A Decentralized Optimization Strategy for Distributed Generators Power Allocation in Microgrids Based on Load Demand–Power Generation Equivalent Forecasting

Shicong Zhang 1,*, Zilong Yu 2, Bowen Zhou 1, Zhile Yang 3, and Dongsheng Yang 1

1 College of Information Science and Engineering, Northeastern University, Shenyang 110819, China; yangdongsheng@mail.neu.edu.cn (D.Y.); zhoubowen@ise.neu.edu.cn (B.Z.)
2 The University of Sydney, Sydney 2008, Australia; 1700686@stu.neu.edu.cn.
3 Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China; zl.yang@siat.ac.cn (Z.Y.)
* Correspondence: 1770492@stu.neu.edu.cn; Tel.: +86-159-9887-6298

Received: 17 December 2019; Accepted: 31 January 2020; Published: 3 February 2020

Abstract: In order to guarantee the economic and reliable operation of renewable Distributed Generators (DGs) in microgrids, a decentralized optimization strategy for DGs power allocation is proposed in this paper. According to the method, all processes and parameters are designed in a fully distributed way. To achieve decentralization and to maintain the balance between power supply and load demand, a load demand–power generation equivalent forecasting method is proposed to improve the strategy through replacing information of load demand by predicted power output, which removes the load prediction center and load sensor devices. The data of historical power generation, which is used for prediction, has already satisfied the balance constraint between power supply and load demand. Therefore, when the balance between the real power output and the predicted power output is gained, the balance constraint of power supply and load demand is achieved. Meanwhile, the uncertainty and forecasting errors of renewable generation are taken into account in the cost functions to optimize the expense of DG operation comprehensively. Then, the proposed algorithm is expounded in detail and the convergence is proved by eigenvalue perturbation theory. Finally, various cases are simulated to verify the accuracy and effectiveness of the proposed method. In summary, the proposed method are effective tools for DGs economic power allocation and the decentralization of microgrid system.

Keywords: microgrid; decentralization; equivalent forecasting; consensus; optimization

1. Introduction

To cope with the serious anthropogenic climate change and to decrease greenhouse gas emissions, more and more renewable energy sources are being applied in the advanced microgrid [1]. Hence, the management of renewable energy sources to guarantee microgrids operating economically and reliably becomes an essential issue, which has drawn great attention from researchers in various areas. Particularly, the question ‘to centralize or to decentralize’ is a highlight on the control and management of microgrids. The centralized optimization algorithms require global information of the whole system, which deeply depend on the control center. With the rapid increasing number of Distributed Generator (DG) units, the centralized methods are not suitable for a large-scale network, due to significance communications and massive computational overhead. Such disadvantages motivate researches on decentralized optimization methods. Literature [2] discussed and reviewed the state-of-the-art solutions regarding Micro Grid Management System
(MGMS) in the primary, secondary, and tertiary levels, and drew a conclusion that the decentralized MGMS framework delivers not only the same control function as the centralized one, but also be with greater scalability, reliability, and resiliency. Meanwhile, the “plug and play” [3] demand of DG drives the tendency toward decentralization as well.

At present, decentralized optimization algorithms mainly depend on consensus or multi-agent systems. Each node/agent in a system only concerns about local decision variables instead of global information, and neighboring nodes/agents communicate with each other to cooperate together for achieving optimization results [4]. For example, a multi-agent system-based hierarchical control method was applied in autonomous microgrids for cooperative frequency control with communication constraints [5]. A state-based consensus method [6] was proposed, which made use of the state information of DGs to calculate the optimized dispatching strategy under a layered architecture. However, such methods still rely on the control center to collect the information and to send to each DG. To fully decentralize the optimization and to remove the control center, [7] proposed a consensus-based approach for economic dispatch problems in a smart grid without any center or leaders. On this basis, [8] proposed a distributed algorithm to solve the resource allocation over dynamic digraphs, which is adaptable to the scenario under the time-varying graph. Noticeably, all of the above methods required the information of total load demand in a microgrid to keep the balance constraint between power supply and load demand in the optimization. Moreover, to consider the impact of load on the cost optimization, [9–11] established the utility function of each load. However, the total load demand information was gained through load forecasting [12], which depended on a prediction center. Hence, the center restrains the decentralization faith. In addition, the setting-up of utility function of each load needs distributed loads’ information, which requires setting sensors at each load. It is difficult to implement in large-number and distributed load scenarios. Therefore, an effective measure should be carried out to solve the above problem.

Meanwhile, the cost function in the decentralized optimization is always treated as a simple quadratic polynomial function. For example, the consensus iterative algorithm based on bias feedback dynamic adjustment method was proposed in [13] for distributed robust economic dispatch. A method relaxed hybrid consensus Alternating Direction Multiplier Method (ADMM) with coupling constraints was proposed in [14] for distributed convex optimization. They all ignored the uncertainty of renewable DGs influencing on the expense. Since the uncertainty is mainly due to randomness and volatility of wind/photovoltaic power, the optimization of MGMS needs to consider the uncertainty of DGs in the cost function. In [15–17], authors analyzed probabilistic forecasting generation of wind turbine (WT) and photovoltaic (PV) power through machine learning during the optimization. However, authors did not analyze the effect of forecasting error on the cost function.

To address the mentioned challenges and to realize the economic and reliable operation of DGs in a microgrid, a decentralized optimization strategy for DGs power allocation basing on load demand–power generation equivalent forecasting is proposed in this paper. The main contributions of this paper are illustrated as follows.

- All processes and parameters of the proposed decentralized optimization strategy are designed in a fully distributed way, which do not need any center or leader to assign global variables or parameters. Each DG only needs to communicate with its neighbors under sparse communication to reach the optimal results.
- A load demand–power generation equivalent forecasting method is proposed to improve the decentralized optimization strategy, which replaces the load prediction center and the load sensor devices. The balance constraint only needs information of predicted power output instead of load demand. The historical power generation data, which is used for prediction, has already satisfied the balance constraint of power supply and load demand. When the sum of real power output equals to the sum of predicted output, the balance between power supply and load demand is achieved.
- The uncertainty of renewable generation is considered in the cost modeling process to optimize the expense comprehensively. A penalty term is introduced into the cost model for adjustment of the expense cost by the fluctuation of renewable energy sources and the forecasting error.
The remainder of this paper is outlined as follows. Section 2 introduces the DGs power allocation problem formulation and mathematical model. In section 3, the decentralized optimization algorithm is presented and proved to be convergent. In Section 4, the accuracy and characteristic of the proposed method are verified through simulations, and compared with a centralized method. Then, the daily optimization results are analyzed to verify the effectiveness of the strategy. Finally, findings are concluded in Section 5.

2. Load Demand–Power Generation Equivalent Forecasting Based Distributed Generators (DGs) Power Allocation Strategy Formulation

The structure of DGs power allocation strategy, including construction, optimization objective, and algorithm, are shown in Figure 1.

1. As mentioned in Section 1, the centralized optimization algorithms deeply depend on the control center, which requires collecting global information from the whole network. With the rapid increasing number of DG units, these approaches are not suitable for a large-scale network, due to significance communications and massive computational overhead. On the contrary, decentralized optimization algorithms mainly rely on consensus or multi-agent systems, which merely need sparse communication with neighbor DGs to achieve information interaction and cooperation [18]. Therefore, the decentralized mode is more suitable to implement the optimization.

![Figure 1: Structure of the microgrid system and optimization strategy.](image)

2. The optimization problem is aiming at the minimum cost of the outputs of DGs. When the total outputs of DGs is determined, a ‘Power Allocation Strategy’ need to be studied to obtain the output of each DG unit. Thereby, we can get the accurate dispatching plan of each DG unit. The Power Allocation Strategy consists of the optimization objective and the optimization algorithm.

3. A load demand–power generation equivalent forecasting method is proposed in the optimization algorithm, which only needs the information of predicted power output instead of load demand. The historical power generation data, which is used for prediction, has already satisfied the balance constraint of power supply and load demand. Therefore, when the balance between the real power output and the predicted output is gained, the equation constraint of power supply and load demand is satisfied. Thereby, the load demand information could be substituted by the predicted power generation in a local way.

4. Due to the uncertainty and fluctuation of the renewable energy sources, the prediction error is considered when gaining the predicted value for fitting a feasibility interval, and the real values of the predicted day are added to the training data to update the training process for a rolling
optimization. Before establishing the model of optimization, the prediction method needs to be analyzed to support the optimization strategy.

2.1. Generation Output Prediction Interval Model

Since the short-term generation prediction of microgrid is a prediction of time series, Long Short-Term Memory Network (LSTM) is more suitable for generation prediction and is used for the prediction of DGs power output in this section. Take the short-term prediction neural network model of PV generation as an example. The impact factors of Photovoltaic (PV) generation, which are treated as input vectors, contain the historical power, light intensity, temperature, and humidity, as well as the temperature and humidity of the predicted day. Meanwhile, the final output of the prediction model is the predicted power of PV, with only one type of data. The neural network model is presented in Figure 2.

\[
\begin{align*}
    &\text{Historical power} \quad p_1 \quad p_2 \quad p_3 \quad p_4 \quad p_5 \quad p_6 \\
    &\text{Historical light intensity} \quad h_1 \quad h_2 \quad h_3 \quad h_4 \quad h_5 \quad h_6 \\
    &\text{Historical temperature} \quad d_1 \quad d_2 \quad d_3 \quad d_4 \quad d_5 \quad d_6 \\
    &\text{Predicted day temperature} \quad d_1 \quad d_2 \quad d_3 \quad d_4 \quad d_5 \quad d_6\end{align*}
\]

where \( x_{f}^{d} \) is represent the prediction output value. The input of model from \( P_1 \) to \( P_6 \) are historical power, light intensity, temperature, humidity and the temperature, humidity of the predicted day, respectively.

According to the above model, PV generation can be predicted. The prediction models for wind turbines, fuel turbines, and hydro turbines are built in a similar way. Some parameters for prediction of a typical region are shown in Figure 3. Figure 3a shows the annual historical temperature in the data collecting area. The annual output power of PV and WT are shown in Figure 3b, c (The data used is one sampling point per hour). It can be seen that the tested region is abundant with renewable energy resources and is suitable for the application of a microgrid system.

\[
\begin{align*}
    &\text{Annual historical temperature (a), Annual output power of PV (b), Annual output power of Wind Turbine (WT) (c).}
\end{align*}
\]
Figure 4. PV output per hour (a), WT output per hour (b), WT Prediction and actual output of 1500 data (c).

The predicted output of WT and PV on a typical day is obtained as shown in Figure 4a, b. We can clearly see that there exists error between the predicted value and the real value. Due to the high randomness and volatility of WT and PV output, the accuracy of prediction plays a vital role in the optimization process [19]. Thus, we take the forecasting error into consideration for a comprehensive optimization. As the fluctuation of WT is stronger, we chose 1500 data of the aimed device to test, which are shown in Figure 4c. We chose the predicted value $\hat{x}_{i,t}$ as a reference value. As the existence of prediction error, the predicted value may not be the same as the real one. By comparing the error between the prediction value and the real one, it can be found that the error obeys a distribution law. According to the data distribution and the model in [17], we can assume that the error obeys a Gaussian distribution, which is defined as the mismatch between the real value and the predicted value $\Delta \hat{x}_{i,t} = x_{i,t} - \hat{x}_{i,t}$. Then, the probability density function can be modeled as follows.

$$f(\Delta \hat{x}_{i,t}) = \frac{1}{\sqrt{2\pi} \sigma_{i,t}} \exp\left(-\frac{(\Delta \hat{x}_{i,t})^2}{2\sigma_{i,t}^2}\right),$$  

where $\sigma_{i,t}$ is the variance; let $100(1 - \alpha_{i,t})\%$ represent confidence level, and $\alpha_{i,t}$ is usually chosen as 0.05.

Based on the probability density function, we can get the corresponding confidence interval $[\Delta \hat{x}_{i,t}^{\text{low}}, \Delta \hat{x}_{i,t}^{\text{up}}]$. When the prediction error is considered, the predicted output turns to be as follows.

$$\tilde{x}_{i,t} = \tilde{x}_{i,t}^{\text{up}} + \Delta \tilde{x}_{i,t}, \quad \tilde{x}_{i,t}^{\text{low}} \leq \tilde{x}_{i,t} \leq \tilde{x}_{i,t}^{\text{min}},$$

$$\tilde{x}_{i,t}^{\text{up}} = \Delta \tilde{x}_{i,t}^{\text{up}} + \tilde{x}_{i,t}^{\text{up}},$$

$$\tilde{x}_{i,t}^{\text{low}} = \Delta \tilde{x}_{i,t}^{\text{low}} + \tilde{x}_{i,t}^{\text{low}},$$

where, $\tilde{x}_{i,t}$ is the $i$th DG predicted value (largest possibility value) when considering the prediction error. $\tilde{x}_{i,t}^{\text{min}}$ and $\tilde{x}_{i,t}^{\text{max}}$ are the lower and upper boundaries of $\tilde{x}_{i,t}$.

Figure 5. The prediction error of WT (a), the error distribution in p. u. (b), comparison of WT output (c).
The error between the real value and the predicted value (1500 data) is presented in Figure 5a. From the analysis, the error distribution of WT obeys a Gaussian distribution and is treated as a p. u. value in Figure 5b. When the prediction error is considered, the output of WT becomes a possible range in Figure 5c.

2.2. Optimization Model and Formulation

2.2.1. Optimization Objective

In this section, the optimization objective is set as the minimum of the total operating cost. Thereby, the cost function is formulated as (3), which mainly includes fuel turbines (Fes), wind turbines (WTs), photovoltaic sets (PVs), and hydropower turbines (HTs).

\[
\min F = \min \left( \sum_{i=1}^{N} F_i(x_{it}) \right) = \sum_{i=1}^{N_1} F_i(x_{it}^{pv}) + \sum_{i=1}^{N_2} F_i(x_{it}^{w}) + \sum_{i=1}^{N_3} F_i(x_{it}^{fe}) + \sum_{i=1}^{N_4} F_i(x_{it}^{HT})
\]  

(3)

where, \( F \) is the cost of generation, which mainly contains the cost of real outputs of DG units; \( F_i(x_{it}) \) is the cost of \( i \)th DG output; \( x_{it} \) is the real output of \( i \)th DG at time \( t \); \( x_{it}^{pv}, x_{it}^{w}, x_{it}^{fe} \) and \( x_{it}^{HT} \) are the real output of each PV, WT, Fe, and HT at time \( t \) in a microgrid, respectively; \( N_1, N_2, N_3 \) and \( N_4 \) are the number of each DGs, respectively. \( N = N_1 + N_2 + N_3 + N_4 \).

2.2.2. Cost Functions and Constraints of DGs

In order to achieve the optimization objective in 2.2.1, the cost function of each DG is established as (4)–(7). Each DG is supposed to operate within the equation constraint of its own capacity and to keep the balance between load demand and power supply in the system.

The cost function of PV is defined as follows.

\[
F_i(x_{it}^{pv}) = a_i x_{it}^{pv} + \epsilon_i \frac{(\bar{x}_{it}^{pv} - x_{it}^{pv})^2}{\bar{x}_{it}^{pv} - \underline{x}_{it}^{pv}},
\]  

(4)

where \( F_i(x_{it}^{pv}) \) is the \( i \)th PV generation cost and \( x_{it}^{pv} \) is the real output value of the \( i \)th PV at time \( t \); \( a_i \) is the cost coefficient; \( \epsilon_i \) is the penalty factor; \( \bar{x}_{it}^{pv} \) is the prediction value (considering prediction error) of the \( i \)th PV at the moment \( t \); \( \underline{x}_{it}^{pv} \) and \( \bar{x}_{it}^{pv} \) are the upper and lower boundaries of \( \bar{x}_{it}^{pv} \).

The cost function of WT is defined as follows.

\[
F_i(x_{it}^{w}) = b_i x_{it}^{w} + \gamma_i \frac{(\bar{x}_{it}^{w} - x_{it}^{w})^2}{\bar{x}_{it}^{w} - \underline{x}_{it}^{w}},
\]  

(5)

where, \( F_i(x_{it}^{w}) \) is the \( i \)th WT generation cost and \( x_{it}^{w} \) is the real output value of the \( i \)th WT at time \( t \); \( b_i \) is the cost coefficient; \( \gamma_i \) is the penalty factor; \( \bar{x}_{it}^{w} \) is the prediction output value of the \( i \)th WT at the moment \( t \). \( \underline{x}_{it}^{w} \) and \( \bar{x}_{it}^{w} \) are the upper and lower boundaries of \( \bar{x}_{it}^{w} \).

Remark 1. Due to the randomness and volatility of the WT and PV generation, the accuracy of the prediction results has a great impact on the optimization objective. Therefore, a penalty term is introduced into the model for adjustment. Specifically, the smaller deviation between the real value and the predicted value is, the smaller impact the penalty term gains. For example, when \( x_{it} = \bar{x}_{it} \), the penalty term is zero, which means the predicted value is identical to the real value. In (4), the cost
function consists of the operation cost $a_i x_{pv,i}^m$ and punishment of the deviation $e_i \left( \frac{x_{pv,i}^m - \tilde{x}_{pv,i}^m}{\tilde{x}_{pv,i}^m} \right)^2$. Since the mismatch between the predicted value and the real value could be positive or negative, both two possible scenarios have an effect on the expense. Therefore, $\left( \tilde{x}_{pv,i}^m - x_{pv,i}^m \right)^2$ is used to represent the error effect, which is easy to calculate as well. $\tilde{x}_{pv,i}^m - \tilde{x}_{pv,i}^m$ is the confidence interval of the predicted value as a constant. ($5$) is the same as ($4$).

The cost function of Fe is expressed as follows.

$$F_i(x_{fe,i}^e) = d_i^{fe} (x_{fe,i}^e)^2 + e_i^{fe} x_{fe,i}^e + f_i^{fe},$$

where, $F_i(x_{fe,i}^e)$ is the $i$th Fe generation cost and $x_{fe,i}^e$ is the real output of the $i$th Fe at time $t$; $d_i^{fe}$, $e_i^{fe}$, and $f_i^{fe}$ are the cost coefficients.

The cost function of HT is expressed as follows.

$$F_i(x_{ht,i}^e) = d_i^{ht} (x_{ht,i}^e)^2 + e_i^{ht} x_{ht,i}^e + f_i^{ht},$$

where, $F_i(x_{ht,i}^e)$ is the $i$th HT generation cost and $x_{ht,i}^e$ is the real output of the $i$th HT at time $t$; $d_i^{ht}$, $e_i^{ht}$, and $f_i^{ht}$ are the cost coefficient. As Fe and HT are stable equipment, the prediction effort will not wildly influence their total expense. Hence, no more punishment is needed.

In order to simplify the solution process, the above cost function is commonly selected as a quadratic polynomial function, which can be obtained by data fitting. Meanwhile, the following supply–demand balance equation constraint and capacity inequality constraints should be satisfied during optimization. The formula is identified as follows.

$$\sum_{i=1}^{N_i} (x_{pv,i}^m) + \sum_{j=1}^{N_j} (x_{w,i}^m) + \sum_{k=1}^{N_k} (x_{fe,i}^e) + \sum_{l=1}^{N_l} (x_{ht,i}^e) = \sum_{i=1}^{N_i} (\tilde{x}_{pv,i}^m) + \sum_{j=1}^{N_j} (\tilde{x}_{w,i}^m) + \sum_{k=1}^{N_k} (\tilde{x}_{fe,i}^e) + \sum_{l=1}^{N_l} (\tilde{x}_{ht,i}^e),$$

where $\tilde{x}_{pv,i}^m$, $\tilde{x}_{w,i}^m$, $\tilde{x}_{fe,i}^e$ and $\tilde{x}_{ht,i}^e$ are the prediction output of the $i$th PV, WT, Fe, and HT at time $t$, respectively.

**Remark 2.** In this paper, the balance constraint equality only needs information of predicted power output instead of load demand. The historical power generation data, which is used for prediction, has already contained the balance constraint of power generation supply and load demand. Therefore, when the balance between the real power output and the predicted output is gained, the equation constraint of power supply and load demand is satisfied. Considering the prediction error, the upper and lower boundaries of the equation are calculated to provide a probability interval for the optimization results.

In addition, the cost function should meet the following inequality constraints as well.

**PV capacity constraint:** $x_{pv,i}^m \leq x_{pv,i}^m \leq x_{pv,i}^m,

**WT capacity constraint:** $x_{w,i}^m \leq x_{w,i}^m \leq x_{w,i}^m,

**Fe capacity constraint:** $x_{fe,i}^e \leq x_{fe,i}^e \leq x_{fe,i}^e,

**HT capacity constraint:** $x_{ht,i}^e \leq x_{ht,i}^e \leq x_{ht,i}^e$.

2.3. Optimization Solution

According to the optimal objective (3) and equality constraint (8), the Lagrangian function can be obtained as follows.
\[ L = \min \left( \sum_{j=1}^{N_i} F_i(x_{i,j}) \right) + \lambda_i \left[ \sum_{j=1}^{N_i} (x_{i,j}^{pp}) + \sum_{j=1}^{N_i} (x_{i,j}^{w}) + \sum_{j=1}^{N_i} (x_{i,j}^{fe}) + \sum_{j=1}^{N_i} (x_{i,j}^{HT}) - \sum_{j=1}^{N_i} (x_{i,j}^{pp}) - \sum_{j=1}^{N_i} (x_{i,j}^{w}) - \sum_{j=1}^{N_i} (x_{i,j}^{fe}) - \sum_{j=1}^{N_i} (x_{i,j}^{HT}) \right] \]  

where \( L \) is the Lagrangian function and \( \lambda_i \) is the Lagrangian multiplier of each DG.

Let \( D = \sum_{j=1}^{N_i} (x_{i,j}^{pp}) + \sum_{j=1}^{N_i} (x_{i,j}^{w}) + \sum_{j=1}^{N_i} (x_{i,j}^{fe}) + \sum_{j=1}^{N_i} (x_{i,j}^{HT}) \). Then, the optimal conditions are provided by:

\[
\begin{align*}
\frac{\partial F}{\partial x_{i,j}^{pp}} &= \frac{\partial F_i(x_{i,j}^{pp})}{\partial x_{i,j}^{pp}} - \lambda_i = 0 \\
\frac{\partial F}{\partial x_{i,j}^{w}} &= \frac{\partial F_i(x_{i,j}^{w})}{\partial x_{i,j}^{w}} - \lambda_i = 0 \\
\frac{\partial F}{\partial x_{i,j}^{fe}} &= \frac{\partial F_i(x_{i,j}^{fe})}{\partial x_{i,j}^{fe}} - \lambda_i = 0 \\
\frac{\partial F}{\partial x_{i,j}^{HT}} &= \frac{\partial F_i(x_{i,j}^{HT})}{\partial x_{i,j}^{HT}} - \lambda_i = 0 \\
\frac{\partial F}{\partial \lambda_i} &= D - \left( \sum_{j=1}^{N_i} (x_{i,j}^{pp}) + \sum_{j=1}^{N_i} (x_{i,j}^{w}) + \sum_{j=1}^{N_i} (x_{i,j}^{fe}) + \sum_{j=1}^{N_i} (x_{i,j}^{HT}) \right) = 0
\end{align*}
\]  

According to (11), we can get the functional relationship between the optimal output \( x_{i,j} \) and the Lagrangian multiplier \( \lambda_i \) that

\[
\frac{\partial F_i(x_{i,j}^{pp})}{\partial x_{i,j}^{pp}} = \frac{\partial F_i(x_{i,j}^{w})}{\partial x_{i,j}^{w}} = \frac{\partial F_i(x_{i,j}^{fe})}{\partial x_{i,j}^{fe}} = \frac{\partial F_i(x_{i,j}^{HT})}{\partial x_{i,j}^{HT}} = \lambda_i.
\]  

Therefore, the Lagrangian multiplier \( \lambda_i \) can be treated as the consensus variable, which communicates between each DG. Moreover, when ignoring the inequality constraints, there is a linear function relationship between \( \lambda_i \) and \( x_{i,j} \), which can be represented as \( \phi(\lambda_i) = x_{i,j} \). When each DG is operating in the optimal consistent condition \( \lambda_i^* \), the optimal power output \( x_{i,j}^* \) can be obtained [20].

When considering the inequality constraints, the optimal conditions are written as follows.

\[
\phi^*(\lambda_i) = \begin{cases} 
\bar{x}_{i,j}, & \text{if } \lambda_i > \bar{\lambda}_i \\
x_{i,j}, & \text{if } \lambda_i \leq \bar{\lambda}_i \leq \bar{\lambda}_i \\
\tilde{x}_{i,j}, & \text{if } \lambda_i < \tilde{\lambda}_i 
\end{cases}
\]

where \( ^* \) means the optimal condition; \( \bar{x}_{i,j} \) and \( \tilde{x}_{i,j} \) represent \( \max_{x_{i,j}} \) and \( \min_{x_{i,j}} \), respectively. \( \bar{\lambda}_i = \phi^{-1}(\bar{x}_{i,j}) \) and \( \tilde{\lambda}_i = \phi^{-1}(\tilde{x}_{i,j}) \).

3. The Decentralized Optimization Method Analysis

In this section, the decentralized optimization method is proposed to deal with the problem formulated in section 2. The method is presented without capacity constraints firstly, and generalized to the situation with constraints. Then, the proof of convergence and the flow of the proposed algorithm are explained in this section. Before elaborating details of the proposed method, some preliminary knowledge is presented briefly.
3.1. Preliminary Knowledge

3.1.1. Graph Theory

Considering a microgrid with \( N \) DGs, a digraph \( \zeta = (\nu, \epsilon, A) \) is used to model the network topology of the system based on the communication infrastructure [21]. The digraph consists of a nonempty finite nodes set \( \nu = \{1, \ldots, i, \ldots, N\} \), and a finite edges set \( \epsilon = \{(i, j)\} \subseteq \nu \times \nu \) which means DG \( j \) can receive information from DG \( i \). \( A \) is the adjacency matrix. When every node in the digraph is reachable from others, \( \zeta \) is said to be strongly connected. Moreover, for each node \( i \in \nu \), the sets of in-neighbors are denoted as \( \mathcal{N}^+_i = \{ j \in \nu \mid (j, i) \in \epsilon \} \), and out-neighbors are \( \mathcal{N}^-_i = \{ j \in \nu \mid (i, j) \in \epsilon \} \). Note that \( \mathcal{N}^+_i \neq \mathcal{N}^-_i \) generally. Let \( d^+_i = |\mathcal{N}^+_i| \) and \( d^-_i = |\mathcal{N}^-_i| \) be the in-degree and out-degree of node \( i \), respectively. In a strongly connected digraph \( d^+_i \neq 0 \) and \( d^-_i \neq 0 \).

3.1.2. The Design of Communication Weight

Let us define two matrices \( W, U \in \mathbb{R}^{N \times N} \) associated with the digraph \( \zeta = (\nu, \epsilon, A) \) as follows.

\[
W_{ij} =
\begin{cases}
\frac{1}{d^+_i}, & \text{if } j \in \mathcal{N}^+_i \\
0, & \text{otherwise}
\end{cases}
\quad \forall i, j \in \nu.
\]

\[
U_{ij} =
\begin{cases}
\frac{1}{d^-_i}, & \text{if } j \in \mathcal{N}^-_i \\
0, & \text{otherwise}
\end{cases}
\quad \forall i, j \in \nu.
\]

(14)

Based on the definition of communication matrices, it is clear to gain that \( W_{ij} \) is a row stochastic matrix and \( U_{ij} \) is a column stochastic matrix. Note that the convergence will not be affected by the communication weights selection process when the above stochastic characteristics are satisfied [20]. Note \( \rho(W) = \rho(U) = 1 \), where \( \rho(\cdot) \) represents the spectral radius of the matrix.

3.1.3. Assumptions and Lemmas

The proposed method should be subjected to the following assumption and lemmas, for convergence.

**Assumption 1.** The digraph \( \zeta = (\nu, \epsilon, A) \) is strongly connected.

When the digraph is strongly connected, every node in the digraph is reachable from others. Then, it can make sure that the global accessibility of the information which requires communication.

**Lemma 1** [8]. The weight matrices \( W, U \) satisfy the following:

1. \( W \) is a right eigenvector of all \( W \), \( W 1 = 1 \) and \( \lim_{k \to \infty} W^k = 1 w^T, w > 0 \) and \( 1^T w = 1 \).

2. \( U \) is a left eigenvector of all \( U \), \( 1^T U = 1^T \) and \( \lim_{k \to \infty} U^k = \mu 1^T, \mu > 0 \) and \( 1^T \mu = 1 \).

where \( 1 \) denotes a vector with all elements being 1 at length \( N \).

**Lemma 2.** The cost function \( F \) of each DG is twice continuously differentiable in \( \mathbb{R} \) (the set of real number) and the second derivative is a real number \( i_1 \).

\[
\frac{\partial F^2(x_{ij})}{\partial x_{ij}^2} = i_1, \forall x_{ij} < x_{i_1} < x_{i_2}.
\]

(15)

**Proof of Lemma 2.** According to (4)-(7), it can be verified that (6) and (7) meet lemma 2. Then, the cost function (4) consists of the operation cost and the punishment of deviation. Obviously, the operation cost \( a_i x_{i_1}^{pv} \) is consistent with lemma 2. For the punishment of deviation, \( \tilde{x}_{ij}^{pv} \) and \( \epsilon_i \) are treated as constants. Then, Equation (4) can be treated as quadratic polynomials, which is submitted to lemma 2 as well. (Equation (5) is the same as (4)).
3.2 Consensus-Based Optimization Algorithm

Definition:

\[
\lambda_i(k+1) = \lambda_i(k) + \sum_{j \in \mathcal{N}_i(k)} W_{ij}(\lambda_j(k) - \lambda_i(k)) + \eta y_i(k)
\]

\[
x_i(k+1) = \phi(\lambda_i(k+1))
\]

\[
y_i(k+1) = \sum_{j \in \mathcal{N}_i(k)} U_{ij} y_j(k) - (x_i(k+1) - x_i(k))
\]

where \( \lambda_i(k+1) \) is the consensus variables of the communication between each DG; \( x_i(k+1) \) is the output of DG; \( y_i(k+1) \) is the mismatch between the total predicted output and the total real output, which is used as a feedback to adjust the iteration dynamically. \( W_{ij} \) and \( U_{ij} \) are the communication weights defined in (14). \( \eta \) is a sufficiently small positive constant gain for mismatch term.

Remark 3. As \( W_{ij} \) is a row stochastic matrix, the iteration could be written as

\[
\lambda_i(k+1) = \sum_{j \in \mathcal{N}_i(k)} W_{ij}(\lambda_j(k)) + \eta y_i(k).
\]

Although the two forms are mathematically equivalent, they have different physical interpretations. The form presented in (16) means that the state of consensus variable \( \lambda_i \) at \( k+1 \) is decided by not only the state of its neighbors at \( k \), but also its own previous state at \( k \).

Initialization:

\[
\begin{align*}
x_i(0) &= \Pi \\
\lambda_i(0) &= \Pi' \\
y_i(0) &= \begin{cases} 
D_i - x_i(0), & \text{if } i \in \mathcal{N}_i^- \\
-x_i(0), & \text{if } i \notin \mathcal{N}_i^-
\end{cases}
\end{align*}
\]

where \( D_i = \frac{D}{N} \), \( D \) is defined as the total value of the predicted output in section 2. The estimation method enables the algorithm to get rid of a leader and the initialized value can be any fixed admissible values \( \Pi \) and \( \Pi' \), without impact on the convergence.

3.3. Proof of Convergence

To analyze the convergence, rewrite (16) as the following matrix form:

\[
\begin{align*}
\lambda(k+1) &= W(\lambda(k) + \eta y(k)) \\
x(k+1) &= \bar{L}(\lambda(k+1)) + \tau \\
y(k+1) &= Uy(k) - (x(k+1) - x(k))
\end{align*}
\]

where, \( x(k+1) = \bar{L}(\lambda(k+1)) + \tau \) is the linear relationship form of \( x_i(k+1) = \phi(\lambda_i(k+1)) \). Therein, \( \bar{L} \) is a diagonal matrix with elements \( l_i \), and \( \bar{L} = \text{diag}[l_1, l_2, \ldots, l_N] \). \( \tau \) is a constant vector. \( \lambda, x, \) and \( y \) are the column stack vectors of \( \lambda_i, x_i, \) and \( y_i \).

According to Lemma 1, it can be verified that:

\[
1^T y(k+1) = 1^T Uy(k) - 1^T (x(k+1) - x(k)) = 1^T y(k) - 1^T x(k+1) + 1^T x(k).
\]

Therefore, \( 1^T (y(k+1) + x(k+1)) = 1^T (y(k) + x(k)) \), which means that the sum of the mismatch and the real output keep constant with the predicted value. Then, replace all \( x \) with \( \hat{\lambda} \) (19) turns to be:
\begin{align}
\lambda(k+1) &= (\Gamma + \eta \Omega) y(k+1) \\
y(k+1) &= y(k)
\end{align}
(19)

where, 
\[
\Gamma = \begin{bmatrix}
W & 0 \\
\Lambda(I-W) & U
\end{bmatrix},
\Omega = \begin{bmatrix}
0 & I \\
0 & -\Lambda
\end{bmatrix}.
\]

It is easy to know that the eigenvalues of \( \Gamma \) is the union of \( W \) and \( U \). According to lemma 1, \( \lambda_1 = \lambda_2 = 1 \). \( \Gamma \) has two eigenvalues which can be defined as \( \sigma_1 = \sigma_2 = 1 \). According to eigenvalue perturbation theory [22], two following matrices are constructed.

\[
M = \begin{bmatrix}
\Gamma^T & \Gamma^T \\
w^T & 0
\end{bmatrix},
V = \begin{bmatrix}
0 & 1 \\
\mu & \mu
\end{bmatrix}
\]
(20)

where, \( l = \sum_{i=1}^{N} l_i \).

Then, we can gain that \( MV = I \). The variation of \( \sigma_1, \sigma_2 \) can be quantified by the eigenvalues of \( M \Omega V \), and

\[
M \Omega V = \begin{bmatrix}
0 & 0 \\
w^T \mu & -lw^T \mu
\end{bmatrix}
\]
(21)

We can get the conclusion \( d\sigma_1 / d\eta = 0, d\sigma_2 / d\eta = -lw^T \mu < 0 \) are two eigenvalues of \( M \Omega V \), which proves that \( d\sigma_1 \) is not changing against \( \eta \), and when \( \eta > 0 \), \( d\sigma_2 \) becomes smaller. Therefore, there exists an upper boundary of \( \eta \), enabling \( |\sigma_2| < 1 \). Hence, all the rest of the eigenvectors of the matrix are within the open unit disk. It can be proved that \( [1^T, 0^T]^T \) is the eigenvector of matrix \( \Gamma + \eta \Omega \). Then we can get \( y_i(k) \to 0, k \to \infty \) and \( \lambda_i \) converges to a common value.

3.4. Generalization to Constrained Condition

When considering the inequality constraints, the algorithm (16) turns to the following form.

\[
\lambda_i(k+1) = \lambda_i(k) + \sum_{j \in N_i^+(k)} W_{ij} (\lambda_j(k) - \lambda_i(k)) + \eta y_i(k)
\]
(22)

\[
x_i(k+1) = \begin{cases}
\bar{x}_i, & \text{if } \lambda_i > \bar{\lambda}_i \\
x_i, & \text{if } \lambda_i \leq \bar{\lambda}_i \leq \bar{\lambda}_i \\
\underline{x}_i, & \text{if } \lambda_i < \underline{\lambda}_i
\end{cases}
\]

\[
y_i(k+1) = \sum_{j \in N_i^+(k)} U_{ij} y_j(k) - (x_i(k+1) - x_i(k))
\]

where the relationship between \( \lambda_i(k+1) \) and \( x_i(k+1) \) is treated as (13). The initial value can be set to any admissible value as well.

**Initialization:**

\[
x_i(0) = \begin{cases}
\underline{x}_i, & \text{if } \underline{x}_i < D_i \\
D_i, & \text{if } \underline{x}_i < D_i < \bar{x}_i \\
\bar{x}_i, & \text{if } D_i > \bar{x}_i
\end{cases}
\]

\[
\lambda_i(0) = \phi^{-1}(x_i(0))
\]

\[
y_i(0) = D_i - x_i(0)
\]
(23)

\( \forall i \in \mathcal{V} \)
When $\sum_{i=1}^{N} x_i \leq D \leq \sum_{i=1}^{N} \bar{x}_i$, at least one DG is not saturated. There still exists an optimization solution in this condition. Thus, this paper only considers the saturated case, which means taking into account the outputs as upper boundaries, where $\sum_{i=1}^{N} y_i(k) = D - \sum_{i=1}^{N} x_i(k)$ is the mismatch between the total generation supply and the load demand. $\eta$ is a feedback gain. $\lambda_i(k)$ will increase as $k$ updates and converges to the same value. As the total power generation is a monotone increasing function of the incremental cost, the total amount of generation will increase as well. Therefore, the feedback item will reduce the mismatch. In this process, some DGs may reach its maximum capacity. Then, the DG output holds saturation after it gets saturated or after a sufficiently long time. Therefore, the communication weights of saturated DG is equivalent to $l_i = 0$, and those unsaturated DGs will still participate in feedback adjustment until $y_i(k) \rightarrow 0, k \rightarrow \infty$ achieves convergence of the algorithm.

3.5. Solution Algorithm Flow

According to the discussion above, the algorithm flow is generated as shown in Figure 6. The specific steps are as follows.

**Figure 6.** Optimization algorithm flow.

Firstly, based on the multi-type data, the LSTM method is used to predict the generation of DGs locally.

Step 1. Collect historical generation data, regional meteorological information of the tested regional microgrid, including the historical power, light intensity, wind speed, temperature, and humidity. The sampling period is chosen as one hour a time, which collected a total of 24 points per day.

Step 2. Preprocess the collected data to form the training data set, which replaces outliers through the adjacent daily interpolation.
Step 3. Train the LSTM neural network model annually with training data set to ensure the LSTM neural network is retrained every year.

Step 4. Collect the information of the predicted day, including the light intensity, wind speed, temperature, and humidity of the predicted day. Use the trained LSTM and the information of the predicted day to predict the local generation of the DG in the regional micro grid including WT, PV, Fe, and HT. The prediction error is considered when gaining the predicted value for fitting a feasibility interval as illustrated in (2), and the real values of the predicted day are added to the training data to update the training process for a rolling optimization.

Secondly, combined with the results of power generation prediction, the decentralization optimization model is established. Moreover, the optimization model does not depend on the load information, which only needs the predicted output. Therefore, local operations can be implemented without a control center.

Step 5. According to the solution of optimization method proposed in (12), define the consensus variable.

Step 6. Initialization—set the initial value of the optimization variable within the valid range.

Step 7. Enter the iterative calculation process. Define the communication weights required for the iteration according to the definition described in (14).

Step 8. According to the iteration strategy (22) update the optimal variables.

Step 9. Determine whether the result is within the constraint or not. If the result is beyond the limit, modify the variables under its constraints and go back to step 5; otherwise, go to the next step.

Step 10. Criticize whether the iteration is end or not. If the mismatch meets the convergence condition, stop the iteration and print the output; otherwise, continue to iterate until convergence.

4. Simulation and Case Study

In this section, several simulations under a text system with six DG units—A1, A2 ... A6 are used to verify the accuracy of the proposed method. The communication topology is shown in Figure 7. Some parameters of DGs are listed in Table 1. The simulation results are obtained under MATLAB environment.

| DG number | \( \alpha_i / \beta_i / d_i^{Fe} / d_i^{HT} \) | \( \varepsilon_i / \gamma_i / e_i^{Fe} / e_i^{HT} \) | \( f_i^{Fe} / f_i^{HT} \) | \( \bar{X}_{Pg} \) (kW) | \( \bar{X}_{Pg} \) (kW) |
|-----------|---------------------------------|---------------------------------|----------------|-----------------|-----------------|
| A1 HT1    | 0.0184                          | 3.5                             | 15             | 180             | 60              |
| A2 Fe1    | 0.1080                          | 2.0                             | 10             | 130             | 40              |
| A3 PV     | 0.0087                          | 1                               | /              | 150             | 0               |
| A4 WT     | 0.0040                          | 0.1                             | /              | 160             | 0               |
| A5 HT2    | 0.0174                          | 2.9                             | 20             | 200             | 85              |
| A6 Fe2    | 0.1250                          | 1.8                             | 25             | 80              | 20              |

A 1-6 are Hydro Turbine 1 (HT1), Fuel Turbine 1 (Fe1), Photovoltaic (PV), Wind Turbine (WT), Hydro Turbine 2 (HT2), and Fuel Turbine 2 (Fe2), respectively.

Figure 7. The communication topology of Distributed Generators (DGs).
Based on Figure 7, the matrices $W$ and $U$ can be defined as

$$ W = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}, \quad U = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} $$

respectively.

4.1. Verification of the Proposed Method

4.1.1. Verification of the Convergence

In this case, the convergence of the proposed method (22) is tested. As the time-scale is an hour, we chose an hour of a day randomly to operate the simulation. The simulation results—the optimal consensus variable $\lambda_i$, mismatch $y_i$, DGs output power $x_i$, and the sum of $x_i$ are presented in Figure 7.

When the total predicted output $D = 440$ kW, the optimal solution is approximately $\lambda^* = 8.65$. Obviously, all consensus variables (Lagrange multipliers) converge to the optimal value through the proposed algorithm. As shown in Figure 8a, they rise monotonously non-decreasing until the upper boundary (Note that $\lambda$ represents the incremental cost which is affected by the coefficient in the cost function). Figure 8b shows values of the feedback (mismatch) $y$ of all DGs convergent to zero which proves that the mismatch is decayed with iteration. Figure 8c indicates the optimal power output. As shown in Figure 8d, the sum of the outputs of all DGs keep in balance with the total predicted value which meets the supply and demand constraints.

![Figure 8](image)

**Figure 8.** The optimal consensus variable $\lambda_i$ (a), Mismatch $y$ (b), DGs output power $x$ (c), the sum of $x$ (d).

4.1.2. Verification of the Accuracy

To verify the accuracy of the proposed method, we chose the results of two single times to compare with the results achieved through the centralized method (quadratic programming). To gain a clearer observation, the values are presented in Table 2. It can be seen that the result of the proposed
method is very close to the centralized method. Meanwhile, the sum of DGs output equals to total prediction, which satisfies the balance condition proposed in section 2. Thus, the optimization maintains the balance between power supply and load demand.

**Table 2. Results comparison.**

| DGs serial number | Time 1 | Time 2 |
|-------------------|--------|--------|
|                   | M1     | M2     | P1   | M1    | M2    | P1   |
| A1 HT1            | 139.99 | 139.98 | 121.26 | 180   | 180   | 175.43 |
| A2 Fe1            | 40     | 40     | 91.04  | 58.79 | 58.95 | 75.8 |
| A3 PV             | 56.26  | 56.26  | 56.26  | 0     | 0     | 0    |
| A4 WT             | 11.08  | 11.08  | 11.07  | 88.32 | 88.32 | 88.32 |
| A5 HT2            | 165.27 | 165.27 | 140.14 | 200   | 200   | 189.44 |
| A6 Fe2            | 27.40  | 27.41  | 20.21  | 51.60 | 51.73 | 50   |
| Total             | 440    | 440    | 439.98 | 578.71 | 579   | 578.99 |

where, M1 is the proposed method; M2 is a central method (a quadratic programming); P1 is the predicted value. Time 1, 2 are two single times of a day. The output unit is kW.

4.2. The Effect of Feedback Gain

The gain of the effect is the parameter that we need to manipulate, which plays an important role in the convergence speed and the performance. If the gain is not selected properly, the convergence will not be guaranteed. Thus, we chose four gains to test the performance, i.e. $\eta = 0.001, 0.005, 0.01,$ and $0.05$. The results are shown in Figure 9.

![Figure 9](image)

Figure 9. The performance of mismatch when $\eta = 0.001$ (a), $0.005$ (b), $0.01$ (c), and $0.05$ (d).

Figure 9 shows that the mismatch $y$ eventually converges to zero within the admissible value range. Noticeably, as $\eta$ increases gradually, the convergence speed rises up. Then, if $\eta$ keeps increasing until well beyond the range, the algorithm will not be convergent. For example, when $\eta \geq 0.1$, the operation loses stability. However, in this paper, the gain $\eta$ needs to have small parameters, which limits the convergence speed. In the future, a wilder parameter selection method and more relaxed convergent conditions will be studied.
4.3. Analysis of Daily Optimization Result on Hourly Time-scale

According to the data of the selected typical regions, the predicted output of DG can be obtained by the method proposed in 2.3, which can be used as the input data of the power balance constraint conditions during the decentralization optimization.

To analysis the distribution of the prediction error, 1500 data of the aimed device were chosen to test, which is shown in Figure 10a. We chose the predicted value as a reference value. By comparing the prediction value and the real one, we can obtain the error. Then, through fitting the data and observe the performance in Figure 10b, it can be found that the error obeys a Gaussian distribution and the error is less that 3.5 percent. When considering the prediction error, the possibility of the prediction values is illustrated in Table 3.

**Table 3. Prediction values of hours considering prediction error in a typical day.**

| DG  | HT1       | Fe1       | PV         | WT         | HT2       | Fe2        |
|-----|-----------|-----------|------------|------------|-----------|------------|
| 4 h | [78.69, 84.83] | [62.08, 66.92] | [0, 1.12]  | [75.79, 81.69] | [96.16, 103.66] | [24.06, 25.94] |
| 5 h | [116.71, 125.84] | [63.32, 68.25] | [0, 1.12]  | [1.73, 1.86]  | [144.37, 155.63] | [20.34, 21.93] |
| 6 h | [163.63, 176.37] | [75.92, 81.84] | [0, 1.12]  | [19.88, 21.43] | [173.35, 186.86] | [29.31, 31.59] |
| 7 h | [173.25, 180] | [76.90, 82.89] | [0, 1.12]  | [28.81, 31.06] | [192.5, 200]    | [38.64, 41.65] |
| 8 h | [163.62, 176.37] | [107.41, 115.77] | [3.40, 3.67] | [39.76, 42.86] | [192.5, 200]    | [56.34, 60.73] |
| 9 h | [163.63, 176.38] | [105.24, 113.44] | [11.73, 12.65] | [39.76, 42.86] | [172.28, 185.71] | [57.88, 62.39] |
| 10 h | [173.25, 180] | [78.56, 84.68] | [25.55, 27.54] | [39.77, 42.85] | [173.25, 186.75] | [72.68, 78.34] |
| 11 h | [173.25, 180] | [100.91, 108.77] | [8.05, 8.56] | [10.66, 11.49] | [192.5, 200]    | [45.48, 49.03] |
| 12 h | [173.25, 180] | [117.02, 126.14] | [34.08, 36.74] | [1.73, 1.86]  | [154, 166]      | [39.66, 42.75] |

Considering the error of prediction, the predicted values become a range rather than a level. In this area, before 6 a.m., the PV does not work. Thus, the results are illustrated from 4 a.m. to 12 p.m.

To analyze the characteristic of the simulation curves, the largest possibility values in the ranges were chosen as an example. The predicted outputs of DGs in a typical day are illustrated in Figure 11. Figure 12 presents the proportion of each DG among the total power generation.
Basing on the predicted outputs of each DG, the optimal power output can be gained by redistribution through the method proposed in 3.2. The results are shown in Figure 13 and Figure 14. As shown in Figure 13, when the prediction error is considered, the optimization results are turned to be a range of interval estimation, which is used to provide advice for day-ahead scheduling policies.

Through the bar chart in Figure 14, the sum of the outputs of each DG is consistent with the total predicted values, which is equivalent to the whole load demand. It can be clearly seen that there is strong uncertainty and volatility of the outputs of renewable energy generation. By contrast, hydro power and fuel turbine generation are more stable, which can be used for supporting the stability of a microgrid system. Therefore, when there is an increase in the proportion of renewable energy such as wind power, maintaining the system stable is one of highlights in current research. Meanwhile, the results are compared in Table 4 to gain a clearer observation.
Table 4. Results comparison of hours in a day.

| DG  | Optimal output (kW) | 4 h   | 5 h   | 6 h   | 7 h   | 8 h   | 9 h   | 10 h  | 11 h  | 12 h  |
|-----|----------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| HT1 | 93.99                | 133.43| 180   | 180   | 180   | 180   | 180   | 180   | 180   | 180   |
| Fe1 | 40                   | 40    | 42.14 | 63.98 | 85.48 | 73.87 | 73.14 | 81.17 | 65.45 |
| PV  | 0                    | 0     | 0     | 0     | 3.54  | 12.19 | 26.54 | 46.81 | 35.41 |
| WT  | 78.74                | 1.79  | 20.65 | 29.94 | 41.32 | 41.31 | 41.31 | 11.07 | 1.8   |
| HT2 | 116.63               | 158.33| 200   | 200   | 200   | 200   | 200   | 200   | 200   |
| Fe2 | 20.63                | 26.44 | 37.21 | 56.08 | 74.66 | 64.62 | 63.99 | 70.93 | 57.35 |
| Total| 349.99               | 359.99| 480   | 530   | 585   | 571.99| 584.98| 589.98| 540.01|

| DG  | Predicted output (kW) | 4 h   | 5 h   | 6 h   | 7 h   | 8 h   | 9 h   | 10 h  | 11 h  | 12 h  |
|-----|------------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| HT1 | 81.76                  | 121.26| 170   | 180   | 170   | 170   | 180   | 180   | 180   | 180   |
| Fe1 | 64.5                   | 65.79 | 78.88 | 79.9  | 111.59| 109.34| 81.62 | 104.84| 121.58|
| PV  | 0                      | 0     | 0     | 0     | 3.54  | 12.19 | 26.55 | 46.82 | 35.41 |
| WT  | 78.74                  | 1.79  | 20.66 | 29.94 | 41.32 | 41.32 | 41.31 | 11.08 | 1.79  |
| HT2 | 99.91                  | 150   | 180.11| 200   | 200   | 179   | 180   | 200   | 160   |
| Fe2 | 25                     | 21.14 | 30.45 | 40.15 | 58.54 | 60.14 | 75.51 | 47.26 | 41.21 |
| D   | 349.91                 | 359.98| 480.10| 529.99| 584.99| 571.99| 584.99| 590   | 539.99|

The simple time is an hour. The data presented is from 4–12 h of one day. Total is the total of optimal output. D is the total of predicted output. The value is taken as the largest possibility value.

In addition, it can be seen from the tendency of lines LD that the sum of the predicted value is consistent with the load demand. Specifically, the peak periods are from 6 a.m. to 12 p.m. and from 5 p.m. to 9 p.m. with larger load demand and power generation. The valley period is from 10 p.m. to 6 a.m. (the next day) and from 1 p.m. to 5 p.m. with smaller load demand and power generation. Meanwhile, to gain clearer observation, the values are presented in Table 4. We can see that the renewable generators WT and PV take the maximum use and the hydro power and the fuel turbines generation to adjust the mismatch flexibly aiming at the optimal objective—the minimum operation cost.

5. Conclusions

In this paper, a decentralized optimal operation method for DGs power allocation based on the equivalent forecasting is proposed. Firstly, a local generation prediction method was introduced into the decentralized optimization strategy to replace the load prediction center or the load sensor device. Based on the multi-type data, LSTM was used to predict the local power generation of each DG. Secondly, the uncertainty and prediction error of renewable generation was taken into account during the cost modeling process to optimize the expense comprehensively. Moreover, the system optimization model was improved based on the predicted power generation results, without load information, so that all process and parameters were designed in a fully distributed way. Then, the algorithm designing procedure was presented and the convergence was proved by eigenvalue perturbation theory. Finally, various cases are simulated and the effectiveness of the proposed method is verified.

However, during the research, there remains some questions requiring further study in the future.

1. More types of equipment, such as energy storage devices and electric vehicles, are supposed to be considered during the operation optimization of a microgrid.
2. Multi-energy utilizations, including cold, heat, power, and gas, should be introduced into the optimization model to fully maximize the benefit.

3. The cost function in practice is more complex instead of a simple polynomial function form. It is usually non-convex, which is difficult to solve numerically. Therefore, a more widely applicable distributed solution algorithm needs to be studied in the future.

Author Contributions: writing—original draft preparation, S.Z.; writing—review and editing, S.Z., Z.Y. (Zilong Yu) and B.Z.; supervision, D.Y. and Z.Y. (Zhile Yang). All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Science Foundation of China (61433004), the Fundamental Research Funds for the Central Universities (N160402003), State Key Laboratory of Alternate Electrical Power Systems with Renewable Energy Sources (LAPS17013), National Natural Science Foundation of China (61703081), the Liaoning Revitalization Talents Program (XLYC1801005), Natural Science Foundation of Liaoning Province (20170520113), and the State Key Laboratory of Alternate Electrical Power System with Renewable Energy Sources (LAPS19005).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Strasser, T.; Andrén, F.; Kathan, J.; Cecati, C.; Buccella, C.; Siano, P.; Mařík, V. A Review of Architectures and Concepts for Intelligence in Future Electric Energy Systems. IEEE Trans. Ind. Electron. 2015, 62, 2424–2438.

2. Cheng, Z.; Duan J, Chow, M. To Centralize or to Distribute: That Is the Question: A Comparison of Advanced Microgrid Management Systems. IEEE Ind. Electron. Mag. 2018, 12, 6–24.

3. Huang, A.; M. L.; Crow, G.T.H. The Future Renewable Electric Energy Delivery and Management (FREEDM) System: The Energy Internet. Proc. IEEE 2011, 99, 133–148.

4. Han, Y.; Zhang, K.; Li, H.; A, E.; Coelho, A.; Guerrero, J.M. MAS-Based Distributed Coordinated Control and Optimization in Microgrid and Microgrid Clusters: A Comprehensive Overview. IEEE Trans. Power Electron. 2018, 33, 6488–6508.

5. Liu, W.; Gu, W.; Sheng, W. Decentralized multi-agent system-based cooperative frequency control for autonomous microgrids with communication constraints. IEEE Trans. Sustain. Energy 2014, 5, 446–456.

6. Divényi, D.; Dán, A.M. Agent-Based Modeling of Distributed Generation in Power System Control. IEEE Trans. Sustain. Energy 2013, 4, 886–893.

7. S. Yang; S. Tan and J. X. Xu. Consensus based approach for economic dispatch problem in a smart grid. IEEE Trans. Power Syst. 2013, 28, 4416–4426.

8. Xu, Y.; Han, T.; Cai, K.; Lin, Z.; Yan, G.; Fu, M.A distributed algorithm for resource allocation over dynamic digraphs. IEEE Trans. Signal Process. 2017, 65, 2600–2612.

9. Ullah, M.H.; Alseyat, A.; Multi-Agent, J.P. System-based Distributed Energy Management in Smart Grid under Uncertainty. In Proceedings of the 2019 IEEE Energy Conversion Congress and Exposition (ECCE); Baltimore, MD, USA, 29 September– 3 October 2019; pp. 3462-3468.

10. Yang, Z.; Wu, R.; Yang, J.; Long, K.; You, P. Economical operation of microgrid with various devices via distributed optimization. IEEE Trans. Smart Grid 2016, 7, 857–867.

11. Zhang, H.; Li, Y.; Gao, D.W.; Zhou, J. Distributed optimal energy management for energy internet. IEEE Trans. Ind. Inform. 2017, 13, 3081–3097.

12. Costa DA, C.; Otto, R.B.; Piardi, A.B.; Ramos, R.A. A Survey of State-of-the-Art on Microgrids: Application in Real Time Simulation Environment. In Proceedings of the 2019 IEEE PES Innovative Smart Grid Technologies Conference—Latin America (ISGT Latin America), Gramado, Brazil, 15–18 September 2019; pp.1-6.

13. Kar, S.; Hug, G. Distributed robust economic dispatch in power systems: A consensus + innovations approach. In Proceedings of the IEEE Power and Energy Society General Meeting, San Diego, CA, USA, 22–26 July 2012.

14. Olama, N.; Bastianello, P.; Mendes, R.C.; Camponogara, E. Relaxed hybrid consensus ADMM for distributed convex optimisation with coupling constraints. IET Control Theory Appl. 2019, 13, 2828–2837.

15. Wan, X.; Xu, Z.; Pinson, P.; Dong, Z.Y.; Wong, K.P. Probabilistic forecasting of wind power generation using extreme learning machine. IEEE Trans. Powers Syst. 2014, 29, 1033–1044.
16. Ospina, J.; Newaz, A.; Faruque, M.O. Forecasting of PV plant output using hybrid wavelet-based LSTM-DNN structure model. *IET Renew. Power Gener* 2019, 13, 1087-1095.

17. Chen, B.; Li, J. Combined probabilistic forecasting method for photovoltaic power using an improved Markov chain. *IET Gener. Transm. Distrib.* 2019, 13, 4364–4373.

18. Olfati-Saber, R.; Fax, J.A.; Murray, R.M. Consensus and Cooperation in Networked Multi-Agent Systems. *Proc. IEEE* 2007, 95, 215–233.

19. Zhang, J.; Cui, M.; He, Y. Robustness and Adaptability Analysis for Equivalent Model of Doubly Fed Induction Generator Wind Farm Using Measured Data. *Appl. Energy* 2020, 261, 114362.

20. Ghamkhari, M.; Mohsenian-Rad, H. A Convex Optimization Framework for Service Rate Allocation in Finite Communications Buffers. *IEEE Commun. Lett.* 2016, 20, 69–72.

21. Xiao, L.; Boyd, S. Fast linear iterations for distributed averaging. In *IEEE Conference on Decision and Control*; Elsevier: Amsterdam, The Netherlands, 2004.

22. Pagnier, L.; Jacquod, P. Optimal Placement of Inertia and Primary Control: A Matrix Perturbation Theory Approach. *IEEE Access* 2019, 7, 145889–145900.

© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).