IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received April 21, 2021, accepted April 26, 2021, date of publication April 29, 2021, date of current version May 10, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3076679

Data-Driven Bidding Strategy for DER Aggregator Based on Gated Recurrent Unit–Enhanced Learning Particle Swarm Optimization

HYUNG JOON KIM, HYUN JOON KANG, AND MUN KYEOM KIM
Department of Energy System Engineering, Chung-Ang University, Seoul 06974, Republic of Korea

Corresponding author: Mun Kyeom Kim (mkim@cau.ac.kr)

This work was supported in part by the National Research Foundation of Korea (NRF) under Grant 2020R1A2C1004743, and in part by the Chung-Ang University Research Scholarship Grant 2021.

ABSTRACT Distributed energy resources (DERs) such as wind turbines (WTs), photovoltaics (PVs), energy storage systems (ESSs), local loads, and demand response (DR) are highly valued for environmental protection. However, their volatility poses several risks to the DER aggregator while formulating a profitable strategy for bidding in the day-ahead power market. This study proposes a data-driven bidding strategy framework for a DER aggregator confronted with various uncertainties. First, a data-driven forecasting model involving gated recurrent unit–enhanced learning particle swarm optimization (GRU-ELPSO) with improved mutual information (IMI) is employed to model renewables and local loads. It is critical for a DER aggregator to accurately estimate these components before bidding in the day-ahead power market. This aids in reducing the penalty costs of forecasting errors. Second, an optimal bidding strategy that is based on the information gap decision theory (IGDT) is formulated to address market price uncertainty. The DER aggregator is assumed to be risk-averse (RA) or risk-seeker (RS), and the corresponding bidding strategies are formulated according to the risk preferences thereof. Then, an hourly bidding profile is created for the DER aggregator to bid successfully in the day-ahead power market. The proposed data-driven bidding framework is evaluated using an illustrative system wherein a dataset is obtained from the PJM market. The results reveal the effectiveness of handling uncertainty by providing accurate forecasting results. In addition, the DER aggregator can bid effectively in the day-ahead power market according to its preference for robustness or high profit, with a suitable bidding profile.

INDEX TERMS Bidding strategy, DER aggregator, gated recurrent unit–enhanced learning particle swarm optimization, information gap decision theory, uncertainty.

NOMENCLATURE

A. SETS

b/N^b	Set / maximum number of PSO iterations
z/N^z	Set / maximum number of particles
p/N^p	Set / maximum number of training data
q/N^q	Set / maximum number of validation data
r/N^r	Set / maximum number of PVs
u/N^u	Set / maximum number of WTs
y/N^y	Set / maximum number of DR customer
u/N^u	Set / maximum number of uncertainties
n	Set of particles dimension
k	Set of hidden layers
L	Column length

The associate editor coordinating the review of this manuscript and approving it for publication was Akshay Kumar Saha.

B. PARAMETERS

$E_{PV}(x, 24)$	PV data of 24 h of x$^{\text{th}}$ day
z_t, r_t	Update and reset gate
h_t, $h^{-}t$	Hidden and candidate hidden state
x_t, y_t	Input and output of the time-step t
σ	Sigmoid function
tanh	Hyperbolic tangent function
W_{xz}, W_{xr}, W_{xh}	Update, reset gate, and candidate hidden state weight matrixes
W_{hz}, W_{hr}	Circular connection weight matrixes
b_z, b_r, b_{h-}	Corresponding bias vector
σ	Sigmoid function
$tanh$	Hyperbolic tangent function
w_P	Inertial weight of ELPSO
h_k	Number of neurons in Hidden layer k

n_{GRU}	Iteration number of GRU
y_p, y_q	Actual and predicted values of training data
y'_p, y'_q	Actual and predicted values of validation data
$E_i, E_j^t, E^m l$	Input random, target, and average variable
$\overline{P^{PV}}, \overline{P^{WT}}$	Maximum power of PV and WT
$\overline{P^{ESS}}$	Maximum power of ESS
η_{ch}, η_{dch}	Charging and discharging efficiency of ESS
$\overline{P^{ESS}_{ch,t}}, P^{ESS}_{dch,t}$	Maximum charging and discharging power of ESS
R_{RN}	Risk-neutral profit
π_1	Penalty coefficients of WTs
π_2	Penalty coefficients of PVs
π_3	Penalty coefficients of local load
f^{min}, f^{max}	Minimum and maximum proportion of estimated baseline load
δ	Profit deviation factor
α, β	Robust and opportunity index
$y_i^{real}, y_i^{pred}, y_i^{mean}$	Real, predicted, and mean values of i variable

C. VARIABLES

E_i, E_j^t, E_l^m	Input discrete random variable, target value, and second target value
S_z	Supplementary variable
$vz, n(b), x_{z,n}(b)$	Velocity and position of particle
w_p	Inertia weight
$c_{1,z}(b), c_2(b)$	Self-cognition and influence coefficient
r_1, r_2	Random variables
$p_{best,z,n}(b)$	Local bests solution of PSO
$g_{best,z,n}(b)$	Global best solution of PSO
R_{Agg}	Revenue of DER aggregator
R_t^{sell}	Revenue from selling power to customers
$R^{DR}t$	Revenue from selling DR
C_t^{PM}	Cost of purchasing power from market
C_t^{DR}	Incentives provided to customers
$C_t^{P,WTs}$	Penalty cost of forecasting errors of WTs
$C_t^{P,PVs}$	Penalty cost of forecasting errors of PVs
$C_t^{P,Base}$	Penalty cost of error in local load forecasting
P^{PV_t}, P_t^{WT}	Power of PV and WT
P^{PM_t}, P_t^{DR}	Energy from power market and DR
P_t^{Base}	Baseline load
P_t^{ESS}	Power of ESS
$P_{t,}^{R,PV}$	Real output power of PVs
$P_t^{R,WT}$	Real output power of WTs
$P_t^{R,Base}$	Real power of baseline load
X_t	State-of-charge of ESS
U_{ch}, U_{dch}	Charging and discharging states of ESS
λ_t^{sell}	Retail price
λ_t^{PDR}	Price at which DR is sold to power market
λ_t^{PM}	Price of day-ahead power market
λ_t^{DR}	Price of incentive payment
λ_t^{RR}	Price of reserve requirement
q	State variable
$G(q) = b$	Equality constraints
$(Hq, x) \leq 0$	Inequality constraints
x	Decision variables
R_{RN}	Revenue of RN strategy
$\lambda_t^{PM}, \lambda_t^{PM*}$	Predicted and actual values of market price

D. ABBREVIATIONS

ABC	Artificial colony bee (algorithm)
CA	Correlation analysis
CPU	Computation time
DERs	Distributed energy resources
EGA	Enhanced genetic algorithm
ELPSO	Enhanced learning PSO
ESS	Energy storage system
GRU	Gated recurrent unit
IDR	Incentive-based demand response
IGDT	Information gap decision theory
IMI	Improved mutual information
LSTM	Long short-term memory
MAE	Mean absolute error
PSO	Particle swarm optimization (algorithm)
PCA	Principle component analysis
PVs	Photovoltaics
RA	Risk-averse
RMSE	Root mean squared error
RO	Robust optimization
RS	Risk-seeker
SE-PSO	Security enhanced-PSO
SM	Spider monkey (algorithm)
SO	Stochastic optimization
SOC	Stage-of-charge of ESS
SVR	Support vector regression
WTs	Wind turbines

I. INTRODUCTION

A. MOTIVATION

The development of distributed energy resources (DERs), which include photovoltaics (PVs), wind turbines (WTs), energy storage systems (ESSs), and demand response (DR), is considered a potential solution and the penetration of DERs is likely to increase owing to environmental policies [1]. The DER aggregator, which aggregates a substantial number of various small-sized renewable sources, has been used to manage the task of fulfilling the expected local load through bidding in the day-ahead power market and scheduling various DERs [2]. As a profit-seeking entity, the DER aggregator is expected to design a day-ahead optimal bidding and scheduling strategy to maximize profits. However, there are factors that could significantly impact the revenue of the aggregator. These include the uncertainties of renewable output power, local load, and day-ahead power market price. Accurate modeling of these factors is crucial for the economic performance of the DER aggregator when it participates in power markets and the accuracy of modeling significantly affects its bidding strategy. Thus, there exists an evident need to develop a forecasting model for predicting the behavior of aggregator components prior to bidding in the power market, in order to effectively handle uncertainties. In addition, an optimal bidding strategy should be determined to effectively procure energy by addressing price uncertainty. In other words, the DER aggregator should prepare various bidding strategies by creating a bidding profile to purchase energy from the power market and satisfy the demand in an economical and robust manner. Therefore, these uncertainties must be considered while developing an optimal bidding strategy for a DER aggregator.

B. LITERATRUE REVIEW

Several studies have focused on managing the uncertainties in the day-ahead bidding process of a DER aggregator. Addressing uncertainty modeling method can be categorized into five classes: stochastic optimization (SO), robust optimization (RO), interval optimization, fuzzy method, and autoregressive moving average (ARIMA) models. In [3], a stochastic-based comprehensive bidding strategy for DER aggregators was developed considering PV and WT output uncertainties. The authors in [4] proposed a stochastic linear programming-based optimal bidding strategy for DER aggregators in a day-ahead market that involves DER uncertainties. In [5], an optimal bidding strategy for renewable-based aggregators in the real-time market was proposed. It is based on bi-level optimization and aims at solving the decision-making problem of DER aggregators and the real-time market clearing problem as a price-maker players. An optimal dispatch problem of a virtual power plant was proposed in [6] where uncertainties of the renewables and load were considered via the scenario-based optimization method. A Stakelberg game theory model was applied in this study to maximize the profit of the virtual power plant while minimizing the cost of purchasing energy from users. A stochastic-based conditional value at risk (CVaR) was also used to manage the uncertainties of renewables in the bidding strategy problem. Because the accurate estimation of the probability distribution function of uncertain factors is problematic in practice, the results obtained through SO in certain cases are unreliable. Unlike SO, RO has been discussed widely which considers the worst-case of uncertainty. A non-cooperative static game-based robust bidding strategy of a DER aggregator was introduced in [7], where the uncertainty of the day-ahead market price was modeled by the robust method. The authors in [8] proposed a hybrid bidding strategy based on dual stochastic/robust optimization. Here, the uncertain output power of renewables and the day-ahead market price were modeled via scenario-based optimization. In addition, RO was used to handle the real-time market price uncertainty. As the RO addresses the worst cases, its results may be overly conservative, as compared with those of SO. In addition, the scope of uncertain parameters needs to be defined in advance, which is a drawback of this optimization method. Interval optimization can also be applied to address the uncertainties where uncertain parameters are represented as interval numbers [9]. A coordinated bidding strategy using interval optimization for the hydro units and wind farms of a generating company was proposed in [10]. The uncertainties of WT power, energy, and intraday energy price were represented by the number of intervals ratherthan probability distributions. In Ref. [11], a three-stage hybrid stochastic/interval optimization was proposed for application in the bidding strategy of microgrids, considering the uncertainties of renewables, loads, and market prices. An economical and robust solution with significantly less computational complexity was provided by adopting hybrid optimization. Notably, interval optimization involves less computational complexity than SO. However, the results of interval optimization are less precise than those of the robust method, because the worst case is considered in the robust method. The fuzzy method is similar to SO in that it adopts a fuzzy membership function to manage uncertainties. A coordinated bidding strategy across the regulation and spinning reserve market was proposed in [12]. Herein, the fuzzy set theory was implemented to model the uncertainties of ancillary service prices. In [13], an optimal DR bidding and pricing mechanism for a virtual power plant was proposed considering renewable output uncertainties under fuzzy optimization. Similar to the SO method, the acquisition of a fuzzy membership function can be difficult. Furthermore, it may be unsuitable for addressing exceptional and highly uncertain circumstances. The authors in [9] presented the optimal bidding of a DER aggregator in the day-ahead frequency regulation market. In their study, seasonal ARIMA was utilized to model market price uncertainty. However, Vatandoust *et al.* [14] revealed that the seasonal ARIMA model could not efficiently manage uncertainties and that it results in inaccurate bids for DER aggregators.

To overcome these shortcomings, new approaches using artificial intelligence have garnered significant attention. Recently, data-driven models based on machine learning and artificial intelligence have been applied in the field of power systems [15]. Numerous uncertain parameters are encountered while formulating the optimal bidding strategy for a DER aggregator. Furthermore, the large volume of data necessitates a strong data-driven forecasting model for high-precision modeling. In general, time-series forecasting models such as long short-term memory (LSTM) or the gated recurrent unit (GRU) have been adopted in state-of-the-art studies for short-term load or renewable output forecasting tasks [16], [17]. In [18], a state-of-charge battery ESS prediction model was proposed by combining LSTM and an unscented Kalman filter. The authors in [19] proposed a data-driven model for handling the EV demand

TABLE 1. Comparison analysis between existing literatures and proposed framework.

References	Uncertainties handled				Problem constraints		Uncertainty modeling method							Player type	
	WTs	PVs	Local load	Price	DR	ESS	Stochastic	Robust	Interval	Fuzzy	ARIMA	Data-driven	IGDT	Price maker	Price taker
[3]	✓	✓	✓			✓	✓								✓
[4]	✓	✓		✓		✓	✓								✓
[5]	✓	✓					✓							✓	
[6]		✓	✓		✓	✓	✓								✓
[7]			✓		✓			✓							✓
[8]	✓	✓	✓			✓	✓	✓							✓
[10]	✓			✓					✓					✓	
[11]	✓	✓	✓	✓	✓	✓		✓	✓						✓
[12]				✓		✓				✓	✓				✓
[13]	✓	✓			✓					✓					✓
[14]				✓		✓	✓				✓				✓
[20]			✓		✓	✓						✓			✓
[21]			✓		✓								✓		✓
[22]			✓	✓	✓	✓							✓		✓
[23]			✓	✓	✓								✓		✓
[24]	✓	✓	✓	✓	✓	✓	✓						✓		✓
Proposed framework	✓	✓	✓	✓	✓	✓						✓	✓		✓

uncertainty. Their results established that the data-driven model can provide a superior framework for handling uncertainty owing to its highly effective memory unit in the network. This would enable a DER aggregator to improve its financial performance while participating in various markets, as compared with other optimization techniques. The information gap decision theory (IGDT) has also emerged as an effective optimization method to model uncertainties without probability distribution functions that are based on historical data or the fluctuation ranges of uncertain variables [20]. When modeling uncertain variables, the IGDT encloses uncertainties as an unbounded gap in which the negative impact of uncertainties can be managed comprehensively. More precisely, the IGDT provides a robust solution that fulfills the specified profit/cost expectations of the operator. In [21], a risk-managing bidding strategy for microgrid operators was presented, considering the impact of load and price uncertainties caused by the IGDT. The authors in [22] proposed IGDT-based bi-level programming for short-term self-scheduling by DR aggregators. A two-level stochastic/IGDT optimization framework was proposed in [23] to address the uncertainties of PVs, WTs, energy, and reserve market prices; probability of calling reserve; and load in the energy and reserve market. In this model, CVaR and IGDT risk-aversion parameters were applied to manage the effects of uncertainties in the bidding problem. It was established that the framework can provide various risk-based strategies to address uncertainties. However, optimization by applying the IGDT alone could not consider the penalty cost of forecasting errors. This yields the actual difference in profit for a DER aggregator when participating in the day-ahead power market.

Table 1 summarizes the comprehensiveness and novelty of this study. It demonstrates the difference between the methods reported in literature and the proposed bidding framework.

C. CONTRIBUTIONS AND PAPER ORGANIZATIONS

This paper proposes a new data-driven bidding framework for DER aggregators that is optimized using the IGDT by considering the uncertainties of PVs, WTs, local loads, and day-ahead market prices. The adopted bidding strategy is superior to other frameworks because the proposed forecasting model yields more accurate day-ahead prediction results through an elaborate modeling of the uncertainty variables via the gated recurrent unit–enhanced learning particle swarm optimization algorithm (GRU-ELPSO) in conjunction with improved mutual information (IMI). In addition, the IGDT-based optimization is used to provide risk-averse (RA) and risk-seeker (RS) bidding strategies for a DER aggregator and submit day-ahead bids to the power market according to its preferences for robustness or high profit. Accordingly, a data-driven bidding strategy based on the IGDT functions as a practical tool for handling uncertainties from the perspective of engineering applications.

The major contributions of this study can be listed as follows:

- The development of an optimal data-driven bidding strategy for DER aggregators to efficiently handle the uncertainties posed by DERs, local loads, and market prices while participating in the day-ahead power market.
- The validity and performance of the proposed data-driven bidding strategy for DER aggregators were

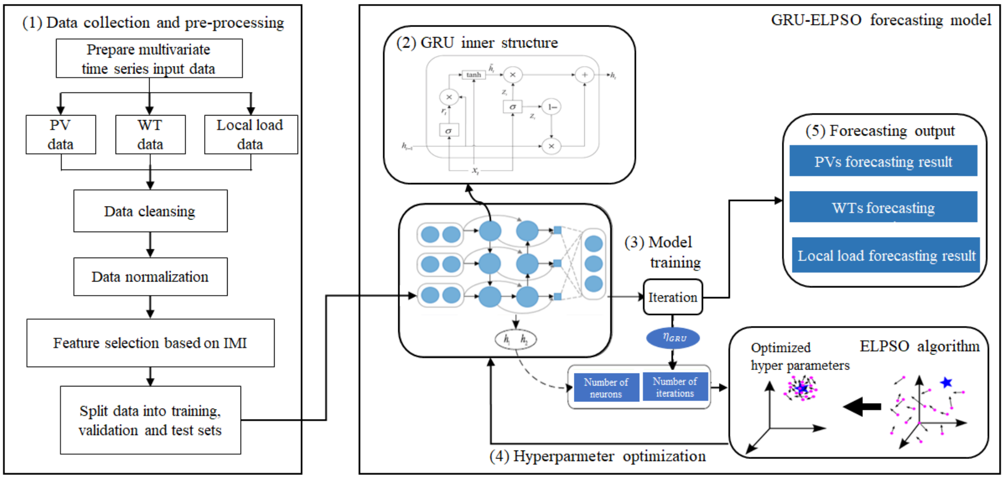


FIGURE 1. Overall structure of proposed data-driven model.

investigated using an illustrative system, where specific data were obtained from the PJM market for considering a realistic environment.

- A novel data-driven forecasting model composed of IMI and GRU-ELPSO is proposed. The IMI functions systematically to rank the candidate input data and extract the best input dataset by analyzing both linear and nonlinear characteristics. To improve the forecasting accuracy and prevent the GRU overfitting problem, an ELPSO is utilized to optimize the hyperparameter sets.
- The IGDT-based optimization method offers both RA and RS strategies. The acquisition of both strategies under varying degrees of market price uncertainty and the risk level that the DER aggregator can endure further demonstrates the superiority of the IGDT method.
- The development of bidding strategy curves according to the aggregator's risk level provides suitable decision-making criteria to be submitted for the day-ahead power market and procure energy at the most reasonable price to maximize profit.

The remainder of this paper is organized as follows: Section II presents the proposed data-driven forecasting model with IMI feature selection. In Section III, an IGDT-based problem formulation for an optimal bidding strategy is presented. The proposed optimal bidding strategy framework is presented in Section IV. Case studies with detailed analyses are provided in Section V. Finally, Section VI concludes the paper.

II. DATA-DRIVEN MODEL FOR HANDLING UNCERTAINTIES

In our work, a data-driven forecasting model, GRU-ELPSO, is proposed to model the various stochastic parameter uncertainties of aggregated renewables and local load. The overall structure of the proposed model for predicting stochastic parameters is shown in Fig. 1. It consists of five parts: (1) collection of historical multivariate time-series data and pre-processing with feature selection, (2) construction of the GRU, (3) GRU training and validation, (4) identification of an optimal GRU hyperparameter by implementing the ELPSO algorithm, and (5) prediction of PVs, WTs, and the local load of the DER aggregator via an individual optimized forecasting model. Based on the proposed data-driven model, the DER aggregator can effectively forecast the aggregated WTs, PVs, and local loads with high accuracy before bidding in the day-ahead power market. In the following section, a detailed description of the proposed forecasting model is presented with mathematical formulations.

A. DATA-PREPROCESSING AND FEATURE SELECTION

Let us assume that E_{PV} is a historical PV datum, which is described in the matrix form. The hourly PV output power data are fed into the pre-processing and feature selection steps:

$$E_{PV} = \begin{bmatrix} E_{PV}(1,\ 1) & E_{PV}(2,\ 1) & \cdots & E_{PV}(x,\ 1) \\ E_{PV}(1,\ 2) & E_{PV}(2,\ 2) & \cdots & E_{PV}(x,\ 2) \\ \vdots & \vdots & \vdots & \vdots \\ E_{PV}(1,\ 24) & E_{PV}(1,\ 24) & \cdots & E_{PV}(x,\ 24) \end{bmatrix} \quad (1)$$

These historical PV data are first pass to the data-cleansing step. Here, missing and defective data are replaced by the mean value of the PV data from preceding days. Then, to ensure that the overall weighted sum lies within the activation function limit, the cleansed PV data undergo the normalization process. This is because the data can include outliers, and the parameter weight matrix is highly marginal. Subsequently, the feature selection process is implemented in machine learning to filter out unimportant or less important input features. The feature selection step is significantly important for prediction accuracy. Furthermore, it prevents overfitting because it aids in circumventing the dimensionality curse [25]. In this regard, the selection of optimal input variables from large dimensions is important. Numerous feature selection techniques exist, such as correlation analysis (CA), principal component analysis (PCA), and mutual information (MI). In this study, an improved MI (IMI) feature selection technique is developed by modifying the conventional entropy-based MI technique [26] to analyze both linear and nonlinear dataset characteristics. This technique ranks inputs according to their information values. The subset of selected features includes the best and most relevant data, which can contribute significantly toward the forecasting accuracy. The proposed IMI technique can be described mathematically as follows:

$$IMI(E, E^t, E^m) = \sum_i \sum_j \sum_l p(E_i,\, E_j^t,\, E_l^m) \times \log_2\left(\frac{p(E_i,\, E_j^t,\, E_l^m)}{p(E_i) \times p(E_j^t) \times p(E_l^m)}\right) \quad \forall E_i,\, E_j^t,\, E_l^m \in \{0, 1\} \tag{2}$$

where $p(E_i,\, E_j^t,\, E_l^m)$ is the joint probability of the three discrete random variables.

In the conventional MI technique, where $MI(E, E_j^t)$ is used, the last sample data among the training sample is tend to be chosen as the target value which is close to the next day. Although this appears logical with respect to time, it could lower the forecasting accuracy owing to average behavior ignorance. Conversely, in the IMI technique, the average value E_l^m is included in addition to the target value, in order to increase prediction accuracy. To obtain the joint and individual probabilities, S_z can be described as

$$S_z = 4E_j^t + 2E_l^m + E_i \tag{3}$$

In (3), it is clear that $S_z \in \{0, 1, 2 \ldots, 7\}$, and S_z, counts the numbers from zero to seven. From the aforementioned discussion, the individual and joint probabilities can be computed as (4). This feature selection technique offers two advantages: (i) it minimizes prediction errors by selecting reasonable and relevant input data, and (ii) it improves the convergence rate by selecting a suitable feature subset. After the IMI feature selection step, the selected input data are split into training, validation, and test data samples for the forecasting model. Finally, the selected key subset features are used as an input for the proposed GRU-ELPSO forecasting model.

$$\begin{aligned}
pr(E=0) &= \frac{S_{0m} + S_{2m} + S_{4m} + S_{6m}}{L}\\
pr(E=1) &= \frac{S_{1m} + S_{3m} + S_{5m} + S_{7m}}{L}\\
pr(E^m=0) &= \frac{S_{0m} + S_{1m} + S_{4m} + S_{5m}}{L}\\
pr(E^m=1) &= \frac{S_{2m} + S_{3m} + S_{6m} + S_{7m}}{L}\\
pr(E^t=0) &= \frac{S_{0m} + S_{1m} + S_{2m} + S_{3m}}{L}\\
pr(E^t=1) &= \frac{S_{4m} + S_{5m} + S_{6m} + S_{7m}}{L}
\end{aligned} \tag{4}$$

B. GATED RECURRENT UNIT

A recurrent neural network (RNN) is a forecasting model that contains a previous information status and passes it to the next time-step where subsequent inputs would be processed. RNNs can be employed in WTs, PVs, and local load forecasting tasks, assuming that these stochastic parameters are time dependent. However, the implementation of a general RNN may be hindered by exploding or vanishing gradient problems [27]. With the development of deep learning technology, LSTM and GRU have been introduced as effective solutions to address these problems. An LSTM is an improved forecasting model based on a simple RNN. It adds "cell state" and "processor" throughout the time series to decide whether the information is required, by using input, forget, and output gates [28]. Although LSTM addressed the gradient exploding and vanishing problems, it encountered these problems in certain cases. The GRU, which is based on the LSTM, is also used to overcome training problems and maintain the internal state throughout a recurrent process. In this study, GRU is adopted as the surrogate emulator in our simulation. This is because it reduces the computational burden and realizes a faster learning curve owing to its more compact structure and fewer parameters, as compared with the RNN and LSTM [29].

The GRU has two gates ("reset gate" and "update gate") as the specific inner GRU structure (see Fig. 1). The forward transmission process can be presented as [30]

$$z_t = \sigma(W_{xz}x_t + W_{hz}h_{t-1} + b_z) \tag{5}$$

$$r_t = \sigma(W_{xr}x_t + W_{hz}h_{t-1} + b_r) \tag{6}$$

$$\overline{h} = \tanh(W_{x\overline{h}}x_t + W_{h\overline{h}}(r_t \odot h_{t-1}) + b_{\overline{h}}) \tag{7}$$

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \overline{h_t} \tag{8}$$

$$y_t = \sigma(W_o \cdot h_t) \tag{9}$$

In (7) and (8), $\circ$ indicates element-wise multiplication. According to (5)–(9), the GRU reset gate determines the method for combining the current input information status with the preorder memory in order to manage the degree to which the state at a past moment is omitted. The higher the reset gate value, the shallower is the degree of omission. The update gate is used to determine the amount of previous information stored in the current time-step. It is implemented to control the extent to which the state memory of the preceding state is processed into the current state. The update gate value affects the extent to which the state memory of the preceding moment is maintained for the current time-step. The two gated vectors determine the output status of the gated loop. In addition, the gate mechanism can maintain information in a long-term sequence without loss over time or being eliminated, although it can be irrelevant to the prediction.

C. ENHANCED LEARNING PARTICLE SWARM OPTIMIZATION

The GRU forecasting model based on a deep neural network is developed with hidden layers. In particular, the number of hidden layers, neurons, and iterations are critical hyperparameters that directly affect the GRU forecasting accuracy. Insufficient or excessive numbers can result in underfitting or overfitting, respectively. Grid and random searches are typical algorithms for identifying these hyperparameters. However, these may involve range limit and bias [31]. The particle swarm optimization (PSO) algorithm is used to overcome these disadvantages and search for suitable hyperparameters. It can search for its own optimal and globally optimized solution in the design space with a satisfactory search time and computational burden. PSO [32] is a meta-heuristic optimization algorithm based on the swarm intelligence of particles that share their explorations among themselves. It is assumed that the z-th particle in the N-dimensional space $x_{P,z} = (x_{P,z1}, x_{P,z2}, \cdots, x_{P,zN})$ has a position and velocity vector. These can be expressed as $p_{P,z} = (p_{P,z1}, p_{P,z2}, \cdots, p_{P,zN})$ and $v_{P,z} = (v_{P,z1}, v_{P,z2}, \cdots, v_{P,zN})$, respectively. The local and global optimal solutions for the b iteration can be denoted as $p_{P,best,z,n}(b)$ and $g_{P,best,z,n}(b)$, respectively. The particle velocity (10) and position (11) can be updated according to $p_{P,best,z,n}(b)$ and $g_{P,best,z,n}(b)$, as follows [32]:

$$\begin{aligned} v_{z,n}(b+1) = w \times v_{z,n}(b) + c_1 \times r_1 \times (p_{best,z,n}(b) - x_{z,n}(b)) \\ + c_2 \times r_2 \times (g_{best,z,n}(b) - x_{z,n}(b)) \end{aligned} \tag{10}$$

$$x_{z,n}(b+1) = x_{z,n}(b) + v_{z,n}(b+1) \tag{11}$$

Improvements are necessary to identify optimized GRU hyperparameters. This is because a conventional PSO converges prematurely or tends to fall into the local optima when the dimension of the optimization problem increases. A new PSO variant (namely, ELPSO) is proposed to solve these problems and thereby guarantee the search speed and ensure the global optima, while maintaining the simplicity of the PSO algorithm after several modifications.

- *Inertia Weight:* This parameter can be employed to balance the local and global search accuracy. A larger weight facilitates a higher global search capability, whereas a small one improves the local search accuracy. Accordingly, this parameter decreases nonlinearly over time according to the following equation:

$$w_P = \frac{1}{2}\left(\cos\left(\left(\frac{b}{N^b}\right)^l \pi\right) + 1\right) \tag{12}$$

- *Self-Cognition Coefficient:* The effectiveness of the guidance of each particle to the particle p_z depends on its experience. The higher its experience, the more valuable is the guidance provided. In this regard, the self-cognition coefficient can be defined to be proportional to the $p_{best,z,n}$ fitness value. In addition, we incorporate randomness in our design to prevent local optima. Furthermore, a separate self-cognition coefficient per particle is defined to improve the search capability. Accordingly, the self-cognition vector $C_1(b) = [c_{11}(b), c_{12}(b), \ldots, c^z_{1,N}(b)]$ is expressed as follows:

$$C_1(b) = F(b)R(b) \tag{13}$$

where R is a non-negative random matrix of dimension $N^z \times N^z$ whose elements are obtained from the range [0, 1]. These elements adjust the amount by which each particle affects the others and are normalized such that the sum of each raw value becomes one. $F(b) = [f_1(b), f_2(b), \ldots, f^z_N(b)]$ indicates the normalized fitness function value. Here, $f_z(b)$ can be computed as

$$f_z(b) = \frac{\frac{1}{p_{best,\,z,n} \cdot cs}}{\sum_{z=1}^{N^z} \frac{1}{p_{best,\,z,n} \cdot cs}} \tag{14}$$

where cs denotes the cost of a given position. The remaining aspect is that the self-cognition coefficient range needs to be determined. It can be derived as follows:

$$\begin{aligned} c_{1,z}(b) = f_1(b)r_{1,z}(b) + f_2(b)r_{2,z}(b) + \cdots \\ + f_{N^z}(b)r_{N^z,z}(b) = \sum_{j=1}^{N^z} f_j(b)r_z(b) \end{aligned} \tag{15}$$

$c_{1z}(b)$ would be within the range [0, 1] because $\sum_{j=1}^{N^z} f_j = 1$ and $0 \le r_{j,z} \le 1$.

- *Social Influence Coefficients:* The parameter $c_2(b)$ specifies the particles' tendency toward $g_{best,z,n}$. It significantly affects the exploitation capability and should be larger than $c_1(b)$ to obtain accurate convergence. Thus, it can be defined as,

$$c_2(b) = \max c_{z,1}(b) \quad 1 \le z \le N^z \tag{16}$$

Based on the proposed parameter-update formulas, the velocity function (10) can be reformulated as follows:

$$\begin{aligned} v_{z,n}(b+1) = w_P \times v_{z,n}(b) + c_{1,z}(b) \times r_1 \times (p_{best,z,n}(b) \\ - x_{z,n}(b)) + c_2(b) \times r_2 \times (g_{best,z,n}(b) - x_{z,n}(b)) \end{aligned} \tag{17}$$

In this study, the number of neurons in hidden layers 1, 2, …, and k and the GRU iteration number are considered as the optimization-seeking variables. Here, the initial particle expression is defined as

$$x_{z,\,n}(0) = (h_1,\ h_2,\ \ldots,\ h_k,\ n_{GRU}) \tag{18}$$

The fitness function of the ELPSO algorithm can be expressed as (19) to optimize the hyperparameters:

$$fit = 0.5\left(\frac{1}{N^p}\sum_{p=1}^{N^p}\frac{\left|y_p - y'_p\right|}{y_p} + \frac{1}{N^q}\sum_{q=1}^{Nq}\frac{\left|y_q - y'_q\right|}{y_q}\right) \tag{19}$$

To protect the forecasting model from overfitting, the fitness function includes the training and validation sample error. Here, the weights of each sample are assigned as 0.5. According to (11)–(19), the optimal GRU model can be determined via repeated iterations.

III. OPTIMAL BIDDING STRATEGY

Although the uncertainties of WTs, PVs, and the local load are handled using the proposed data-driven model, the market price uncertainty remains when the DER aggregator submits day-ahead bids to the power market. In the present study, the IGDT is employed to capture bidding strategy risks associated with market prices. It provides an optimal bidding strategy curve for the DER aggregator to maximize profit. In this section, a problem formulation of the risk-neutral (RN) optimization method is provided. Thereafter, market price uncertainty is considered using the IGDT method. Subsequently, the process of obtaining the risk-constraint bidding strategy and the corresponding bidding profile is presented to the DER aggregator.

A. RN OPTIMIZATION

In the RN strategy, the market price uncertainty is neglected, whereas the variations in WTs, PVs, and local loads are considered via a data-driven model. The objective function is to achieve the DER aggregator's bidding strategy, the DR amount, charging/discharging schedule of the ESS, and especially to maximize profit earned during this process.

1) OBJECTIVE FUNCTION

The optimal bidding strategy is to maximize the profit of the DER aggregator (20):

$$Max\ R_{Agg} = \sum_{t=1}^{N^t} \left\{ \begin{array}{c} \left(R_t^{sell} + R_t^{DR}\right) - \\ \left(C_t^{PM} + C_t^{DR} + C_t^{P,WTs} \right. \\ \left. + C_t^{P,PVs} + C_t^{P,lc}\right) \end{array} \right\} \tag{20}$$

$$R_t^{sell} = P_t^{sell} \times \lambda_t^{sell} \quad \forall t \tag{21}$$

$$R_t^{DR} = P_t^{DR} \times \lambda_t^{PDR} \quad \forall t \tag{22}$$

$$C_t^{PM} = P_t^{PM} \times \lambda_t^{PM} \quad \forall t \tag{23}$$

$$C_t^{DR} = P_t^{DR} \times \lambda_t^{DR} \quad \forall t \tag{24}$$

$$\begin{aligned} C_t^{P,PVs} &= \sum_{t=1}^{N^t} \lambda_t^{RR} \times \max[0, \sum_{r=1}^{N^u} (P_{r,t}^{PV} - P_{r,t}^{R,PV})] \\ &\quad + \sum_{t=1}^{N^t} \pi_2 \times \max[0, \sum_{r=1}^{N^u} (P_{r,t}^{R,PV} - P_{r,t}^{PV})] \quad \forall r, t \end{aligned} \tag{25}$$

$$\begin{aligned} C_t^{P,WTs} &= \sum_{t=1}^{N^t} \lambda_t^{RR} \times \max[0, \sum_{u=1}^{N^u} (P_{u,t}^{WT} - P_{u,t}^{R,WT})] \\ &\quad + \sum_{t=1}^{N^t} \pi_1 \times \max[0, \sum_{u=1}^{N^u} (P_{u,t}^{R,WT} - P_{u,t}^{WT})] \quad \forall u, t \end{aligned} \tag{26}$$

$$\begin{aligned} C_t^{P,\,Base} &= \sum_{t=1}^{N^t} \lambda_t^{RR} \times \max[0, (P_t^{Base} - P_t^{R,\,Base})] \\ &\quad + \sum_{t=1}^{N^t} \pi_3 \times \max[0, P_t^{R,\,Base} - P_t^{Base})] \quad \forall t \end{aligned} \tag{27}$$

The revenue R_{Agg} of DER aggregator comprises several parts: the revenue from providing power to retail customers (21), revenue from providing DR (22), cost of purchasing power from the power market (23), and incentive provided to DR customers (24). The penalty cost of PVs (25) and WTs (26) includes the reserve requirement and spillage costs. The former represents the overestimation where the forecasted day-ahead power from DERs exceeds the actual output power on the operation day; in this case, the aggregator must pay for the additional power. The latter represents the underestimation where the forecasted power is less than that required on the actual dispatch day; in this case, the aggregator needs to pay for the reserve power to satisfy the requirements on the scheduled day. The penalty cost of local load forecasting errors is described in (27).

2) POWER BALANCE CONSTRAINTS

The supply–demand constraint (28) ensures that the power generated by DERs and that purchased from the day-ahead power market is equal to the baseline load after DR, if implemented:

$$P_t^{PM} + P_t^{WT} + P_t^{PV} + P_t^{DR} + P_t^{ESS} = P_t^{Base} \quad \forall t \tag{28}$$

3) DER CONSTRAINTS

The sums of the output power of PVs and WTs are indicated in (29) and (30), respectively. The energy output for both PVs and WTs is limited by the upper bounds given in (31) and (32), respectively [9]:

$$P_t^{PV} = \sum_{r=1}^{N^r} P_{r,t}^{PV} \tag{29}$$

$$P_t^{WT} = \sum_{u=1}^{N^u} P_{u,t}^{WT} \tag{30}$$

$$0 \leq P_{r,t}^{PV} \leq \overline{P_r^{PV}} \quad \forall r, t \tag{31}$$

$$0 \leq P_{u,t}^{WT} \leq \overline{P_u^{WT}} \quad \forall u, t \tag{32}$$

4) ESS CONSTRAINTS

Equations (33)–(39) present the constraints of ESSs [24]. The charging/discharging and state-of-charge (SOC) of ESSs can be computed according to (33) and (34), respectively. The energy limit is ensured in (35), and (36)–(38) represent the SOC bound and charging/discharging limit. It should be noted that the charging and discharging of ESSs cannot occur simultaneously in (39):

$$P_t^{ESS} = P_{ch,t}^{ESS} \times \eta^{ch} - \frac{P_{dch,t}^{ESS}}{\eta^{dch}} \quad \forall t \tag{33}$$

$$X_t = X_{t-1} + P_{ch,t}^{ESS} - P_{dch,t}^{ESS} \quad \forall t \tag{34}$$

$$X_{t0} \leq X_t \leq X_{nt} \quad \forall t \tag{35}$$

$$\underline{X} \leq X_t \leq \overline{X} \quad \forall t \tag{36}$$

$$0 \leq P_{ch,t}^{ESS} \leq \overline{P_{ch,t}^{ESS}} \times U_t^{ch} \quad \forall t \tag{37}$$

$$0 \leq P_{dch,t}^{ESS} \leq \overline{P_{dch,t}^{ESS}} \times U_t^{dch} \quad \forall t \tag{38}$$

$$U_t^{ch} + U_t^{dch} \leq 1, U_t^{ch}, \ U_t^{dch} \in \{0, 1\} \quad \forall t \tag{39}$$

5) DR CONSTRAINTS

In this study, the incentive-based load reduction DR is implemented because it is suitable for current market practices. Furthermore, the DR programs integrated in the wholesale power market are load reductions, as stated in [33]. In this regard, incentive-based DR (IDR) is considered for trading in the day-ahead power market without the need to modify traditional market trading floors. The load curtailment of IDR participants is limited by the upper and lower bounds, as shown in (40), and the total reduction from the IDR participants is given as (41)

$$f^{\min} \times P_t^{base} \leq P_{y,t}^{DR} \leq f^{\max} \times P_t^{base} \forall y, t \tag{40}$$

$$P_t^{DR} = \sum_{y=1}^{N^y} P_{y,t}^{DR} \quad \forall y, t \tag{41}$$

B. IGDT OPTIMIZATION METHOD

1) BACKGROUND OF IGDT

The IGDT-based optimization method is an innovative method for handling uncertainty. It is aimed at determining the maximum level of risk that a decision-maker can endure in the day-ahead power market while satisfying a certain level of profit expectation [22]. To achieve this, the IGDT method focuses on the difference between the real and forecasted values of uncertain parameters. This can be formulated as follows:

$$U(\alpha, \lambda_t^{PM*}) = \left\{ \lambda_t^{PM} : \frac{\left|\lambda_t^{PM} - \lambda_t^{PM*}\right|}{\lambda_t^{PM}} \leq \alpha \right\}, \alpha \geq 0 \tag{42}$$

Equation (42) enables the DER aggregator to make decisions that are RA and/or RS based on the convenience of expressing the difference between the actual and predicted values, without considering the probability distribution of historical data.

There are two processes for solving the optimization problem considering market uncertainty [23], [24]. First, an RN optimization problem is solved, as indicated by (43). Here, the market price uncertainty is not considered:

$$\begin{aligned} &\max \ R_{Agg} = f(q) \\ &s.t. \ G(q) = b \\ &\qquad H(q, x) \leq 0 \end{aligned} \tag{43}$$

In the second process, the market price uncertainty is considered via the IGDT, and the profit deviation factor is then provided. The target of the second process is to maximize/minimize the level of uncertainty while the expected revenue is obtained. These can be formulated as an RA and RS strategy for a DER aggregator, as described in (44) and (45), which are based on the robustness and opportunity functions.

$$\begin{aligned} &\alpha = \max \ \alpha \\ &s.t \ f(q) \geq (1+\delta) \cdot R_{RN} \\ &\quad G(q) = b \\ &\quad b = (1 \pm \alpha) b_{RN} \\ &\quad H(q, x) \leq 0 \end{aligned} \tag{44}$$

$$\begin{aligned} &\beta = \min \ \alpha \\ &s.t \ f(q) \geq (1-\delta) \cdot R_{RN} \\ &\quad G(q) = b \\ &\quad b = (1 \mp \alpha) b_{RN} \\ &\quad H(q, x) \leq 0 \end{aligned} \tag{45}$$

From (44) and (45), it is evident that the RA strategy tends to prevent risks by considering a higher market price even if the DER aggregator earns a profit that is less than that earned using the RN strategy. Meanwhile, in the RS strategy, stability can be forsaken to a certain extent to obtain higher revenue.

2) RA OPTIMIZATION

As mentioned earlier, the robustness function (44) is related to higher market prices. It denotes the largest uncertainty variable value such that the minimum aggregator revenue is higher than the desired revenue target. From (42), a higher market price can be described as

$$\lambda_t^{PM} = \lambda_t^{PM*} + \alpha \lambda_t^{PM*} \tag{46}$$

According to the robustness function described in (44), the maximum aggregator profit is

$$\begin{aligned} &\max \ \alpha \\ &s.t \ R \geq (1-\delta) R_{RN} \\ &\quad \lambda_t^{PM} = (1+\alpha)\lambda_t^{PM*} \end{aligned} \tag{47}$$

It should be noted that the RA objective function involves the acquisition of the risk resistance capability of the DER aggregator. Thus, the worst case should be considered in this strategy. It is assumed to occur when the costs of purchasing from the power market are high.

3) RS OPTIMIZATION

Because the RS strategy is related to a low market price or higher profits of the DER aggregator, it expects to earn higher profits with low market prices. In other words, the RS strategy assesses the possibility of benefiting significantly from low market prices by reducing the purchasing costs. This can be presented as

$$\lambda_t^{PM} = \lambda_t^{PM*} - \alpha \lambda_t^{PM*} \tag{48}$$

The RS strategy can be described as follows based on the opportunistic function and the adopted market price.

$$\begin{aligned} &\min \ \alpha \\ &s.t \ R \leq (1+\delta) R_{RN} \\ &\quad \lambda_t^{EM} = (1-\alpha)\lambda_t^{EM*} \end{aligned} \tag{49}$$

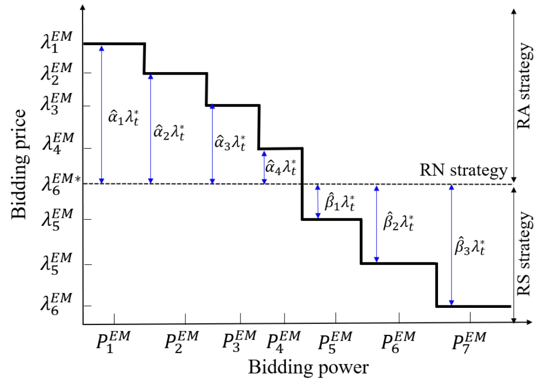


FIGURE 2. Optimal bidding strategy curve.

4) OPTIMAL BIDDING STRATEGY CURVE

To transact energy in the day-ahead power market, the DER aggregator should submit the hourly bidding or offering curves. In this study, the local load of the designated DER aggregator exceeds the DER generation capacity. Therefore, the aggregator participates in the day-ahead power market to procure energy as a consumer. In addition, it is assumed that the DER aggregator functions as a price-taker, considering its relatively small capacity, and only submits non-priced quantity-only bids [14]. To achieve this, a suitable bidding strategy curve is required for the DER aggregator. In addition, the IGDT-based optimization method is used to create an optimal staircase bidding curve based on the robustness and opportunistic functions. As an example, Fig. 2 presents the procedure for creating a seven-level staircase bidding curve for the DER aggregator. At each level, the adopted IGDT-based optimization method is used to determine the confidence level when specific profit-levels above and below the expected profit are selected. Accordingly, the purchasing power corresponding to the cleared price is developed for each profit level. Based on this staircase bidding profile, the DER aggregator can bid for the most reasonable quantity in the day-ahead power market in order to maximize profit.

IV. OVERALL FRAMEWORK

The overall decision-making framework of the proposed data-driven bidding strategy includes two stages, as shown in Fig. 3. It can be summarized as follows:

Step 1: Collect the multivariate DER and local load historical data and perform data pre-processing through cleansing, normalization, and feature selection.

Step 2: Construct the forecasting model GRU and select the optimized hyperparameter from the ELPSO algorithm.

Step 3: Forecast the hourly WT output, PV output, and local load of the DER aggregator as input data for the second stage.

Step 4: Solve the RN optimization problem by maximizing the aggregator's profit without considering the market price uncertainty.

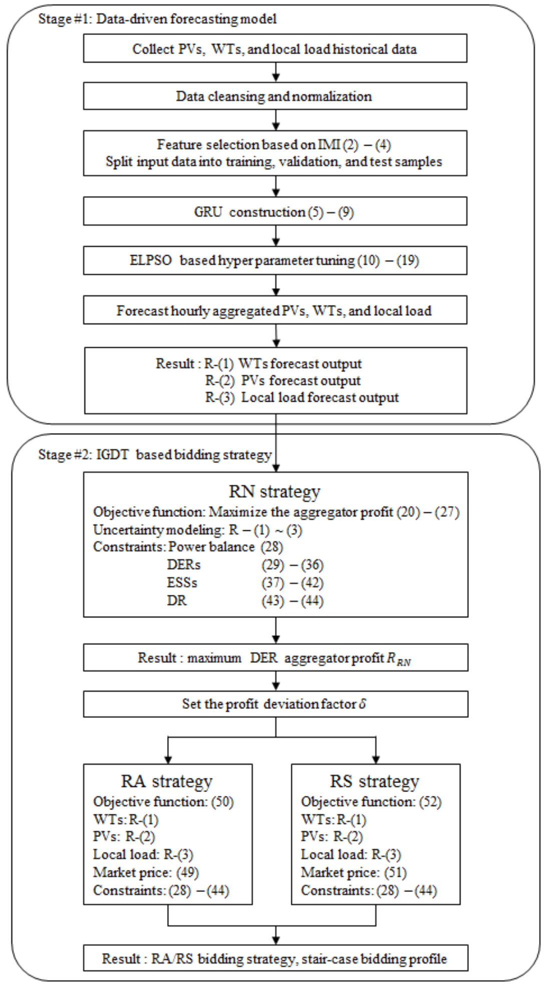


FIGURE 3. Illustration of the proposed bidding strategy architecture.

Step 5: Obtain the RN bidding strategy results in conjunction with the maximized profit value and deviation factor.

Step 6: Consider the market price uncertainty using the IGDT and obtain both the RA and RS bidding strategies in conjunction with the corresponding acceptable uncertainty levels.

Step 7: Develop the hourly staircase bidding profile based on the RA/RS results.

V. CASE STUDY

A. DESCRIPTION OF BENCHMARK DATASET

In the current study, the performance of the proposed data-driven bidding strategy is verified using an illustrative system. Here, the DER aggregator involves 500 WTs, 400 PVs, 100 ESSs, and 100 customers for bidding in the day-ahead power market. To predict the DERs and local load of the aggregator and to evaluate the proposed data-driven model performance, hourly historical data from 2019 are acquired from the publicly available PJM market and modified [37]. Multivariate variables including the temperature,

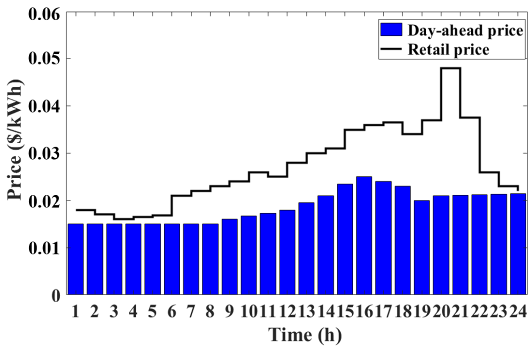


FIGURE 4. Day-ahead market and retail prices.

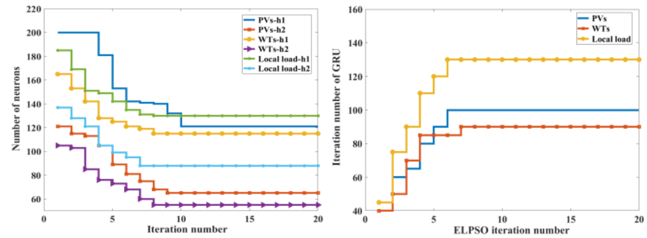
TABLE 2. Parameter settings.

Parameter	Value	Parameter	value
$\overline{X}, \underline{X}$ (kWh)	90, 10	η^{ch}, η^{dch}	0.95
$\overline{P^{ESS}_{ch,t,s}}, \overline{P^{ESS}_{dch,t,s}}$ (kWh)	20	$U^{ch}_{t,s}$, $U^{dch}_{t,s}$ (kW)	20
f^{min}, f^{max}	0, 0.3	π_1, π_2, π_3 ($/kWh)	0.025
N^{z}, N^{b}	2000, 20	δ	0.3

humidity, hour, and dew point are typical inputs used to train the forecasting model. To forecast each component, the input data sequence is set to the previous 168 h ($= 7 \times 24$) of data. Then, the prepared dataset is subjected to data processing, which includes cleansing, normalization, and feature selection. Thus, the optimal input variables are extracted from the provided dataset. Of the entire dataset, 70%, 20%, and 10% are allocated to training, validation, and test samples, respectively. The initial particle positions are in the ranges of [1, 150] and [40, 200] for h_k and n_{GRU}, respectively. In this study, adam, relu, and mse were set as the optimizer, activation, and loss functions, respectively, for the GRU model. The deep learning simulation is implemented based on the Python TensorFlow framework. Additionally, to illustrate the performance of the proposed data-driven model, we compare its results with those of three well-established forecasting models (namely, LSTM, RNN, and SVR) [38]. Fig. 4 shows the day-ahead electricity and retail prices that are adopted [34], and Table 2 depicts several important parameter settings for bidding in the day-ahead power market [35], [36]. ELPSO is utilized to solve the optimization problem. The simulations are performed using a 3.4 GHz CPU with 16 GB RAM and run on a Windows 10 Pro 64-bit operating system.

B. DATA-DRIVEN FORECASTING RESULTS

A random day from our test samples is adopted to analyze and evaluate the performance of the proposed model in terms of handling uncertainties. In the deep learning simulation, the two-layer GRU model shows the highest performance because it is sufficient for the relationships between the input and output variables of the adopted historical data. Hence, the optimized hyperparameter sets for the PVs, WTs, and local load forecasting models (i.e., h_1, h_2, and n_{GRU},

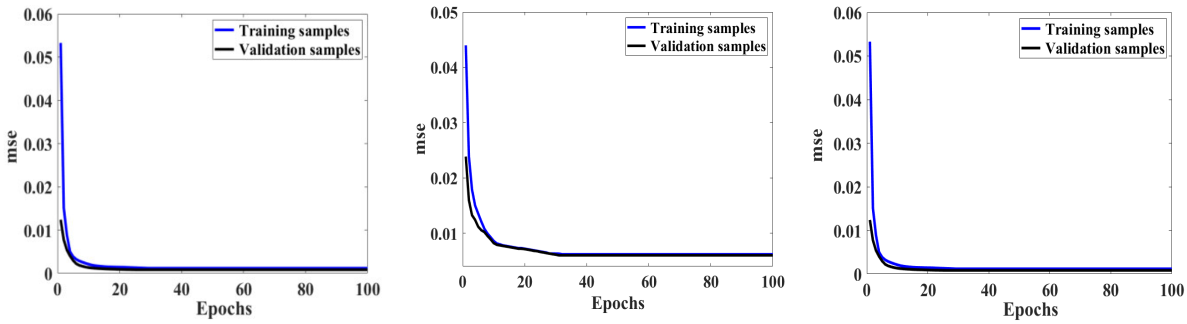


(a) Hidden layer neurons, (b) GRU Iteration number

FIGURE 5. Hyperparameter optimization results.

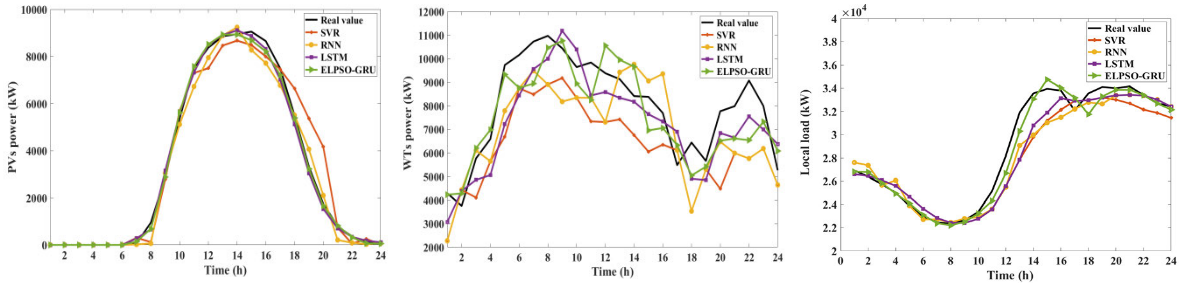
respectively) are obtained using iterations of the ELPSO algorithm as depicted in Fig. 5.

The overfitting and convergence problems are prevented by optimizing the GRU model hyperparameters, as shown in Fig. 6. The validation error decreases gradually as the training error does for the PVs, WTs, and local loads. Accordingly, the proposed model resolves the overfitting problem, and the error between the training and validation is marginal. Indeed, there is no variance or bias in the renewable output and local load forecasting. Fig. 7 shows the day-ahead forecasted DERs output power and local load with hourly resolution of the proposed data-driven model with other benchmark models. From the graphical illustration, it is obvious that all forecasting models are capable of capturing the nonlinear behavior of historical data and forecasting future output power and local load owing to the IMI feature selection technique. It is also clear that the proposed model closely follows the real value as compared to other benchmark models in terms of DER output and local load forecasting because of the IMI feature extraction and optimized hyperparameter settings. Fig. 8 presents a comparative analysis of the coefficient of determination, R^2, to illustrate the performance of the proposed data-driven model. Here, the x- and y-axes represent the actual and forecasted values, respectively. Fig. 8(a)–(d) show the actual and forecasted PV outputs, whereas Fig. 8(e)–(h) and 8(i)–(l) depict the results of the WTs and local load, respectively. R^2 is one of the statistical evaluation indicators for comparing the prediction performance of variables. It can be computed as $R^2 = 1 - \sum_u \left(y_u^{predict} - y_u^{real}\right)^2 / \sum_u \left(y_u^{average} - y_u^{real}\right)^2$. This measure indicates the extent to which the variance of the forecasted variable validates the variance of the actual variable. For example, an R^2 value of 1 implies that the predictions of the forecasting model perfectly fit the actual data. From Fig. 8, it is evident that the overall R^2 value obtained via the proposed model is higher than those from the other models. This demonstrates its superiority in terms of the forecasting accuracy. Table 3 summarizes the forecasting error of each benchmark model in terms of two well-established error criteria: the root mean square error (RMSE) and the mean absolute error (MAE). As shown in this Table, the proposed model displays a reasonable performance in terms of the prediction of stochastic variables. The error in the terms of RMSE and MAE is also reduced significantly, which results



(a) PVs (b) WTs (c) Local loads

FIGURE 6. Learning curve results.



(a) PVs (b) WTs (c) Local loads

FIGURE 7. Day-ahead prediction results with respect to various forecasting approaches.

TABLE 3. Forecasting error criteria.

Forecasting model	PVs		WTs		Local load	
	RMSE	MAE	RMSE	MAE	RMSE	MAE
SVR	0.1010	0.0731	0.1221	0.0828	0.0693	0.0578
RNN	0.0954	0.0604	0.1144	0.0742	0.0541	0.0428
LSTM	0.0678	0.0395	0.0972	0.0614	0.0463	0.0311
Proposed model	0.0542	0.0303	0.0891	0.0443	0.0365	0.0224

TABLE 4. Comparison analysis of different feature selections.

Feature selection	RMSE			MAE		
	PVs	WTs	Local load	PVs	WTs	Local load
CA	0.0619	0.1081	0.0501	0.0414	0.0502	0.0297
PCA	0.0558	0.0981	0.0441	0.0397	0.0471	0.0277
MI	0.0571	0.0955	0.0475	0.0381	0.0498	0.0261
IMI	0.0542	0.0891	0.0365	0.0303	0.0443	0.0224

in improved accuracy. Table 4 illustrates the forecasting error with respect to those of the benchmark feature selection techniques, i.e., CA, PCA, and MI. Evidently, the proposed IMI technique achieves the lowest RMSE and MAE for each uncertain component because it can analyze both linear and nonlinear dataset characteristics considering the target and average variables. By contrast, the other three techniques do not always guarantee the capture of effective input features. Furthermore, a negative prediction performance effect could also be observed in the prediction of the renewables and local load. According to Tables 3 and 4, this improvement is achieved because the proposed model can optimize the hyperparameter settings in conjunction with a suitable feature selection technique (IMI), without being affected by the overfitting and convergence problems. Therefore, this forecasting model is the most suitable for handling uncertainties and for the accurate modeling of DER aggregator components prior to bidding in the day-ahead power market.

C. OPTIMAL BIDDING STRATEGY RESULTS

The day-ahead RN bidding strategy of the DER aggregator according to the proposed data-driven forecasting results is presented in Fig. 9. Excluding the impact of the uncertainty in the day-ahead market price, the total DER aggregator profit is \$12513.11. Based on the RN bidding result, 34870.17 kW of the baseline load is decreased by the IDR program, and 27250.47 kW is reduced during the peak period when the day-ahead market price is high. The DER aggregator procures 389974.47 kW of electricity from the day-ahead power market to fulfill the aggregated local load and incurs a penalty of 36352.51 kW due to forecasting errors of the renewables and local load.

The impact of the day-ahead forecasting accuracy of the renewables and local load on the RN aggregator's bidding strategy is shown in Table 5. For the SO method, the Weibull, beta, and normal probability distribution functions are used to model the uncertainties of WTs, PVs, and local load, respectively. Thereafter, the multi-scenario tree method and differential evolution clustering algorithm are applied for scenario generation and reduction [9]. In RO, the uncertainties are fully considered, such that the most conservative solution is achieved [8]. Table 5 shows that the RO solution presents the lowest profit at the expense of its high preferences for maximum robustness; by contrast, the profit of the DER aggregator

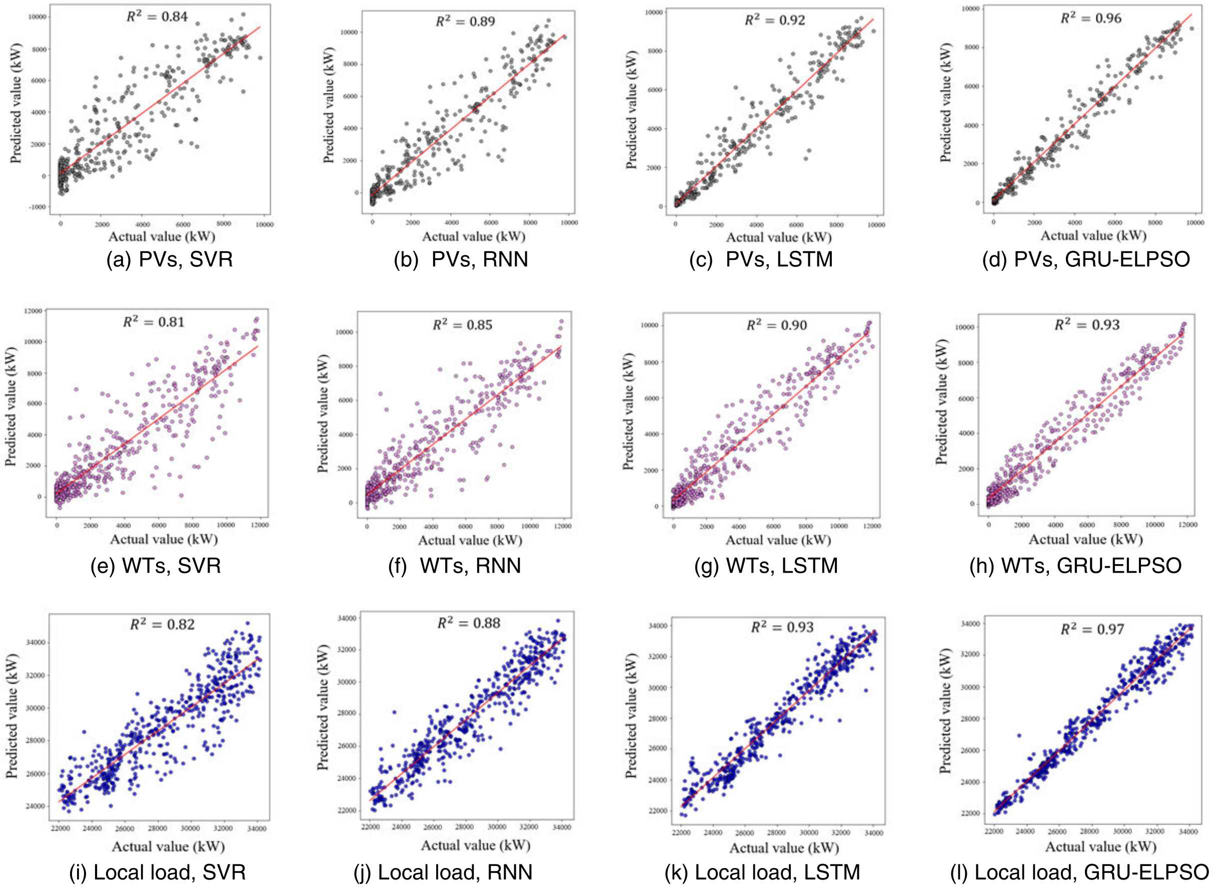


FIGURE 8. R^2 comparison for different forecasting approaches.

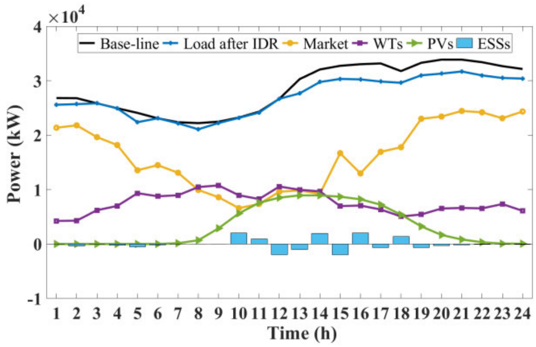


FIGURE 9. Day-ahead RN bidding strategy result.

TABLE 5. Profit comparison for methods of handling uncertainty.

Uncertainty modeling method	Total profit ($)	Penalty cost ($)			
		PVs	WTs	Local load	Total
RO	10 512	305.84	1325.33	714.44	2346.12
SO	11 298	207.12	1141.15	380.61	1728.88
SVR	10 840	249.99	1071.22	685.29	2006.5
RNN	11 360	180.23	859.99	573.66	1711.89
LSTM	12 215	82.84	638.92	522.09	1243.85
Proposed model	12 513	56.27	538.43	280.98	875.68
Actual value	13 512	-	-	-	-

is increased when using SO. With regard to the SO solution, although scenario generation and reduction are employed, the economy and reliability issues are not fully addressed in the bidding strategy problem. The obtained results rely fully on the historical probability distribution functions, which may vary substantially according to system conditions. As a result, the robustness of SO could deteriorate if the obtained probability distribution functions differ significantly from the actual values. By contrast, the data-driven approach is applicable for addressing uncertainties where the probability distribution functions are not required. RNNs, LSTM, and the proposed model provide high-profit solutions for the DER aggregator, as compared with the RO and SO solutions. It is also worth mentioning that the total DER aggregator profit is maximized when the proposed data-driven model is adopted, because the penalty costs incurred due to the forecasting errors are reduced significantly. In addition, the proposed model yields a profit that is closest to the actual profit because it incurs the lowest penalty costs. Considering the present circumstances, the actual financial profit and cost of the DER aggregator would be significantly higher over a wider

TABLE 6. Comparison of results of optimization algorithms.

Algorithm	Best solution ($)	Worst solution ($)	Average ($)	Average CPU (s)
GA	13 506	13 502	13 504	191.35
ABC	13 506	13 503	13 504	180.12
SM	13 507	13 505	13 506	109.51
Whale	13 508	13 505	13 506	102.85
EGA	13 509	13 507	13 508	103.48
SE-PSO	13 511	13 508	13 509	105.11
ELPSO	13 512	13 511	13 511	82.44

time-step or scale. This represents the considerable effect of the proposed data-driven forecasting model on practical works for achieving accurate uncertainty variable modeling.

Performance analysis is conducted with various meta-heuristic algorithms to demonstrate the superiority of the proposed ELPSO in solving the optimal bidding strategy problem of DER aggregators. For an unbiased comparison, we adopt the RN bidding strategy of the DER aggregator without considering the uncertainty in market price. In total, seven meta-heuristic algorithms are considered: genetic algorithm (GA) [39], artificial bee colony (ABC) [40], spider monkey (SM) [41], whale [42], enhanced GA (EGA) [39], security enhanced-PSO (SE-PSO) [43], and the proposed ELPSO algorithm. The simulations are implemented 15 times, and the results for the best, worst, mean, and corresponding average computation time (CPU) are obtained. Table 6 presents the comparison results for the seven optimization algorithms. As is evident, the proposed ELPSO provides the most profitable solution in conjunction with a reasonable CPU time. This indicates that it can be the most reasonable optimization solver when applied to real and large systems. Another advantage of the proposed ELPSO is that the difference between the best and worst solutions is minimum. Thus, these quantitative results reveal the global search capability, high accuracy, and low computational load of the proposed ELPSO algorithm, thereby demonstrating its efficiency in determining the optimal bidding strategy for DER aggregators.

Fig. 10 shows the robust and opportunity indexes of the DER aggregator's RA and RS bidding strategies. The specific profit and corresponding risk level are achieved by varying the profit deviation factor from the RN strategy by 0.3. For the RA strategy, the DER aggregator can tolerate a risk level of 0.4306 for a maximum profit deviation factor of 0.21. Meanwhile, for the RS strategy, the risk level is 0–0.433 for a maximum profit deviation factor 0.18. Unless the robust or opportunity index exceeds its boundary value, the RA aggregator is more robust and the RS aggregator gains more profit. It should be noted that no feasible solution for the DER aggregator's bidding strategy can be obtained if the risk index exceeds its threshold.

Fig. 11 illustrates several RA and RS bidding strategies based on the risk level shown in Fig. 10. Two regularities in the bidding strategy can be observed as the risk level

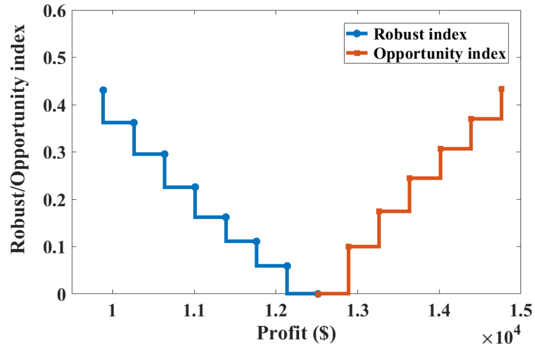


FIGURE 10. Robustness and opportunistic curves.

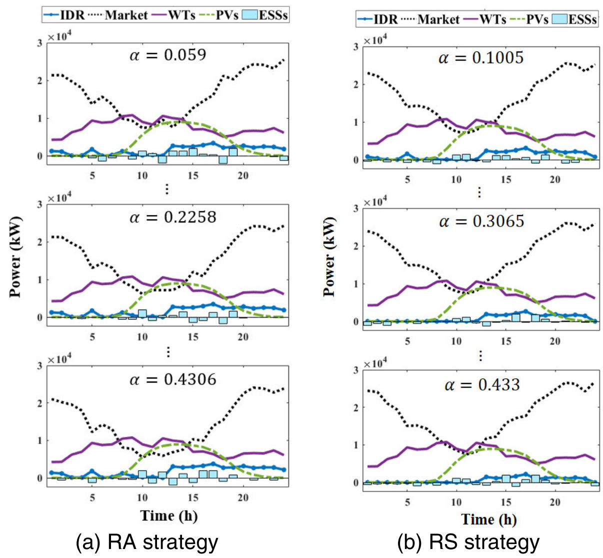


FIGURE 11. RA/RS bidding strategies under different risk levels.

varies. First, the power procured from the day-ahead market, represented by the black dotted line, exhibits a downward trend as the DER aggregator prefers a robust strategy because of the high purchasing cost. Second, in contrast to the market purchasing amount, the IDR exhibits an upward trend to fulfill the required local load. Thus, it can be concluded that the RA aggregator's bidding strategy becomes more robust at the expense of a higher interest by reducing the power procured, which exhibits uncertainty. Meanwhile, the RS aggregator pursues higher profits, wherein higher risk must be tolerated in exchange for high revenues. The 13-level optimal day-ahead bidding profile of the DER aggregator with the obtained RA/RS bidding strategy results is shown in Fig. 12. As a price-taker, the RA aggregator reduces the purchasing amount where announced purchasing cost is high and increases in RS strategy to earn high profit. Specifically, at $t = 2$, the bidding quantity varies from 20210.08 kW to 23912.63 kW according to the DER aggregator's risk tolerance. Therefore, it should be noted that the obtained bidding profile can assist the DER aggregator in terms of purchasing

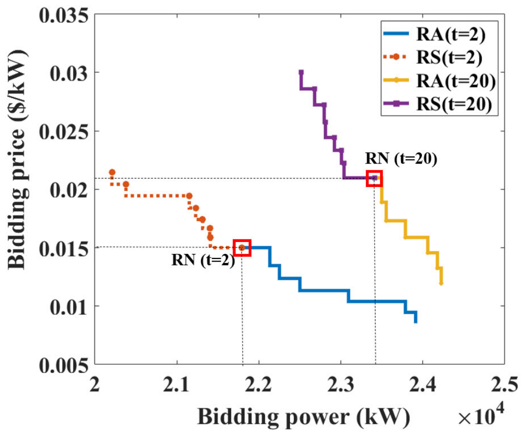


FIGURE 12. Optimal bidding profile.

power from the day-ahead market at the most reasonable price according to its risk preference.

VI. CONCLUSION

This paper proposes the framework of a data-driven bidding strategy for a DER aggregator in the day-ahead power market considering uncertainties. Prior to bidding in the day-ahead power market, a precise modeling of renewables and the local load was implemented using a data-driven forecasting model. The GRU model was used for training and forecasting, whereas the ELPSO algorithm was utilized to optimize the GRU hyperparameter. To achieve high accuracy, the IMI feature selection technique was adopted to extract the best input dataset. The proposed data-driven model was validated by comparing it with three benchmark forecasting models: LSTM, RNN, and SVR. The comparison results of the prediction, R^2, and the errors revealed that the proposed model outperforms the other models in terms of forecasting accuracy and contributes toward reducing the penalty costs. The IGDT-based optimization method was applied to manage market price uncertainty and successfully bid in the day-ahead power market. The RN bidding strategy was solved beforehand to obtain the profit standard. This value serves as the basis for determining the profit variation when the deviation factors are provided under uncertain cases. Subsequently, by varying the risk level, the DER aggregator succeeded in providing two bidding strategies (namely, RA and RS) according to its preference for robustness or high profit. These results demonstrated that the RA aggregator tends to increase robustness at the expense of profit and vice versa. A staircase bidding curve was developed based on the obtained RA and RS bidding strategy results. It could serve as a guideline for the DER aggregator to purchase power at the most reasonable price, while maximizing profit. Therefore, the proposed bidding framework can assist the DER aggregator in the decision-making process to bid in the day-ahead power market by addressing the economic and robustness issues associated with various uncertainties. This work can be improved further by considering plug-in hybrid electric vehicles and microgrids and by exploring their impacts on the bidding strategy of DER aggregators. Moving forward, the bidding and offering strategies of DER aggregators in reserve, intraday, and real-time markets could also be investigated for applying more practical power market scenarios.

REFERENCES

[1] G. E. Asimakopoulou and N. D. Hatziargyriou, "Evaluation of economic benefits of DER aggregation," *IEEE Trans. Sustain. Energy*, vol. 9, no. 2, pp. 499–510, Apr. 2018.

[2] X. Lu, K. Li, H. Xu, F. Wang, Z. Zhou, and Y. Zhang, "Fundamentals and business model for resource aggregator of demand response in electricity markets," *Energy*, vol. 204, Aug. 2020, Art. no. 117885.

[3] P. Fazlalipour, M. Ehsan, and B. Mohammadi-Ivatloo, "Risk-aware stochastic bidding strategy of renewable micro-grids in day-ahead and real-time markets," *Energy*, vol. 171, pp. 689–700, Mar. 2019.

[4] M. D. Somma, G. Graditi, and P. Siano, "Optimal bidding strategy for a DER aggregator in the day-ahead market in the presence of demand flexibility," *IEEE Trans. Ind. Electron.*, vol. 66, no. 2, pp. 1509–1519, Feb. 2019.

[5] P. Sheikhahmadi and S. Bahramara, "The participation of a renewable energy-based aggregator in real-time market: A bi-level approach," *J. Cleaner Prod.*, vol. 276, Dec. 2020, Art. no. 123149.

[6] H. Wu, X. Liu, B. Ye, and B. Xu, "Optimal dispatch and bidding strategy of a virtual power plant based on a stackelberg game," *IET Gener., Transmiss. Distribution*, vol. 14, no. 4, pp. 552–563, Feb. 2020.

[7] S. Abapour, B. Mohammadi-Ivatloo, and M. T. Hagh, "Robust bidding strategy for demand response aggregators in electricity market based on game theory," *J. Cleaner Prod.*, vol. 243, Jan. 2020, Art. no. 118393.

[8] G. Liu, Y. Xu, and K. Tomsovic, "Bidding strategy for microgrid in day-ahead market based on hybrid stochastic/robust optimization," *IEEE Trans. Smart Grid*, vol. 7, no. 1, pp. 227–237, Jan. 2016.

[9] H. J. Kim, M. K. Kim, and J. W. Lee, "A two-stage stochastic p-robust optimal energy trading management in microgrid operation considering uncertainty with hybrid demand response," *Int. J. Electr. Power Energy Syst.*, vol. 124, Jan. 2021, Art. no. 106422.

[10] Y. Liu, C. Jiang, J. Shen, and J. Hu, "Coordination of hydro units with wind power generation using interval optimization," *IEEE Trans. Sustain. Energy*, vol. 6, no. 2, pp. 443–453, Apr. 2015.

[11] H. Nezamabadi and V. Vahidinasab, "Market bidding strategy of the microgrids considering demand response and energy storage potential flexibilities," *IET Gener., Transmiss. Distrib.*, vol. 13, no. 8, pp. 1346–1357, Apr. 2019.

[12] M. Ansari, A. T. Al-Awami, E. Sortomme, and M. A. Abido, "Coordinated bidding of ancillary services for vehicle-to-grid using fuzzy optimization," *IEEE Trans. Smart Grid*, vol. 6, no. 1, pp. 261–270, Jan. 2015.

[13] A. T. Al-Awami, N. A. Amleh, and A. M. Muqbel, "Optimal demand response bidding and pricing mechanism with fuzzy optimization: Application for a virtual power plant," *IEEE Trans. Ind. Appl.*, vol. 53, no. 5, pp. 5051–5061, Oct. 2017.

[14] B. Vatandoust, A. Ahmadian, M. A. Golkar, A. Elkamel, A. Almansoori, and M. Ghaljehei, "Risk-averse optimal bidding of electric vehicles and energy storage aggregator in day-ahead frequency regulation market," *IEEE Trans. Power Syst.*, vol. 34, no. 3, pp. 2036–2047, May 2019.

[15] H. Jahangir, H. Tayarani, S. Baghali, A. Ahmadian, A. Elkamel, M. A. Golkar, and M. Castilla, "A novel electricity price forecasting approach based on dimension reduction strategy and rough artificial neural networks," *IEEE Trans. Ind. Informat.*, vol. 16, no. 4, pp. 2369–2381, Apr. 2020, doi: 10.1109/TII.2019.2933009.

[16] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436–444, May 2015.

[17] J. Li and J. Wang, "Forecasting of energy futures market and synchronization based on stochastic gated recurrent unit model," *Energy*, vol. 213, Dec. 2020, Art. no. 118787.

[18] Z. Niu, Z. Yu, W. Tang, Q. Wu, and M. Reformat, "Wind power forecasting using attention-based gated recurrent unit network," *Energy*, vol. 196, Apr. 2020, Art. no. 117081.

[19] F. Yang, S. Zhang, W. Li, and Q. Miao, "State-of-charge estimation if lithium-ion batteries using LSTM and UKF," *Energy*, vol. 201, Jun. 2020, Jun. Jun. 117664.

[20] H. Jahangir, S. S. Gougheri, B. Vatandoust, M. A. Golkar, A. Ahmadian, and A. Hajizadeh, ''Plug-in electric vehicle behavior modeling in energy market: A novel deep learning-based approach with clustering technique,'' *IEEE Trans. Smart Grid*, vol. 11, no. 6, pp. 4738–4748, Nov. 2020, doi: 10.1109/TSG.2020.2998072.

[21] S. M. Moghaddas-Tafreshi, M. Jafari, S. Mohseni, and S. Kelly, ''Optimal operation of an energy hub considering the uncertainty associated with the power consumption of plug-in hybrid electric vehicles using information gap decision theory,'' *Int. J. Electr. Power Energy Syst.*, vol. 112, pp. 92–108, Apr. 2019, doi: 10.1016/j.ijepes.2019.04.040.

[22] N. Rezaei, A. Ahmadi, A. Khazali, and J. Aghaei, ''Multiobjective riskconstrained optimal bidding strategy of smart microgrids: An IGDT based normal boundary intersection approach,'' *IEEE Trans. Ind. Informat.*, vol. 15, no. 3, pp. 1532–1543, Mar. 2019, doi: 10.1109/TII.2018.2850533.

[23] M. Vahid-Ghavidel, N. Mahmoudi, and B. Mohammadi-Ivatloo, ''Self-scheduling of demand response aggregators in short-term markets based on information gap decision theory,'' *IEEE Trans. Smart Grid*, vol. 10, no. 2, pp. 2115–2126, Mar. 2019, doi: 10.1109/TSG.2017.2788890.

[24] R. Mafakheri, P. Sheikhahmadi, and S. Bahramara, ''A two-level model for the participation of microgrids in energy and reserve markets using hybrid stochastic-IGDT approach,'' *Int. J. Elect. Power Energy Syst.*, vol. 119, Jul. 2020, Art. no. 105977.

[25] C. Zhang, J. Zhou, C. Li, W. Fu, and T. Peng, ''A compound structure of ELM based on feature selection and parameter optimization using hybrid backtracking search algorithm for wind speed forecasting,'' *Energy Convers. Manage.*, vol. 143, pp. 360–376, Jul. 2017.

[26] N. Amjady and F. Keynia, ''Day-ahead price forecasting of electricity markets by mutual information technique and cascaded neuro-evolutionary algorithm,'' *IEEE Trans. Power Syst.*, vol. 24, no. 1, pp. 306–318, Feb. 2009, doi: 10.1109/TPWRS.2008.2006997.

[27] S. Li, W. Li, C. Cook, C. Zhu, and Y. Gao, ''Independently recurrent neural network (IndRNN): Building a longer and deeper RNN,'' in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2018, pp. 5457–5466.

[28] S. Hochreiter and J. Schmidhuber, ''Long short-term memory,'' *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.

[29] K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H, Schwenk, and Y. Bengio, ''Learning phrase representations using RNN encoder-decoder for statistical machine translation,'' in *Proc. Conf. Empirical Methods Natural Lang. Process. (EMNLP)*, 2014, pp. 1724–1734.

[30] P. A. González and J. M. Zamarreño, ''Prediction of hourly energy consumption in buildings based on a feedback artificial neural network,'' *Energy Buildings*, vol. 37, no. 6, pp. 595–601, 2005.

[31] J. Kennedy and R. Eberhart, ''Particle swarm optimization,'' in *Proc. Int. Conf. Neural Netw. (ICNN)*, Perth, WA, Australia, vol. 4, Nov. 1995, pp. 1942–1948, doi: 10.1109/ICNN.1995.488968.

[32] M. Parvania, M. Fotuhi-Firuzabad, and M. Shahidehpour, ''Optimal demand response aggregation in wholesale electricity markets,'' *IEEE Trans. Smart Grid*, vol. 4, no. 4, pp. 1957–1965, Dec. 2013.

[33] T. V. Sohrabi, M. A. Jirdehi, and R. Hemmati, ''Energy management in microgrid based on the multi objective stochastic programming incorporating portable renewable energy resource as demand response option,'' *Energy*, vol. 118, no. 1, pp. 827–839, Jan. 2017.

[34] M. R. Ebrahimi and N. Amjady, ''Adaptive robust optimization framework for day-ahead microgrid scheduling,'' *Int. J. Elect. Power Energy Syst.*, vol. 107, pp. 213–223, May 2019.

[35] M. A. Hossain, H. R. Pota, S. Squartini, and A. F. Abdou, ''Modified PSO algorithm for real-time energy management in grid-connected microgrid,'' *Renew. Energy*, vol. 136, pp. 746–757, Jun. 2019.

[36] *PJM Electricity Market*. [Online]. Available: https://www.pjm.com

[37] H. Jahangir, M. A. Golkar, F. Alhameli, A. Mazouz, A. Ahmadian, and A. Elkamel, ''Short-term wind speed forecasting framework based on stacked denoising auto-encoders with rough ANN,'' *Sustain. Energy Technol. Assessments*, vol. 38, Apr. 2020, Art. no. 100601.

[38] D. L. Duan, X. D. Ling, X. Y. Wu, and B. Zhong, ''Reconfiguration of distribution network for loss reduction and reliability improvement based on an enhanced genetic algorithm,'' *Int. J. Elect. Power Energy Syst.*, vol. 64, pp. 88–95, Jan. 2015.

[39] D. Karaboga and B. Basturk, ''A powerful and efficient algorithm for numerical function optimization: Artificial bee colony (ABC) algorithm,'' *J. Global Optim.*, vol. 39, no. 3, pp. 459–471, 2007.

[40] G. Deb, K. Chakraborty, and S. Deb, ''Spider monkey optimization technique–based allocation of distributed generation for demand side management,'' *Int. Trans. Elect. Energy Syst.*, vol. 29, May 2019, Art. no. e120009.

[41] L. A. Wong, V. K. Ramachandaramurthy, S. L. Walker, P. Taylor, and M. J. Sanjari, ''Optimal placement and sizing of battery energy storage system for losses reduction using whale optimization algorithm,'' *J. Energy Storage*, vol. 26, Dec. 2019, Art. no. 100892.

[42] B. M. Hussein, ''A new scaled fuzzy method using PSO segmentation (SePSO) applied for two area power system,'' *Int. J. Elect. Comput. Eng.*, vol. 9, no. 2, pp. 815–825, 2019.

HYUNG JOON KIM received the B.S. and M.S. degrees from the Department of Energy System Engineering, Chung-Ang University, Seoul, South Korea, in 2018 and 2021, respectively, where he is currently pursuing the Ph.D. degree. His research interests include data-driven approaches for power system analysis, renewable energy resources, microgrids, demand response, and energy management systems.

HYUN JOON KANG received the B.S. degree from the Department of Energy System Engineering, Chung-Ang University, Seoul, South Korea, in 2021, where he is currently pursuing the M.S. degree. His research interests include electric vehicles and vehicle-to-grid.

MUN KYEOM KIM received the Ph.D. degree in electrical and computer engineering from Seoul National University. He is currently a Professor with the School of Energy System Engineering, Chung-Ang University, Seoul, South Korea. His research interests include operational techniques in hybrid AC/DC power systems, AI-based smart power networks, big data-based demand response, real-time market design, and multi-agent-based smart city intelligence.