Cooperative trading of a price-maker wind power producer: A data-driven approach considering uncertainty

Rongquan Zhang, Saddam Aziz, Member, IEEE, Sadiq Ahmad, Member, IEEE
Rizwan Qureshi, Member, IEEE, Gangqiang Li Senior Member, IEEE, and Siqi Bu, Senior Member, IEEE

Abstract—This paper presents a novel framework for cooperative trading in a price-maker wind power producer, that participates in the short-term electricity balance markets. In this framework, market price uncertainty is first modeled using a price uncertainty predictor, consisting of ridge regression (RR), nonpooling convolutional neural network (NPCNN), and linear quantile regression (LQR). RR is employed to select the correlated features to the corresponding forecast day, NPCNN is employed to extract the nonlinear features, and LQR is employed to estimate the price uncertainty. Then, an improved firefly algorithm (IFA) is proposed to solve the optimization problem. IFA uses the adaptive moment estimation method to improve the convergence speed and search for the global solution. Finally, the Shapley value is employed for the profit distribution of cooperative power producers. Illustrative examples show the effectiveness of the proposed framework and optimization model.

Index Terms—wind power; uncertainty; cooperative; convolutional neural network; improved firefly algorithm

I. INTRODUCTION

In recent years, sustainable energy projects, such as wind power gained popularity, as a way to minimise air pollution and global climate change [1]. However, the uncertainty and the limited predictive capability of wind power production may cause significant challenges to the power market. The latest data from China’s National Energy Administration indicates that the installed number of wind power in northwest China, such as Gansu province, accounts for more than 20% of the total power capacity with a steady growing rate [2].

As a result of this development, electricity prices have risen sharply, as the costs of production deviations, from the generation dispatching plan [3]. Facing this circumstances, wind power producers are requested by market operators to participate in the power market as a price-maker (this is, bidding price and quantity) [4]. As a result, understanding how to create fair bidding behaviour and maximise profitability is critical, for a wind power producer.

There are several possible ways of improving the economic profits of the wind power producer. The first is based on data-driven approaches to reduce the market uncertainties, i.e., wind power generation (WPG) and electricity price (EP). Data-driven approaches can broadly split into statistical models, machine learning models, and deep learning models. Many efforts are made in the past for devising optimal bidding strategies [5], [6]. In [7], Li et al. analyzed the trading model for a wind power producer considering uncertainties of WEG and EP by adopting four different statistical models. It showed that advanced forecasting methods not only reduce penalty costs but also improve the expected profits. However, the statistical models reflect the dynamic trend by linear curve fitting, thus maybe not suitable for forecasting wind power generation and electricity price in high-dimension [8].

In [9], Catalao et al. used a hybrid intelligent approach consisting of the wavelet transform, particle swarm optimization, and adaptive-network-based fuzzy inference system (machine learning models) to generate the scenarios for electricity prices and wind power production, and then proposed a two-stage stochastic-programming-based bidding model for the wind power producer. The drawback is that, machine learning models are prone to over-fitting and gradient vanishing as the volume and dimension of data increases [10].

Deep learning, perhaps the most powerful forecasting technology, is widely used nowadays, due to the large datasets and computational power available today. In [11], Zhang et al. proposed a hybrid deep-learning framework to predict day-ahead electricity prices. In [12], Wang et al. exploited a deep-learning-based ensemble approach for probabilistic wind power forecasting. The experimental results showed [11], [12] superior forecasting performance of deep learning models, compared to statistical models and machine learning models.

In [13], a stochastic-based decision-making framework for wind power producers and demand response aggregators was modeled for participating in the day-ahead market bidding. The results showed that the cooperative bidding of WES and demand response (DR) aggregators could improve its expected profits, and alleviate the uncertainty risk for wind power production. A survey of optimization approaches in smart grid is presented in [14], and communication technologies in [15]. In [16], a stochastic bidding model for cooperative operation between wind power producer and energy storage was developed to maximize their profits. To the authors’ knowledge, cooperative transaction among power producers in previous studies has received a little attention.

In [17], an optimization model for a price-maker wind
power producer in the day-ahead market has been employed to analyze its effect on prices and maximize its profits. In [18], a profit maximization model for a price-maker wind power producer in the real-time market was developed as a mathematical program model with equilibrium constraints, and convert into a mixed-integer linear program model, and then resolved by off-the-shelf optimization software. In [19], a bi-level optimization model for a price-maker wind power producer was established to obtain the optimal bidding strategy in the day-ahead and real-time market.

However, the bidding models for wind power producers in previous studies did not consider opponent bidding behaviors and network constraints. Electricity market games are perceived to be an effective method to circumvent these issues. Electricity market games can be specifically classified into game theory and heuristic approaches. Because heuristic approaches such as genetic algorithm (GA), particle swarm optimization algorithms (PSO), and competitive co-evolutionary algorithms (CEA) can capture the dynamic adaptive nature of the opponents’ behavior than equilibrium models, it has received increasing attention to analyzing bidding strategies for traditional power producers in power market [20].

In [21], GA is used to determine bidding strategies for traditional power producers in electricity markets with elastic demand. Results show that the heuristic approach is different from the game theory due to adaptive learning. In [22], Ahmed et al. proposed a new decomposition-based PSO algorithm to solve bid prices and quantities for traditional power producers. In [23], Ladjic et al. used competitive co-evolutionary algorithms (CEA) to obtain traditional power producers’ bidding strategies in a deregulated electricity market. However, GA, PSO, and CEA have some disadvantages, such as fall into the local optima and slow convergence [24]. Thus, the capability of these ways to obtain a bidding strategy is severely restricted.

The main contributions of this work can be summarized as follows:

1) An uncertainty predictor consisting of ridge regression (RR), non-pooling convolutional neural network (NPCNN), and linear quantile regression is proposed to forecast the day-ahead electricity price. Ridge regression is used to select the correlated features to the corresponding different hours of forecast day, CNN is employed to extract complicated nonlinear features of electricity price, and linear quantile regression is used to evaluate the uncertainty of day-ahead electricity price. All of these are designed to improve the prediction accuracy.

2) A cooperative bidding model for a price-maker wind power producers considering day-ahead electricity price uncertainty is formulated to maximize its profits. Besides, Shapley value [25] is used to distribute the profits for cooperative power producers.

3) Improved firefly algorithm (IFA), inspired by the adaptive moment estimation method is proposed to solve the cooperative bidding model. IFA is capable of improving the convergence rate and searching for a global solution compared to the traditional heuristic algorithm.

The rest of this paper is organized as follows. The cooperative bidding model for a wind power producer with uncertainty is given in Section 2. The framework is given in Section 3. Case studies are presented and analyzed in Section 4, and conclusions are given in Section 5.

II. PROPOSED METHODOLOGY

A. Electricity Balancing Market

In this paper, we consider that wind power producers in electricity trading participate in two-settlement electricity markets. Wind power producers is a price-taker in the day-ahead market but a price-maker in the real-time market. Day ahead (DA) market settlement is based on hourly pre-scheduled quantities and RT prices, and real time (RT) market settlement is based on the deviations between hourly actual scheduled quantities and hourly pre-scheduled quantities and RT prices. Besides, wind power producers must pay penalty cost due to the limited predictive capability of wind power production.

Under this market framework, this paper proposes a bi-level stochastic cooperative bidding model for a price-maker wind power producer in the real-time market. The top-layer model helps to determine the optimal bidding strategy to maximize the profits with the uncertainties of DA prices and wind power production. The lower-layer model aims to the ISO’s optimization. Given space limitations, DA prices are obtained by using an uncertainty predictor, but wind power productions are obtained using Monte Carlo method. A complete framework is shown in Figure 1. Besides, the additional assumptions are also introduced for the sake of clarity:

- Up/down ramp limit and startup/ shutdown costs of traditional power producers are ignored;
- Conventional power producers in the DA market and contract trading are not considered;
- It is assumed that deviation penalty rates are known.

B. The Top-level model

The top-level model to maximize the profits for the wind power producer with the DA market price uncertainty and cooperation can be described as follows

\[
F = \max \left\{ g^n_{in}, \left\{ \sum_{o1} \sum_{o2} \pi^{o1,02} g^{o1,02} + \sum_{i} g^n_i \right\} \right\}
\]

subject to

\[
\begin{align*}
    g^n_i &= \rho_i q_i - C_i \\
    q^{01,02} &= \begin{cases} 0 & z > g^{01,02} \\
                        z & z \geq g^{01,02} \end{cases}
\end{align*}
\]
where \( i \) and \( w \) denote a conventional and wind power producer, respectively. The superscript/subscript \( n \) and \( c \) are the non-cooperative and cooperative behavior, respectively; \( O_1, O_2 \) and \( \pi^{0_1,0_2} \) represent a collection of price scenarios, a collection of wind power generation scenarios and their joint probability; \( q_i \) represents the profits for conventional power producer \( i \); \( f_{0_1,0_2} \) represents the profits for wind power producer, \( w \) in price scenario \( 0_1 \) and wind power generation \( 0_2 \). \( pen \) is deviation penalty rates for deviation between DA schedule and RT delivery; \( \rho_i(q_w) \) and \( q_i(q_w) \) are the LMP and awarded generation quantity for power producer \( i(w) \) in the RT market. \( \rho_{w}^{0_1,0_2}(b,o) \) and \( q_{w}^{0_1,0_2}(b,o) \) are the DA market price, and power bought from or sold to the DA market for wind power producer, \( w \) in scenario \( o \), respectively. \( z_0 \) and \( \xi \) are the targeted profit and the acceptable downside risk tolerance, respectively.

\[ C_{i,t}(C_{w,t}) \text{ denotes the cost function for power producer, } i(w). \] \( C_{w} \) is the cost coefficient for wind power producer \( w \). \( \alpha_i, \beta_i \) and \( \gamma_i \) are the cost coefficients for conventional power producer \( i \). In equation 3, the first and second equality constraints are the total profits for conventional power producer \( i \) and wind power producer \( w \), respectively. The third and fourth equality constraints denote the risk constraint; The fifth and sixth equality constraints denote the operation cost for conventional power producer \( i \) and wind power producer \( w \), respectively. The sum of all scenarios’ occurrence rate as the fifth equality constraints is 1.

**C. The lower-level model**

The lower-level model is cleared based on the security constraints economic dispatch to minimize the bidding costs for all power producers, if the bidding strategy from the top-level model is submitted to the ISO, as follows:

\[
\min \sum_{f=1}^{f} \lambda f q_i + \lambda w q_w^0 \tag{3}
\]

s.t

\[
\begin{align*}
\sum_{L-1}^{L} q_L + q_w^0 - \sum_{B-1}^{B} q_b - p_{\text{loss}} &= 0 \\
\rho_{\text{loss}} &= (H)^T \left( \sum_{L-1}^{L} q_L + q_w^0 - \sum_{B-1}^{B} q_b \right) + \psi_L \\
SF_L(q_L + q_w^0 - q_b) &= F_L \\
q_{L_{\text{min}}}^l &\leq q_L \leq q_{L_{\text{max}}}^l \\
q_{w_{\text{min}}}^l &\leq q_w \leq q_{w_{\text{max}}}^l \\
F_{L_{\text{min}}} &\leq F_L \leq F_{L_{\text{max}}}
\end{align*}
\]

\( L \) and \( B \) are the numbers of lines and buses in the power system, respectively; \( O \) is the aggregations of scenarios, \( \lambda_{i,s} \) and \( q_{i,s} \) are the bidding price and quantity for segments \( s \) of conventional supplier \( i \); and are the bidding price and the awarded quantity for segments \( s \) of wind power producer \( w \) in scenario \( o \); \( q_b \) denotes the nodal load for bus \( b \); \( p_{\text{loss}} \) denotes the system loss; \( H \) and \( \psi_l \) are sensitivity and linearization offset in system loss.

\( SF \) is the shifting factor; \( F_L \) represents line flow on line \( l \); and are the minimum and maximum output for producer \( i, q_{i_{\text{min}}}^l \) and \( q_{i_{\text{max}}}^l \) are the minimum and maximum output for producer \( i \). \( F_{L_{\text{min}}} \) and \( F_{L_{\text{max}}} \) are the lower and upper limit of real power flow on line \( l \); In equation (3), the first and second equality constraint is the total awarded generation quantity for power producer. The third equality is the power balance constraint in the whole system. More details about the above constraints could be found and explained in [20].

Once the SCED model is cleared, the LMP of each power producer can be formulated based on the Kuhn-Tucker condition as follows,

\[
\rho = \eta e - \eta \frac{\partial p_{\text{loss}}}{\partial q_i} + T_i^T (\Upsilon^+ - \Upsilon^-) \tag{4}
\]

where \( \eta \) is the Lagrangian multiplier for the power balance constraint; \( \Upsilon^+ \) and \( \Upsilon^- \) are the forward and reverse Lagrangian multipliers for the line flow constraint.

**III. Solution Method**

In this Section, a forecasting framework consisting of RR, NPCNN, and LQR is first introduced in wind power bidding to evaluate day-ahead price uncertainty. Then, IfA, using the momentum theory and the adaptive learning rate method is proposed to address the cooperative bidding model. At last, the profits of cooperative power producers are distributed by the Shapley value [21].

**A. Uncertainty Predictor**

In this section, a deep learning-based uncertainty evaluator consisting of feature selection, deterministic and probabilistic forecasting is presented to reduce the uncertainty.
1) **Feature Selection:** Ridge regression-recursive feature elimination (Ridge regression) method is used to select the correlate features to the corresponding forecast day for Day ahead Electricity Price, DAEP. First, Ridge model tackles the problem of multi co-linearity by estimating regression coefficients, is used to select the most correlated feature of the pre-selection feature subset. Then, the remaining correlated features are obtained by iterating over the rest of the pre-selection feature subset. The largest correlated parameter in each iteration can be described as follows,

\[
\beta_{\text{optimal}}^k = \max \left\{ \arg \min \left( \| y - \sum_{j=1}^{P} \beta_j^k x_j \|_2^2 + \| I \beta_j^k \|_2^2 \right) \right\}
\]

where \( \beta_j^k \) is the ridge parameter of the pre-selection feature for the \( j \)th feature in the \( k \)th iteration and \( I \) is the identity matrix.

Note that, there are two ways to predict day-ahead electricity prices. One is to build a single NPCNN model for predicting 24-ahead electricity prices through correlate features of one day, the other is to build multiple NPCNN models through correlate features of different time domain (dividing one day into multiple time domain) to predict 24-ahead electricity prices, and make the latter as the method of this work.

B. Non-pooling convolutional neural network based deterministic forecasting

The convolutional neural network (CNN) is a feed-forward neural network that employs convolutional and pooling operations as a hidden layer to feature extraction and information filtering, which efficiently reduces the number of network parameters and feature dimension reduction. A typical CNN consists of an input layer, stacking convolutional and pooling layers, a fully connected layer, and an output layer. Numerous studies have suggested that CNN has been widely applied in research fields, like text recognition, fault location, and pattern recognition.

Moreover, due to the power feature extraction capability, it has also been successful in the field of the power system in recent years. In [26], the authors proposed an improved residual-based deep CNN to forecast wind power. In [27], the authors proposed a novel hybrid model based on a novel modified evolutionary algorithm and CNN for electricity load forecasting. The simulation results also demonstrated that the forecasting performance of NPCNN outperforms other compared forecasting models, in terms of different error evaluation indexes.

In this paper, a NPCNN (without pooling layer) is proposed for the day-ahead electricity price forecasting, due to the number of correlated features for electricity price, such as system loads, weekly cycle, historical electricity prices is too little to provide information filtering and dimension reduction of the feature map.

1) **Input layer:** The input layer is designed to convert the 1-dimensional time series data into 2-dimensional, that can be used as the input of the convolutional layer. The processing is as follows: 1) The size of the convolutional layer \( m \times m \) should be determined by the number of correlated features \( N \) from the feature selection; 2) Putting the first \( m \) correlated features as the first row, the next \( m \) correlated features as the second column. 3) Taking the maximally correlated feature to fill in the gaps of null elements \( m \times m - N \) for the last row of the image.

2) **Convolutional layer:** As the kernel part of the NPCNN, convolutional layer is to extract the inherent feature in its previous layer by using convolutional operation. Each kernel has a small receptive field and corresponds to a full depth of the input volume map.

\[
y^l_i = f \left( \sum_{j=1}^{I_m} x^l_j - 1 + k^l_{ij} + b^l_j \right)
\]

where \( I_m \) is the number of feature map, \( k^l_{ij} \) and \( b^l_j \) are respectively the weight and offset for the \( j \)th input map and the \( j \)th output map corresponding to the \( l \)th convolution layer respectively, \( f \) is a user-defined activation function, and \( * \) represent a convolutional operation.

3) **Linear quantile regression based probabilistic forecasting:** In this work, the uncertainty (probability density) with different confident intervals for DAEP are estimated using linear quantile regression model (LQR) [28]. LQR is formulated to make the linear relationship between the point forecasting results (taken as explanatory variables), and the conditional distribution of the actual values (taken as the outcome variable). This uncertainty for DAEP can therefore be trained, if giving a set of point forecasting results, and is given in equation 7.

\[
y_j = f \left( \sum_{i=1}^{I_m} x_j - 1 + k^l_{ij} + b^l_j \right)
\]

where \( y_j \) and \( x_j \) are the forecasting and real value for sample \( j \) respectively, \( \beta \) is an optimal vector according to \( \tau \) (quantile value).

C. Firefly algorithm

1) **Standard firefly algorithm:** Firefly algorithm (FA) proposed by Yang is a type of numerical computational algorithm with limited function evaluations. It was inspired by the regularity of firefly activity that established a simplified optimization model. In FA, each firefly assumed as the unisex individual is compelled to attract brighter fireflies in each iteration. Besides, the attraction exponentially decreases with increasing distance between the fireflies, and individual moves at random if no fireflies are brighter than the one. Therefore, the movement of FA \( k^l_{ij} \) between firefly \( k^l_i \) and brighter fireflies \( k^l_j \) can be calculated as follows,

\[
k^l_{ij} + 1 = k^l_i + r_0 e^{-r \gamma} (k^l_j - k^l_i) + a \xi
\]

where the second term is attraction intensity with the distance \( r_{ij} \) between \( k^l_i \) and \( k^l_j \), and the third term is randomization by using the normal distribution, \( \xi \) multiplied by a fixed step \( a \), \( r_0 \) is the attractiveness at the distance \( r = 0 \), and \( \gamma \) is the light absorption coefficient. Here, \( r_0 \) and \( \gamma \) are both set to 1.
2) Improved firefly algorithm: It can be seen from the Equation 8 that each firefly of the FA in each iteration is moved by all brighter fireflies’ attraction and randomization. However, there are two problems with this approach: 1) too many attractions may lead to oscillations and slow convergence rate during the search process [29]; 2) fine-tuning in exploitation (randomization) only is easy to be trapped in local optima and converge low accuracy in optimizing multimodal function [30]. Therefore, to efficiently overcome these shortcomings, an improved FA (IFA) using the adaptive moment estimation method is proposed in this paper. In this approach, the momentum parameter is firstly introduced into IFA to reduce oscillations, and the adaptive moment is defined using the estimate of the second moment to make the convergence faster and more stable. Accordingly, the movement rule of IFA can be expressed as follows:

\[
\begin{align*}
    a_{iter}^{iter+1} &= \mu_0 a_{iter} \\
    T_i^{iter+1} &= r_0 e^{-\tau^2} (k_{iter}^{iter} - k_i^{iter}) + a_t \\
    V_i^{iter+1} &= c_1^{iter} V_i^{iter} + (1 - c_1^{iter}) T_i^{iter+1} \\
    W_i^{iter+1} &= c_2^{iter} W_i^{iter} + (1 - c_2^{iter}) (T_i^{iter+1})^2 \\
    k_i^{iter+1} &= k_i^{iter} + \frac{\mu_0}{\sqrt{W_i^{iter+1} + \epsilon}} V_i^{iter+1}
\end{align*}
\]

where \( T_i^{iter+1} \) denotes the current movement; \( V_i^{iter+1} \) denotes the momentum; \( W_i^{iter} \) is the second moment of the movements, and \( c_1^{iter} \) and \( c_2^{iter} \) denote the weights for \( V_i^{iter} \) and \( W_i^{iter} \), respectively. \( \mu_0 \) is the initial learning rate, and \( \mu_{0} / \sqrt{W_i^{iter+1} + \epsilon} \) is the learning rate. From equation 11 and equation 13, it can be seen that \( V_i^{iter+1} \) will be greater, if \( T_i^{iter+1} \) is in the same direction as \( V_i^{iter} \) (\( V_i^{iter+1} \) will be smaller, if their direction are reversed). This means that the momentum helps accelerate FA in the relevant direction and dampens oscillations. From equation 12 and equation 13, it can also be seen that the learning rate would increase the momentum changes so as to jump from local optimal solution, if \( W_i^{iter+1} \) dramatically drops. This means that the adaptive learning rate helps FA to enhance the accuracy convergence.

D. Shapley value

The profit distribution of cooperative power producers is a crucial problem, which directly affects the stability of long-term cooperative suppliers. In this paper, Shapley value \( S_{V_z} \) is used to distribute power producers’ profits [21], as follows:

\[
S_{V_z} = \sum_{\sigma} \frac{|S_c|!(|Z_1| - 1 - |S_c|)!}{|Z_1|!} m_{z}(v)
\]

where \( S_c \) denote a sub-coalition consisting of cooperative power producers. \( m_{z}(v) \) denotes the marginal contribution for power producer [21]. In this paper, firstly, a bi-layer cooperative bidding model for a wind power producer as price-maker is proposed. Then, a forecasting framework consisting of NPCNN and LQR is introduced to reduce the uncertainty. In addition, FA is used to solve the bi-layer bidding model. Ultimately, the profits of cooperative power suppliers are distributed by Shapley value.

IV. EMPIRICAL RESULTS

A. Experimental settings

The modified IEEE 30-bus system serving as a system model is made up of 24 load buses, 41 transmission lines and 6 generators (a wind power producer and 5 conventional power producers), and the grid topology is shown in [31]. Furthermore, a wind power producer with 35MW (G6) rated capacity is fixed at bus-13. The remaining conventional power producers (G1, G2, G3, G4, and G5) are connected at bus 1, bus 2, bus 5, bus 8, and bus 11, respectively, their corresponding capacity are set to 40 MW, 45 MW, 40 MW, 30 MW, and 48 MW. It is obvious that wind power penetration is approximately 14.71% of the total installed generation capacity in the system. This penetration of wind power is a perfect reflection in the real grid, such as most states systems in America. The second-order cost coefficients for conventional power producers, which suppose that the constant terms are set at 100, are individually set to 0.02, 0.0175, 0.0625, 0.00834 and 0.025, and their corresponding first-order cost coefficients are 2, 1.75, 1, 3.25 and 2, respectively. In addition, the cost coefficient and the constant term for wind supplier is set to 3 and 150, respectively.

The proposed deep learning-based ensemble uncertainty evaluator for real-time EPF are evaluated using PJM market data. The train and test data set covers a period from June 2018 to July 2019, in which two test weeks, March 2–8, 2019 denoting low load and July 2–8, 2019 denoting high load, are predicted. Also, one month before the forecast (February 2019 and June 2019) is applied to select the correlated features. In this work, ridge regression-recursive (RR) is applied to construct 24 NPCNN input structures for different time domains. Table 1 gives nine correlation values from four NPCNN models. CNN structure and other supporting information are given in [17]. In order to verify the effectiveness of the deep learning-based ensemble uncertainty evaluator, the forecasting results are compared with the light gradient boosting machine (LGBM) [32], XBGOOST, CABOOST, Random Forest(RF) [33], Decision Tree(DT), BPNN [20], k-nearest neighbor (KNN), and SVR [18].

In IFA, full-coalitions model (which is G1-G6) and 3 well-known benchmark functions (f2-f4) were introduced and analyzed to evaluate the optimization performance. Table I shows a detailed description of f2-f4. f2-f4 are uni-modal, multi-modal, and noisy quadratic function, respectively. Meanwhile, in order to verify the effectiveness of IFA, the optimization results are contrasted with the standard FA, Memetic FA, and FA with chaos, and these optimization algorithm were independently executed 20 times to make a fair comparison.
TABLE I

BENCHMARK FUNCTIONS, WHERE D IS THE DIMENSION

| Name                    | Function                                                                 | Range  |
|-------------------------|--------------------------------------------------------------------------|--------|
| Sphere                  | $f_2 = \sum_{d=1}^{D} x_d^2$                                             | [-100,100] |
| Ackley                  | $f_3 = 20 \times e^{-(\frac{1}{\pi} \sum_{d=1}^{D} x_d^2 + \frac{1}{2} \sum_{d=1}^{D} \cos(2\pi x_d))} + 20 + e$ | [-32,32] |
| Quartic with noise       | $f_4 = \sum_{d=1}^{D} d \times x_d^2 + \text{random}[0,1]$               | [-1.28,1.28] |

Fig. 2. Actual load curve for Mar 2 - 8 and Jul 2 - 8

B. Case studies

In order to analyze the feasibility and efficiency of cooperative bidding result, the overall performance of PUP is firstly evaluated in terms of mean absolute error (MAPE), Average coverage percentage (ACP) and Interval sharpness (IS). Afterward the convergence error and the stability of IFA, in terms of optimization objectives is measured.

1) Performance evaluation of UP: Two test weeks, March 1-7, 2019 denoting low load and July 2-8, 2019 denoting high load, are predicted. Load curve for these two weeks are shown in Fig. 2. Tables II compare the performance of the proposed method with eight contrast forecasting models (KNN, SVR, TREE, RF, XGBOOST, CATBOOST, LGBM, and BPNN) for Feb 26-Mar 4 and Jul 2-7, 2019, and Fig. 3 and Fig 4 show their forecasting and actual results. It can be seen that the daily MAPE value of NPCNN varies from 0.0587 to 0.0988, with an average of 0.0814 and a variance of 8.6E-5, for Feb 26 - Mar 4, and from 0.0458 to 0.0981, with an average of 0.0732 and a variance of 4.8E-4.

The average MAPE results of Feb 26-Mar 4 and Jul 2-7 2019 for eight contrast forecasting models are 0.1401, 0.2600, 0.1207, 0.1236, 0.1202, 0.1183, 0.1239, and 0.1303, respectively. For these two weeks, the average MAPE results and variance of NPCNN are significantly increased. Especially in the week of Jul 2-7, 2019, the average MAPE results of NPCNN improved by 121.72%, 353.31%, 90.44%, 79.64%, 87.70%, 93.44%, and 106.56%, respectively, compared with eight contrast forecasting models. It is clear that the daily MAPE results of NPCNN performs best, followed by CATBOOST, XGBOOST, TREE, RF, LGBM, BPNN, KNN, and SVR. The reason may be that the NPCNN model are well versed in search of change trend through better nonlinear mapping capability.

However, NPCNN based point forecasting for BMP cannot be applied to the optimal strategy of wind supplier directly, because its error cannot be eliminated. Therefore, QR is introduced in this paper to model the uncertainty for BMP. TABLE III shows the constructed prediction interval (PI) with the prediction interval nominal confidence (PINC) 90% by using the proposed and contrast models.

In TABLE III, the ISs are significantly increased under roughly the same average coverage percentage (ACP) for these two weeks. Compared with eight contrast forecasting models, the ISs are 61.63%, 31.30%, 35.29%, 38.73%, 45.42%, 44.11%, 31.82%, and 23.04% for Feb 26-Mar 4 , 2019, respectively, and 47.18%, 32.95, 43.80%, 26.64%, 29.66% 12.07%, 16.70%, and 26.04% for Jul 2-7, 2019, respectively. This numerical result is consistent with the conclusions from the point prediction results. This is because the errors acquired from the deterministic forecast for BMP have fewer uncertainties compared to eight contrast forecasting models. The constructed prediction intervals with confidence level 80% using the proposed model for these two weeks are also presented in Figs. 5, and 6. It can be seen that the real data for BMP are within the established lower and upper bounds for a great percentage.

2) Optimal performance evaluation of IFA: This section aims to evaluate the overall performance of the IFA according to different optimization objectives. Table IV summarizes the computational results of the proposed and contrast optimization algorithms. For all the algorithms, population size and the number of iterations for $f_1$ were set to 30 and 60, respectively, and population size and the number of iterations for $f_2 - f_4$ were set to 30 and 5.0E+03, respectively.

In Table IV, for $f_1$, the averages of FA, MFA, CFA, and IFA are $USD 4997.76$, $USD 4997.87$, $USD 4997.80$, and $USD 4998.02$, respectively, and their corresponding variance is 3.21E-2, 9.2E-3, 2.59E-2, and 1.28E-4, respectively. Compared with FA, MFA, and CFA, the averages of IFA are improved by $USD 0.26$, $USD 0.15$, and $USD 0.22$, respectively, and the variance is reduced by 0.0320, 0.0091, and 0.0258, respectively. It can be seen that the accuracy and stability of IFA is best compared to the benchmarks. Similarly, the experiment results for $f_2 - f_4$ show that IFA can also obtain a stable optimal solution compared with FA, MFA, and CFA. The main reason is that the adaptive moment estimation method with controlling the step size stores the square of the past movement memories to escape from the local optimum.

Fig. 7 presents convergence curves of various algorithms for $f_1 - f_4$ before 60 iterations (considering vision effects). At the beginning stages, IFA shows a slow convergence rate compared to some algorithms. The main reason is that accu-
Fig. 3. Day-ahead electricity forecasting results of different forecasting models for Feb 26-Mar 4, 2019

Fig. 4. Day-ahead electricity forecasting results of different forecasting models for Jul 2-7, 2019

### TABLE II

| Week            | Daily MAPE | KNN  | SVR  | TREE | RF   | XGBOOST | CATBOOST | LGBM  | BPNN  | NPCNN |
|-----------------|------------|------|------|------|------|---------|----------|-------|-------|-------|
| Feb 26–Mar 4, 2019 | Max        | 0.1839 | 0.3435 | 0.1432 | 0.1554 | 0.1512 | 0.1498 | 0.1345 | 0.1391 | 0.081  |
|                 | Min        | 0.0717 | 0.0957 | 0.0855 | 0.0774 | 0.0729 | 0.0603 | 0.0674 | 0.0674 | 0.0597 |
|                 | Average    | 0.1179 | 0.1884 | 0.1028 | 0.1079 | 0.1089 | 0.0993 | 0.1062 | 0.1093 | 0.0685 |
|                 | Variance   | 1.70E-03 | 8.50E-03 | 5.90E-04 | 1.10E-03 | 6.20E-04 | 1.10E-03 | 6.50E-04 | 1.08E-03 | 8.60E-05 |
| July 2-8, 2019  | Max        | 0.2293 | 0.316  | 0.0866 | 0.233  | 0.2366 | 0.225  | 0.2433 | 0.0765 | 0.1091 |
|                 | Min        | 0.0873 | 0.2003 | 0.2497 | 0.1791 | 0.1707 | 0.1694 | 0.1907 | 0.2768 | 0.0458 |
|                 | Average    | 0.1623 | 0.3316 | 0.1385 | 0.1394 | 0.1315 | 0.1374 | 0.1416 | 0.1512 | 0.0732 |
|                 | Variance   | 3.40E-03 | 1.30E-02 | 4.90E-03 | 4.30E-03 | 5.00E-03 | 3.10E-03 | 4.80E-03 | 5.20E-03 | 4.80E-04 |

### TABLE III

| Week           | ACP (%) | IS     |
|----------------|---------|--------|
| Feb 26–Mar 4, 2019 | ACP (%) | IS     |
| ACP (QR)       | 80.95   | -2.19  |
| SVR+QR         | 81.55   | -3.17  |
| TREE+QR        | 81.55   | -2.66  |
| RF+QR          | 81.55   | -2.72  |
| XGBOOST+QR     | 80.95   | -2.86  |
| CATBOOST+QR    | 80.95   | -2.83  |
| LGBM+QR        | 80.95   | -2.92  |
| BPNN+QR        | 80.95   | -2.42  |
| NPCNN+QR       | 80.95   | -1.96  |
| Jul 2-8, 2019  | ACP (%) | IS     |
| ACP (QR)       | 80.36   | -2.56  |
| SVR+QR         | 80.15   | -2.85  |
| TREE+QR        | 80.15   | -2.85  |
| RF+QR          | 80.15   | -2.85  |
| XGBOOST+QR     | 80.15   | -2.85  |
| CATBOOST+QR    | 80.15   | -2.85  |
| LGBM+QR        | 80.15   | -2.85  |
| BPNN+QR        | 80.15   | -2.85  |
| NPCNN+QR       | 80.15   | -2.85  |

### TABLE IV

| Objective Function | FA | MFA | CFA | IFA |
|-------------------|----|-----|-----|-----|
|                   | Mean | Var | Mean | Var | Mean | Var | Mean | Var |
| f1                | 4997.76 | 3.21E-02 | 4997.87 | 9.20E-03 | 4997.8 | 2.59E-02 | 4998.02 | 1.28E-04 |
| f2                | 3.26E-42 | 1.37E-04 | 2.04E-42 | 4.37E-05 | 3.77E-44 | 1.45E-05 | 2.12E-86 | 3.34E-173 |
| f3                | 1.04E-14 | 1.68E-29 | 7.99E-15 | 0 | 7.99E-15 | 0 | 7.99E-15 | 0 |
| f4                | 0.94 | 9.05E-03 | 1.08E-03 | 2.31E-03 | 1.19 | 7.17E-02 | 0.1 | 9.40E-05 |
The profit comparisons with and without coalitions are listed in Table VII. Table 5 shows that power producers (G2, G3, and G5) are marginal units because their strategic bids equal their LMPs, and the output is between the upper and lower limits of power output. Similarly, Table 6 shows that G1 of coalition C1 and G5 of coalition C5 are marginal units. In other words, the cooperative bidding behavior would directly affect market pricing. It should be noted that G6 only changes the LMP without impacting the output, regardless of using bidding behavior with which coalition.

The reason is that the day-ahead LMP of the wind power producer is below the real-time LMP at the equilibrium reached by bidding power producers. Comparisons with and without coalitions of Table 7 show that the extra profits of the wind power producer under coalition are USD 697.7, USD -21.3, USD -407.2, USD 70.8, and USD 28.8, respectively. And the extra profits of its corresponding cooperative power producers are USD -360.0, USD 1691.7, USD 367.8, USD 547.5, and USD 111.6, respectively. This means cooperative members under coalition 4 can both make more profits, most likely make coalitions. So we can conclude that, in general, the wind power producer with coalition can make more profits compared to their bids individually.

| Results                  | G1 | G2  | G3  | G4  | G5  | G6  |
|--------------------------|----|-----|-----|-----|-----|-----|
| Bids (MW/$)              | 20.2 | 31.2 | 30.4 | 23.2 | 50  | 30.5 |
| Outputs (MW)             | 40  | 11.9 | 36.7 | 30  | 31.43 | 35  |
| LMPs (MW/$)              | 30.91 | 31.2 | 30.4 | 30.37 | 50  | 49.2 |
| Costs ($)                | 212 | 123.5 | 221.1 | 205  | 187.6 | 412 |
| Profits ($)              | 1024.6 | 251.1 | 896.41 | 706.1 | 1384 | 1168.1 |

V. Conclusion

In this paper, the cooperative trading model for a price-maker wind power producer considering uncertainty was solved with a data-driven framework. The proposed framework firstly presented an uncertainty predictor that utilized a combination of convolutional neural network, and linear quantile regression to forecast the market price uncertainty, in which ridge regression is used for selecting the correlated features to the corresponding forecast day. Moreover, the framework employed the improved firefly algorithm to optimize the cooperative trading model, as well as the Shapley value to distribute the cooperative profits.

Finally, the effectiveness of the proposed framework and optimization model are comprehensively evaluated in the illustrative examples. The numerical results show that the daily deterministic and probabilistic forecasting results of the proposed uncertainty predictor is more accurate and efficient compared with eight benchmarks. Meanwhile, the improved firefly algorithm is demonstrated to exhibit better stability and robustness compared to the frequently-used benchmarks.

The superiority of the proposed optimization algorithm is due to the techniques that used the momentum theory and the adaptive moment theory. It also demonstrates that the cooperative trading for a wind power producer exhibits higher profits than bidding individually. It is, therefore, apparent that the proposed framework and model have high application value and development potential in energy systems and power market.

References

[1] H. Wang, Z. Lei, X. Zhang, B. Zhou, and J. Peng, “A review of deep learning for renewable energy forecasting,” Energy Conversion and Management, vol. 198, p. 111799, 2019.
Fig. 7. Convergence curves of various algorithms for $f_1$ – $f_4$ before 60 iterations

| TABLE VI |
| BIDDING RESULTS FOR EACH COMBINATION UNDER COALITIONS |
| Results | C1 | C2 | C3 | C4 | C5 |
| C1 | G1 | G6 | G2 | G6 | G3 | G6 | G4 | G6 | G5 | G6 |
| Bids (MWh) | 17.1 | 45.9 | 20.5 | 30 | 19.5 | 25.7 | 17.6 | 21.7 | 37.9 | 18.5 |
| Outputs (MW) | 13 | 35 | 45 | 35 | 40 | 35 | 50 | 35 | 34.9 | 35 |
| LMPs (M$/h) | 11.1 | 66.3 | 47.9 | 49.7 | 43 | 45.2 | 48.6 | 50 | 49.7 | 49.7 |
| Costs ($) | 194.5 | 312 | 214.2 | 312 | 240 | 312 | 205 | 312 | 200.3 | 312 |
| Shapley value ($) | 664.6 | 1865.8 | 1942.8 | 1146.8 | 1264.2 | 760.9 | 1253.6 | 1238.9 | 1495.6 | 1196.9 |

| TABLE VII |
| PROFIT COMPARISONS WITH AND WITHOUT COALITIONS |
| Coalitions | C1 | C2 | C3 | C4 | C5 |
| C1 | G1 | G6 | G2 | G6 | G3 | G6 | G4 | G6 | G5 | G6 |
| Extra profit ($) | -360 | 697.7 | 1691.7 | -21.3 | 367.8 | -407.2 | 547.5 | 70.8 | 111.6 | 28.8 |
| Total extra profit ($) | 337.7 | 1670.4 | -39.4 | 618.3 | 140.4 |

[2] F. Liu, F. Sun, W. Liu, T. Wang, H. Wang, X. Wang, and W. H. Lim, “On wind speed pattern and energy potential in china,” Applied Energy, vol. 236, pp. 867–876, 2019.

[3] L. Baringo and A. J. Conejo, “Strategic offering for a wind power producer,” IEEE Transactions on Power Systems, vol. 28, no. 4, pp. 4645–4654, 2013.

[4] Y. Xiao, X. Wang, X. Wang, C. Dang, and M. Lu, “Behavior analysis of wind power producer in electricity market,” Applied Energy, vol. 171, pp. 325–335, 2016.

[5] S. Shafiee, H. Zareipour, and A. M. Knight, “Developing bidding and offering curves of a price-maker energy storage facility based on robust optimization,” IEEE Transactions on Energy Storage, vol. 10, no. 1, pp. 650–660, 2017.

[6] J. Arteaga and H. Zareipour, “A price-maker/price-taker model for the operation of battery storage systems in electricity markets,” IEEE Transactions on Smart Grid, vol. 10, no. 6, pp. 6912–6920, 2019.

[7] S. Li and C. S. Park, “Wind power bidding strategy in the short-term electricity market,” Energy Economics, vol. 75, pp. 336–344, 2018.

[8] Z. Qian, Y. Pei, H. Zareipour, and N. Chen, “A review and discussion of decomposition-based hybrid models for wind energy forecasting applications,” Applied energy, vol. 235, pp. 939–953, 2019.

[9] J. P. Catalao, H. M. Pousinho, and V. M. Mendes, “Optimal offering strategies for wind power producers considering uncertainty and risk,” IEEE Systems Journal, vol. 6, no. 2, pp. 270–277, 2011.

[10] L. Cheng, H. Zang, T. Ding, R. Sun, M. Wang, Z. Wei, and G. Sun, “Ensemble recurrent neural network based probabilistic wind speed forecasting approach,” Energies, vol. 11, no. 8, p. 1958, 2018.

[11] R. Zhang, G. Li, and Z. Ma, “A deep learning based hybrid architecture for day-ahead electricity price forecasting,” IEEE Access, vol. 6, pp. 143423–143430, 2020.

[12] H.-z. Wang, G.-q. Li, G.-b. Wang, J.-c. Peng, H. Jiang, and Y.-t. Liu, “Deep learning based ensemble approach for probabilistic wind power forecasting,” Applied energy, vol. 188, pp. 56–70, 2017.

[13] J. C. do Prado and W. Qiao, “A stochastic decision-making model for an electricity retailer with intermittent renewable energy and short-term demand response,” IEEE Transactions on Smart Grid, vol. 10, no. 3, pp. 2581–2592, 2018.

[14] S. A. Malik, T. M. Gondal, S. Ahmad, M. Adil, and R. Qureshi, “Towards optimization approaches in smart grid a review,” in 2019 2nd International Conference on Computing, Mathematics and Engineering Technologies (iCoMET) IEEE, 2019, pp. 1–7.

[15] S. Ahmed, T. M. Gondal, M. Adil, S. A. Malik, and R. Qureshi, “A survey on communication technologies in smart grid,” in 2019 IEEE PES GTD Grand International Conference and Exposition Asia (GTD Asia). IEEE, 2019, pp. 7–12.

[16] H. Ding, Z. Hu, and Y. Song, “Rolling optimization of wind farm and energy storage system in electricity markets,” IEEE Transactions on Power Systems, vol. 30, no. 5, pp. 2676–2684, 2014.

[17] A. A. S. de la Nieta, J. Contreras, J. I. Muño, and M. O’Malley, “Modeling the impact of a wind power producer as a price-maker,” IEEE Transactions on Power Systems, vol. 29, no. 6, pp. 2723–2732, 2014.

[18] M. Zugno, J. M. Morales, P. Pinson, and H. Madsen, “Pool strategy of a price-maker wind power producer,” IEEE Transactions on Power Systems, vol. 28, no. 3, pp. 3440–3450, 2013.

[19] T. Dai and W. Qiao, “Optimal bidding strategy of a strategic wind power producer in the short-term market,” IEEE Transactions on Sustainable Energy, vol. 6, no. 3, pp. 707–719, 2015.

[20] J. Wang, Z. Zhou, and A. Botterud, “An evolutionary game approach to analyzing bidding strategies in electricity markets with elastic demand,” Energy, vol. 36, no. 5, pp. 3459–3467, 2011.

[21] A. D. Yuekaya, J. Valenzuela, and G. Dozier, “Strategic bidding in electricity markets using particle swarm optimization,” Electric Power Systems Research, vol. 79, no. 2, pp. 335–345, 2009.

[22] M. A. Soliman, H. M. Hasanien, H. Z. Azazi, E. E. El-Kholy, and S. A. Mahmoud, “Hybrid anfis-ga-based control scheme for performance...
enhancement of a grid-connected wind generator,” *IET Renewable Power Generation*, vol. 12, no. 7, pp. 832–843, 2018.

[23] T. Li and M. Shahidehpour, “Strategic bidding of transmission-constrained gencos with incomplete information,” *IEEE Transactions on Power Systems*, vol. 20, no. 1, pp. 437–447, 2005.

[24] R. Zhang, S. Aziz, M. U. Farooq, K. N. Hasan, N. Mohammed, S. Ahmad, and N. Ibadah, “A wind energy supplier bidding strategy using combined ega-inspired hpsoifa optimizer and deep learning predictor,” *Energies*, vol. 14, no. 11, 2021. [Online]. Available: https://www.mdpi.com/1996-1073/14/11/3059

[25] A. Stern and A. Tettenhorst, “Hodge decomposition and the shapley value of a cooperative game,” *Games and Economic Behavior*, vol. 113, pp. 186–198, 2019.

[26] M. S. Kiran and M. Gunduz, “A recombination-based hybridization of particle swarm optimization and artificial bee colony algorithm for continuous optimization problems,” *Applied Soft Computing*, vol. 13, no. 4, pp. 2188–2203, 2013.

[27] X.-S. Yang and X. He, “Firefly algorithm: recent advances and applications,” *International Journal of Swarm Intelligence*, vol. 1, no. 1, pp. 36–50, 2013.

[28] W. Zhang, H. Quan, and D. Srinivasan, “An improved quantile regression neural network for probabilistic load forecasting,” *IEEE Transactions on Smart Grid*, vol. 10, no. 4, pp. 4425–4434, 2018.

[29] K. Chandrasekaran and S. P. Simon, “Optimal deviation based firefly algorithm tuned fuzzy design for multi-objective ucp,” *IEEE Transactions on Power Systems*, vol. 28, no. 1, pp. 460–471, 2012.

[30] I. B. Aydilek, “A hybrid firefly and particle swarm optimization algorithm for computationally expensive numerical problems,” *Applied Soft Computing*, vol. 66, pp. 232–249, 2018.

[31] N. K. Yong and N. Awang, “Quantile regression for analysing pm10 concentrations in petaling jaya,” *Malaysian Journal of Fundamental and Applied Sciences*, vol. 13, no. 2, pp. 86–90, 2017.

[32] Y. Wang, J. Chen, X. Chen, X. Zeng, Y. Kong, S. Sun, Y. Guo, and Y. Liu, “Short-term load forecasting for industrial customers based on tcn-lightgbm,” *IEEE Transactions on Power Systems*, vol. 36, no. 3, pp. 1984–1997, 2020.

[33] H. Aprillia, H.-T. Yang, and C.-M. Huang, “Statistical load forecasting using optimal quantile regression random forest and risk assessment index,” *IEEE Transactions on Smart Grid*, vol. 12, no. 2, pp. 1467–1480, 2020.