Tucker (KKT) conditions hold for this convex quadratic optimization problem, which can be written as:

$$h + Hp^* + A^\top \upsilon^* + G^\top \mu^* = 0 \tag{9a}$$

$$D(\mu^*)(Gp^* - d) = 0 \tag{9b}$$

$$Ap^* - b = 0 \tag{9c}$$

where p^* denotes the optimal decision variables. υ^* and μ^* are the optimal dual variables for the equality and inequality constraints, respectively. $D(\cdot)$ represents the diagonal matrix.

As observed in Eq. (9), it can be regarded as a implicit function $J(\mathcal{P}, \Theta) = 0$, where $\mathcal{P}$ denotes the collection of the optimal primal and dual variables. Therefore, on the basis of the implict function theorem, the derivative of C with respect to Θ can be calculated as:

$$\frac{\partial C}{\partial \Theta} = \frac{\partial C}{\partial \mathcal{P}} [-(\frac{\partial J}{\partial \mathcal{P}})^{-1} \frac{\partial J}{\partial \Theta}] \tag{10}$$

where these three fractions can all be computed by the partial differentiation operation. The final form is not presented here due to page limits. Note that $\hat{\boldsymbol{p}}_i^{\text{load}}$ is actually part of Θ, thus the derivative of C with respect to $\hat{\boldsymbol{p}}_i^{\text{load}}$ can be directly obtained by extracting the relevant values from Eq. (10).

IV. Ex-Post Profit Allocation of Load Forecast Sharing Participants

To incentivize the load forecast sharing process, profit allocation is performed after the operational cost is received. As a widely recognized profit allocation method, the Shapley value is employed in this paper to allocate the reduced operational cost based on each participant's marginal contribution.

Specifically, the load forecast sharing coalition $\mathcal{N}$ can be formed by all participants with different joining sequences. In a given joining sequence, the corresponding marginal contribution of the participant i is the difference in coalition value before and after its inclusion. Thus, the marginal contribution of the participant i can be calculated as:

$$\nu_i = \frac{1}{|\mathcal{O}_{\mathcal{N}}|} \sum_{o \in \mathcal{O}_{\mathcal{N}}} \triangle_o(i) \tag{11}$$

where $|\mathcal{O}_{\mathcal{N}}|$ represents the number of all possible joining sequences. $\triangle_o(i)$ denotes the incremental value after the participant i joins in the coalition obeying the sequence o.

979-8-3315-8138-1

On the basis, Eq. (11) can be equivalently transformed into the following form for Shapley value computation:

$$\nu_i = \sum_{\mathcal{C} \subseteq \mathcal{N} \setminus \{i\}} \frac{|\mathcal{C}|!(|\mathcal{N}| - |\mathcal{C}| - 1)!}{|\mathcal{N}|!} (\mathcal{F}(\mathcal{C} \cup \{i\}) - \mathcal{F}(\mathcal{C})) \quad (12)$$

Lemma 1: The efficiency condition of the load forecast sharing coalition $\mathcal{N}$, i.e., $\sum_{i \in \mathcal{N}} \nu_i = \mathcal{F}(\mathcal{N})$ in Eq. (4a), consistently holds.

Proof: In each joining sequence o, each participant in the load forecast sharing coalition joins the coalition sequentially, ultimately forming the coalition $\mathcal{N}$. Obviously, in this sequence o, the sum of all participants' marginal contributions equals the total reduced operational cost $\mathcal{F}(\mathcal{N})$ of coalition $\mathcal{N}$. Therefore, the summation of ν_i represents the average of the sum of participants' marginal contributions across all sequences, which always equals $\mathcal{F}(\mathcal{N})$.

Lemma 2: The individual rationality condition of the load forecast sharing coalition, i.e., $\nu_i \geq \mathcal{F}(i), i \in \mathcal{N}$, holds.

Proof: For a coalition solely comprising any individual participant, the reduced operational cost is zero, i.e., $\mathcal{F}(i) = 0, \forall i \in \mathcal{N}$. Regarding the coalition $\mathcal{N}$ formed by all participants, even if the forecasts shared by all participants negatively impact operational cost reduction, the proposed method still degenerates to the user's own model, guaranteeing a non-negative coalition profit. Specifically, we can set all weights and intercepts in the weighted averaging ensemble model to zero, except for those associated with the user. Accordingly, the forecast results of coalition $\mathcal{N}$ are identical to those of the user, and consequently, the allocated profit $\{\nu_i\}_{i \in \mathcal{N}}$ for each participant will be zero. Thus, the individual rationality condition remains satisfied even in this extreme scenario.

Regarding the coalition rationality condition, it may not be satisfied when the forecasts from different participants is highly homogeneous. Thus, selecting appropriate participants is crucial for maintaining coalition stability, but this is beyond the scope of this paper. In the case study presented below, in which participants use different models and features, the results indicate that the coalition stability is maintained.

V. Case Study

A. Experimental Setup

1) Test System Setup: The load user possesses an ESS with a capacity of 0.3 MW and a maximum charge/discharge power of 0.15 MW, where both charging and discharging efficiencies are set to 0.93. The linear and quadratic coefficients of the ESS degradation cost are 2 \$/MW and 2 \$/MW^2, respectively. In the day-ahead scheduling stage, the user procures electricity to meet the load demands for the subsequent day. In the intra-day stage, penalties arise from imperfect load forecasts. Specifically, a penalty of 1.5 times the day-ahead price is incurred for excessive electricity purchases resulting from overestimating the load, while a penalty of 2 times the day-ahead price is imposed for procuring power balance services due to underestimating the load.

2) Dataset Description: The real-world load dataset from the Global Energy Forecasting Competition 2012 (GEFCom2012) is employed in this case study [10]. The load data, covering the years 2004 to 2007 with an hourly resolution, is selected as the user's power consumption data, scaled to 1 MW. In addition, corresponding hourly outdoor temperature data and calendar information, including the cosine and sine values for the month of the year, day of the week, day of the year, and holiday data for these four years, are also obtained from GEFCom2012. The day-ahead price p_t^{g} for the user is collected from the day-ahead hourly locational marginal prices (LMPs) in the PJM market [11]. The above data is split into training, validation and test datasets in a 6:2:2 ratio, where all load forecast sharing participants share the same data partition.

3) Load Forecast Sharing Setup: By releasing the load forecasting task, the load user and three participants form a coalition to jointly engage in the load forecast sharing process. Due to differences in forecasting knowledge, preferences, and available data, participants tend to choose different forecasting models and input features. The detailed setup of the load forecast sharing participants is shown in Table I. The load user and Participant 1 employ a three-layer multi-layer perceptron (MLP) model and a long short-term memory (LSTM) model with two hidden layers, respectively. Participants 2 and 3 both use a dual model, which includes one LSTM model for processing time-series input features and one MLP model for extracting non-temporal input features. A linear output layer is used to establish the relationship between the concatenated extracted features from the MLP and LSTM models and the final load forecasting result.

TABLE I
Detailed Setup of Load Forecast Sharing Participants

Participant	Model	Input Features	Hyperparameters
Load User	MLP	Historical 7-day load data	Training epochs = 100 Batch size = 16 Adam optimizer = True Learning rate = 1e-3 Hidden dimension = 64
Participant 1	LSTM	Historical 3-day load and temp. data	
Participant 2	Dual-M	Historical 5-day load data, calendar data	
Participant 3	Dual-M	Historical 3-day load and temp. data, calendar data	

Based on the above dataset and the trained models of each participant, both the cost-oriented and accuracy-oriented ensemble models are trained using identical hyperparameters and neural network architecture.

B. Effectiveness Evaluation of Cost-oriented Load Forecast Sharing and Aggregation

To highlight the significance of load forecast sharing and demonstrate the advantages of the proposed cost-oriented method, the forecasting models of each participant, along with the cost-oriented and accuracy-oriented forecast aggregation methods, are evaluated in terms of prediction accuracy and operational cost on the test set, as shown in Fig. 3. The prediction accuracy is measured using the Mean Absolute

979-8-3315-8138-1

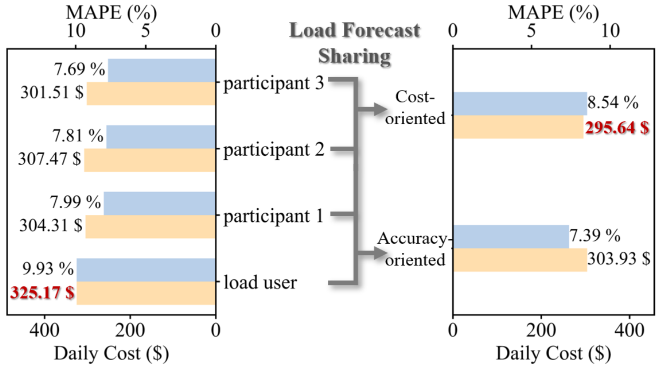

Fig. 3. Performance evaluation of cost-oriented and accuracy-oriented forecast aggregation methods, as well as forecasting models of individual participants

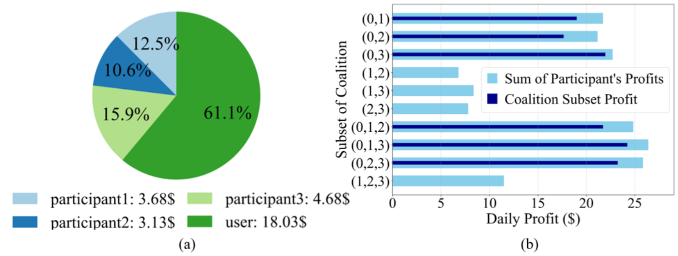

Fig. 4. Results of profit allocation. (a) Allocated profits for each participant. (b) Comparison of coalition subset profit with the sum of the corresponding participants' allocated profits, where the number 0 refers to the load user.

Percentage Error (MAPE) metric, and the operational cost is the sum of the day-ahead and intra-day operational costs based on the day-ahead load forecast results.

As observed in Fig. 3, the proposed cost-oriented load forecast sharing method outperforms the individual forecasting models of all participants in terms of operational cost. Compared to the user's own forecasting model without data sharing, the proposed method reduces the daily operational cost from 325.17$ to 295.64$. Additionally, due to the different optimization objectives of the ensemble models for load forecast aggregation, the cost-oriented method achieves the lowest operational cost, while the accuracy-oriented method results in the lowest MAPE. This phenomenon suggests that the relationship between forecast error and operational cost is not intuitively positively correlated. In other words, smaller forecast errors do not necessarily lead to lower operational costs. Thus, the proposed cost-oriented method can more effectively leverage the value of shared load forecasts.

C. Analysis of Profit Allocation and Coalition Stability

Fig. 4 shows the profit allocation outcomes for all load forecast sharing participants using a pie chart, along with a comparison of the profits for all coalition subsets (excluding the empty set and subsets consisting of only a single participant) and the sum of the allocated profits for the corresponding participants, presented through a bullet chart.

As shown in Fig. 4(a), since the user initiated the forecasting task and provided historical load data, with the forecast results guiding the user's day-ahead scheduling plan, the forecasts from other participants generate no economic benefit in the user's absence. As a result, the user receives more than 60% of the profits. Moreover, the reduction in operational costs due to improved forecasting results is allocated to all participants, satisfying the coalition's efficiency condition, which aligns with the analysis in Lemma 1. Moreover, since any coalition consisting of a single participant always has a profit of zero, and all participants in Fig. 4(a) receive positive allocated profits, the individual rationality condition is satisfied as well.

As observed in Fig. 4(b), the profits of all coalition subsets are smaller than the sum of the allocated profits for the corresponding participants, indicating that the load forecasts provided by each participant contribute heterogeneously and complementarily to reducing operational costs. Therefore, the coalition rationality condition is satisfied. Based on the above numerical analysis, the coalition stability under the proposed load forecasting sharing method and among these forecast-sharing participants is confirmed.

VI. Conclusions

This paper proposes a cost-oriented coalition game approach to promote load forecast sharing. Unlike accuracy-oriented schemes, the operational costs induced by forecasting errors are directly leveraged to learn an ensemble model that combines load forecasts from multiple participants. Then, a profit allocation method based on the Shapley value provides economic incentives for each participant. On the above basis, the coalition stability is verified through a combination of theoretical analysis and experimental validation. Future work will focus on data valuation for numerous data owners from the perspective of energy system operations.

References

[1] L. Tang, H. Xie, X. Wang, and Z. Bie, "Privacy-preserving knowledge sharing for few-shot building energy prediction: A federated learning approach," *Appl. Energy*, vol. 337, p. 120860, May 2023.

[2] H.-Y. Su and C.-C. Lai, "Toward improved load forecasting in smart grids: A robust deep ensemble learning framework," *IEEE Trans. Smart Grid*, vol. 15, no. 4, pp. 4292-4296, Jul. 2024.

[3] B. Wang, Q. Guo, and Y. Yu, "Mechanism design for data sharing: An electricity retail perspective," *Appl. Energy*, vol. 314, p. 118871, May 2022.

[4] C. Goncalves, P. Pinson, and R. J. Bessa, "Towards data markets in renewable energy forecasting," *IEEE Trans. Sustain. Energy*, vol. 12, no. 1, pp. 533–542, Jan. 2021.

[5] P. Pinson, L. Han, and J. Kazempour, "Regression markets and application to energy forecasting," *TOP*, vol. 30, no. 3, pp. 533–573, May 2022.

[6] L. Han, P. Pinson, and J. Kazempour, "Trading data for wind power forecasting: A regression market with lasso regularization," Electr. Power Syst. Res., vol. 212, p. 108442, Nov. 2022.

[7] D. Qin, P. Pinson, and Y. Wang, "Load forecasting model trading: A cost-oriented and auction-based approach," *IEEE Trans. Smart Grid*, *early access*, 2025.

[8] Z. Sun, L. Von Krannichfeldt, and Y. Wang, "Trading and valuation of day-ahead load forecasts in an ensemble model," *IEEE Trans. Ind. Appl.*, vol. 59, no. 3, pp. 2686–2695, May 2023.

[9] A. Porras, R. Fern´andez-Blanco, J. M. Morales, and S. Pineda, "An efficient robust approach to the day-ahead operation of an aggregator of electric vehicles," *IEEE Trans. Smart Grid*, vol. 11, no. 6, pp. 4960-4970, Nov. 2020.

[10] T. Hong, P. Pinson, and S. Fan, "Global energy forecasting competition 2012," *Int. J. Forecasting*, vol. 30, no. 2, pp. 357–363, Apr. 2014.

[11] PJM, "Day-Ahead Hourly LMPs," [Online]. Available: https://dataminer2.pjm.com/feed/da hrl lmps/definition.

979-8-3315-8138-1