frontiers | Research Topics

Planning, operation and control of modern power system with large-scale renewable energy generations

Edited by
Youbo Liu, Hao Xiao, Xia Lei, Yaser Qudaih and Xin Zhang

Published in
Frontiers in Energy Research

ISSN 1664-8714
ISBN 978-2-8325-3225-6
DOI 10.3389/978-2-8325-3225-6

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Planning, operation and control of modern power system with large-scale renewable energy generations

Topic editors

Youbo Liu — Sichuan University, China

Hao Xiao — Institute of Electrical Engineering, Chinese Academy of Sciences (CAS), China

Xia Lei — Xihua University, China

Yaser Qudaih — Higher Colleges of Technology, United Arab Emirates

Xin Zhang — Brunel University London, United Kingdom

Citation

Liu, Y., Xiao, H., Lei, X., Qudaih, Y., Zhang, X., eds. (2023). *Planning, operation and control of modern power system with large-scale renewable energy generations*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-3225-6

Table of contents

frontiers | Frontiers in Energy Research

TYPE Original Research
PUBLISHED 30 August 2022
DOI 10.3389/fenrg.2022.990509

OPEN ACCESS

EDITED BY
Youbo Liu,
Sichuan University, China

REVIEWED BY
Miao Yu,
Zhejiang University, China
Jun Hu,
Tsinghua University, China
Shenxi Zhang,
Shanghai Jiao Tong University, China

*CORRESPONDENCE
Qi Huang,
hwong@uestc.edu.cn

SPECIALTY SECTION
This article was submitted to Smart Grids,
a section of the journal
Frontiers in Energy Research

RECEIVED 10 July 2022
ACCEPTED 01 August 2022
PUBLISHED 30 August 2022

CITATION
Duan X, Huang Q, Chen Z, Li J and Ren X (2022), A rapid warning method for wildfire occurrences along the transmission corridor driven by power system monitoring data.
Front. Energy Res. 10:990509.
doi: 10.3389/fenrg.2022.990509

A rapid warning method for wildfire occurrences along the transmission corridor driven by power system monitoring data

Xiangxi Duan[1,2,3], Qi Huang[1,3]*, Zhe Chen[1,3], Jian Li[1,3] and Xibi Ren[1,3]

[1]University of Electronic Science and Technology of China, Chengdu, China, [2]State Grid Sichuan Electric Power Research Institute, Chengdu, China, [3]Power system wide area measurement and control of the Key Laboratory of Sichuan Province, Chengdu, China

This paper focuses on the problems of frequent wildfire occurrences along the transmission corridor and the lack of accurate and timely monitoring means for early warnings. Furthermore, this paper evaluates the rapid warning method for wildfire occurrences along the transmission corridor driven by power system monitoring data. First, we established the relationship between the historical data of wildfires along the transmission corridor and the operating state information of a power grid based on the Apriori association rule algorithm; the characteristic signals of the transmission line when wildfires occur were mined. Second, based on the characteristics of the time distribution of wildfire occurrences along the transmission corridor, a nonlinear regression model was created to further improve the prediction accuracy. Finally, by combining the characteristic signals and time distribution characteristics, we developed an early warning method. This method not only addresses the challenges faced by meteorological satellite remote sensors caused by the weather, the long transit time interval, and the high cost of adding sensors, but it also realizes the remote and rapid early warning of wildfires along the transmission corridor. Finally, a case study of practical data of a certain area in southwest China is used to verify the proposed method. The results show the high accuracy and timeliness of the proposed method.

KEYWORDS

wildfire occurrence along transmission corridor, nonlinear regression model, apriori association rule algorithm, monitor data, weather satellite

1 Introduction

Most renewable energy sources are concentrated in areas far from load centers and transported over long distances by transmission lines. These renewable energy transmission corridors are prone to wildfires. The safe and stable operation of those transmission lines is related to power grid security and people's livelihoods. However, as those transmission lines often cross vegetation-rich areas, the transmission corridor is prone to outbreaks of wildfires due to the combined effects of human activities and meteorological factors, which in turn

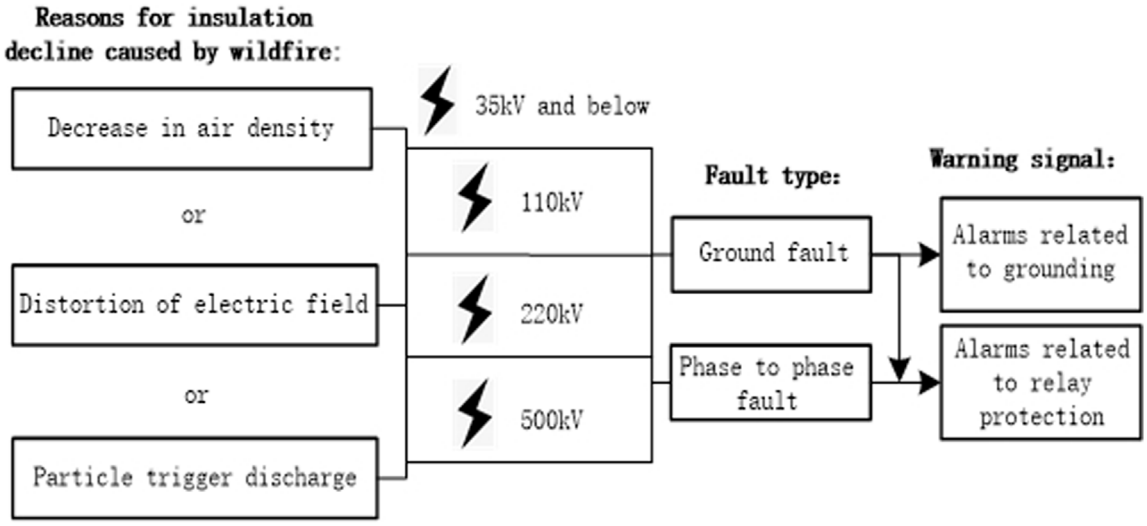


FIGURE 1
Mechanism of generating alarm signals.

threaten the safety of the power grid in the area and affect the normal power supply (Lu et al., 2017; Liu et al., 2018). Under wildfire conditions, the breakdown voltage of the transmission line will be significantly reduced. This dielectric strength reduction of the transmission line gap is the result of the combined action of the flame temperature, electrons and ions, and solid particles (Wu et al., 2011; Antonov, 2021). When the transmission line is grounded, short-circuited, or disconnected, it may also induce wildfires. For example, the direct cause of the March.30 forest fire in Liangshan, Sichuan Province in China is that wires reserved for connecting wires on both sides of the tower overlapped with the hoop of the pole cross-arm support frame under the action of the specific wind direction. This resulted in a permanent grounding discharge fault, causing the aluminum metal of the wire body to melt. The insulating material caught fire and burned, resulting in a significant loss of life and property (Accident investigation team, 2021). In another example, the wildfires in California were mainly caused by the failure of the power transmission equipment (Muhs et al., 2020).

The traditional transmission line wildfire warning mainly relies on manual inspections, which require considerable manpower and material resources. Currently, the research on early warnings for transmission line wildfires mainly focuses on the installation of smoke and fire detection devices at fixed locations, the installation of video surveillance devices for transmission lines, meteorological radar monitoring, and meteorological satellite remote sensing image recognition. Based on polar-orbiting satellite and geostationary satellite images, researchers (Liang et al., 2020) improved the threshold conditions, and adaptive dynamic threshold conditions were used for fire spot identification. In one study (Liu et al., 2020), the authors proposed a transmission line wildfire fault detection method based on millimeter-wave radar technology that combined infrared multi-spectral technology with millimeter-wave radar technology. In another study (Dian et al., 2019), a cellular automata model based on the principle of disaster geography was used to predict the spatiotemporal process of wildfires, and the proposed line interruption model was used to illustrate the mechanism of wildfire damage to transmission lines. A spatiotemporal context algorithm based on Himawari-8 geostationary satellite data was proposed in the literature (Chen et al., 2019). Other researchers (Muhs et al., 2021) proposed a probabilistic statistics-based hot spot identification algorithm for transmission lines. The automatic identification algorithm of wildfires in transmission lines was studied based on the weather radar echo database and the network monitoring of a new generation weather radar (Shu et al., 2020). Melnikov et al. (Melnikov et al., 2008) used S-band dual-polarization radar to analyze the radar polarization parameter characteristics of a forest fire echo in Oklahoma, United States. In another work (Beltramone et al., 2017), researchers used three risk factors of transmission line historical faults, real-time monitoring data, and meteorological forecast information, to develop an early warning evaluation model for transmission line fire prevention based on AHP.

Meteorological satellite remote sensing is the most widely used wildfire monitoring technology, which has the advantages of a wide monitoring range, short monitoring period, and high spatial resolution. However, there are disadvantages, such as a long time interval for satellite transit, blind spots in the scanning coverage, and easy interference in image recognition. Fires cannot be usually found at first, especially in the case of wildfires caused by transmission line faults. Installation of smoke and fire detection devices at fixed locations is costly,

FIGURE 2
The transmission lines were broken by a lightning strike.

FIGURE 3
Substandard tying process.

and maintenance costs are considerable. The use of weather radar to detect wildfires is still in its infancy, and it is susceptible to clutter interference, such as that from factories and wind power. Currently, the power grid has a relatively complete information monitoring system, and it is possible to carry out fire warning and analysis based on the information of the power grid monitoring system. In this paper, we propose a rapid early warning method for wildfires along the transmission corridors based on power grid monitoring information. Then the mechanism's relationship between wildfires along the transmission corridors and warning signals is investigated. The characteristic signals were mined using the Apriori association rule algorithm based on historical event records and related alarm signals of wildfires along the transmission corridor. Early warnings based on characteristic signals are susceptible to interference from similar warning signals caused by other power grid equipment failures. To further improve early warning accuracy, this paper analyzes the time distribution of regional wildfires along the transmission corridor. A nonlinear regression model of time distribution was established, and the optimization of the probability early warning value was carried out. Finally, by combining the characteristic signals and time distribution nonlinear regression model, the early warning method based on the characteristic signals and time distribution nonlinear regression model was created, and the remote and rapid early warning of wildfires along the transmission corridor was realized. This work eliminate the interference of meteorological conditions on meteorological satellite remote sensing without additional installment of sensors. It greatly contributes to manpower and financial reduction. Meanwhile the safe and stable delivery of large-scale renewable energy power is guaranteed.

In the first section of this paper, the mechanism's relationship between the transmission corridor fire and alarm signal is studied, and the Apriori association rule algorithm is introduced. In the second section, the nonlinear regression model of the time distribution of the wildfire along the transmission corridor is introduced to solve the problem of easy interference and low accuracy of the warnings based on the characteristic signals alone. In the third section, the probability warning value optimization of the time distribution nonlinear regression model is introduced. The early warning method based on the characteristic signal of wildfires and the time distribution nonlinear regression model is formed, and the remote and rapid early warning of wildfire is realized. In the fourth section, the feasibility of the method is verified by numerical examples. Based on the calculations for Example 1, and using the Liangshan forest fire (30 March) as an example, the high accuracy and timeliness are demonstrated by comparing the satellite image recognition method with the method described in this paper without considering the temporal distribution characteristics of mountain fires. In the fifth section, we discuss our conclusions, summarize the advantages of our method compared with the existing methods, and suggest the applications of the method.

2 Mining characteristic signals of wildfires along the transmission corridor based on the apriori association rule algorithm

2.1 Signal generation mechanism analysis

2.1.1 Power grid failure caused by wildfires

A wildfire is a kind of multiphase weak plasma containing solid and liquid particles with different particle sizes at high temperatures. When a wildfire spreads to the vicinity of the transmission line, the insulation strength of the transmission line

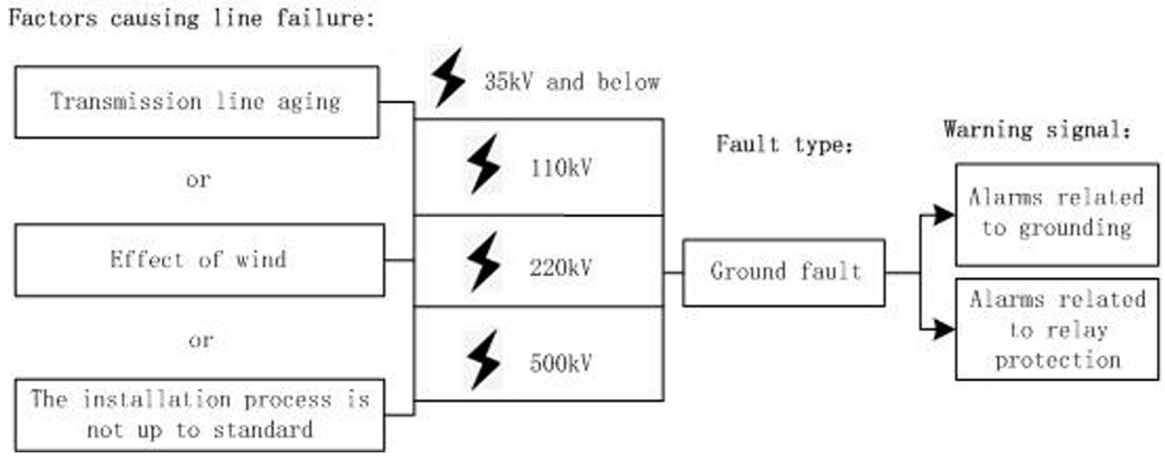


FIGURE 4
Mechanism of generating alarm signals.

decreases, and it is easy to cause the trip of the line due to the phase-to-phase fault and ground fault. The reasons for the decline of transmission line insulation strength caused by wildfire are as follows: 1) The high temperature of the flame reduces the air density, which leads to the decrease of the insulation level; 2) the electric field near the transmission line is distorted by the charge in the flame; and 3) particle contact leads to a discharge (Wu et al., 2011). Figure 1 shows the signal generation mechanism of a transmission line failure caused by a wildfire using the transmission lines of various voltage levels in southwest China as an example.

Most transmission lines of 35 kV and below are non-effective grounded systems. When a single-phase grounding fault occurs, the transmission lines will not trip directly and can still run for approximately 2 h. If a grounding fault occurs on transmission lines of 35 kV and below, alarms related to grounding will be triggered. When the ground fault and phase-to-phase fault occur on transmission lines of 110 kV and above, the relay protection machine action will be triggered and a trip will be caused.

2.1.2 Wildfires caused by transmission line faults

In the actual operation of transmission lines, faults such as grounding, short circuits, and transmission line disconnections may occur due to aging, substandard processes, wind action, lightning strikes, and other factors. As shown in Figures 2, 3, the transmission lines were broken by a lightning strike, and the installation process was not up to standard: these occurrences may both cause a wildfire.

When a wildfire is caused by a transmission line fault, the triggering mechanism of the alarm signals occurs as shown in Figure 4.

As shown in Figure 4, when the above factors cause a ground fault of the transmission line, the relevant alarm signals will be triggered. Similar to the alarm signals of a power grid failure caused by wildfire, alarms related to grounding and relay protection will be triggered. These alarm signals are sent to the power grid dispatch center or centralized control station through the remote devices in the substation. Therefore, it is possible to use the alarm signal of the control center for the rapid warning for wildfires.

TABLE 1 Related alarm signals.

No.	Alarm signal
1	Grounding alarm
2	The circuit breaker switch is misaligned
3	Protection trigger of relay protection device
4	Out-of-limit voltage
5	Action of harmonic elimination device
6	Action of traveling wave distance measurement
7	Alarm of excessive zero sequence current

2.2 Characteristic signal mining

The analysis of the signal triggering mechanism and the characteristics of the power grid equipment in southwest China revealed the related alarms that may be triggered. These are listed in Table 1. To further mine the data for the characteristic signals of wildfires in the transmission corridor during the actual operation of the power grid, it is necessary to match the recorded data of wildfires in the transmission corridor with the substation or power plant on both sides of the line, and associate the alarm data from 5 h before the start of wildfire to the end of wildfire in the transmission corridor.

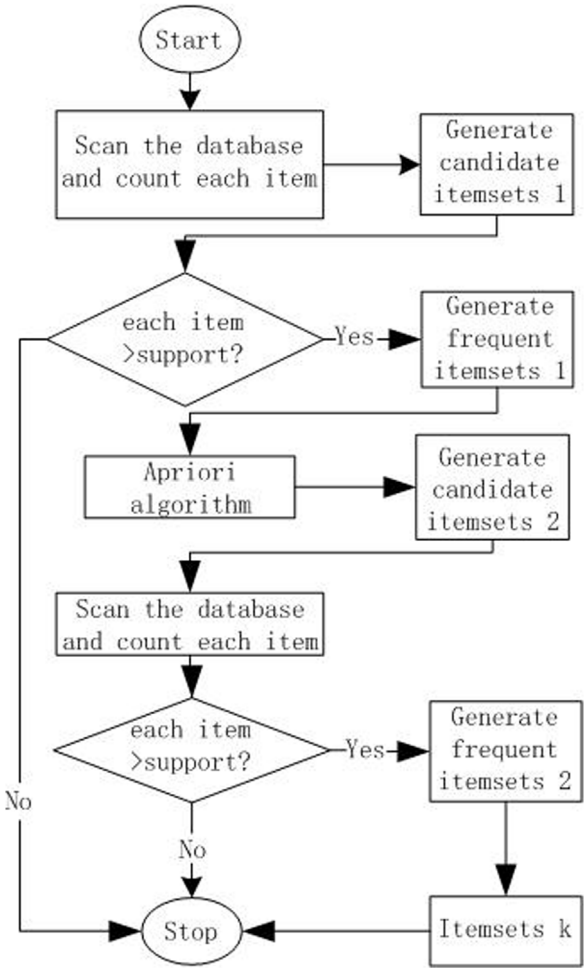


FIGURE 5
The computational process of the Apriori association rule algorithm.

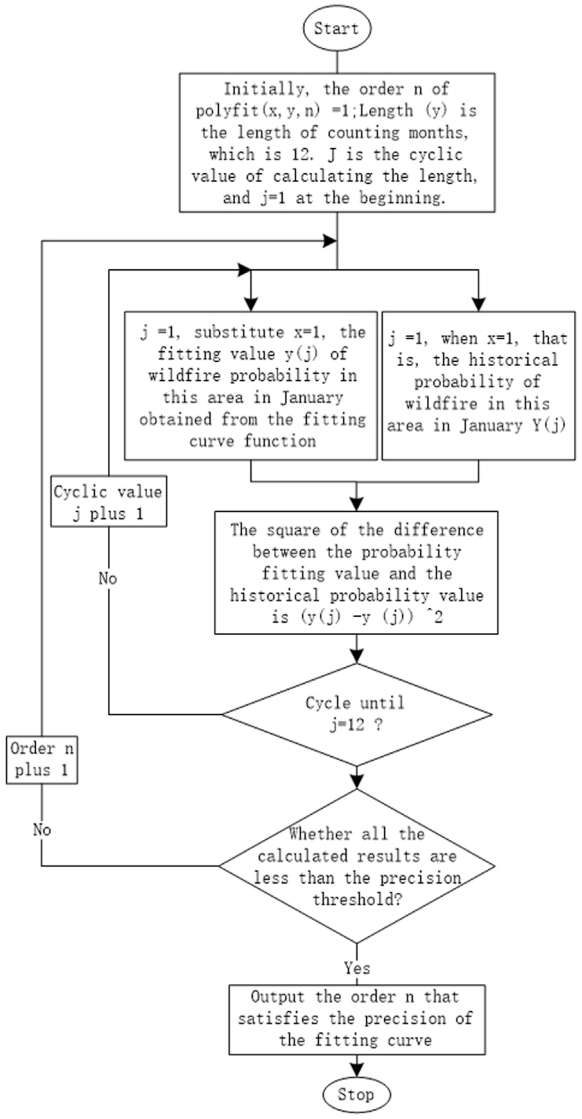


FIGURE 6
The calculation process of the order determination of fitting curve function.

The Apriori association rule algorithm is used to mine characteristic signals of transmission lines with different voltage levels, and the characteristic signals are mined from a large number of historical alarm data related to wildfire occurrence along the transmission line. The Apriori association rule algorithm is an iterative method of the layer-by-layer search. Its core idea is to generate candidate items and their support through connection and then generate frequent itemsets through pruning. The algorithm flow is shown in Figure 5.

The Apriori association rule algorithm is used to mine characteristic signals of transmission lines with different voltage levels. This is mined from a large number of historical alarm data related to wildfire occurrence along the transmission corridor. The Apriori association rule algorithm is a layer-by-layer iterative search method that generates candidate items and their support by concatenation and then generates frequent itemsets by pruning (Tian et al., 2020).

The alarm signal is expressed as follows:

$$Z_i\,(1 \le i \le 7,\ i \in N) \tag{1}$$

where i is the serial number of a characteristic signal.

Support = $P(Z_1\,Z_2)$ is the probability of event Z_1 and event Z_2 happening at the same time.

Confidence = $P(Z_2|Z_1) = P(Z_1\,Z_2)/P(Z_1)$ is the probability of event Z_2 occurring based on event Z_1.

Itemset k: If an event contains k elements, the event is called itemset k.

Frequent itemset k: The event met the minimum support threshold.

Strong association rules: The k-dimensional data itemset is a necessary condition for a frequent itemset, which is the set of all its k − 1 dimensional subitems, and is also a frequent itemset. These are rules that meet both minimum support thresholds and minimum confidence thresholds.

The computational process of the Apriori association rule algorithm is shown in Figure 5.

As shown in Figure 5, the one-dimensional itemset is scanned, and the one-dimensional frequent itemset is generated if the minimum support is satisfied. The one-dimensional itemset that satisfies the minimum support continues the minimum support calculation for the two-dimensional itemset until the k itemsets do not satisfy the minimum support. In the frequent itemset k − 1, the itemsets that do not meet the minimum support have been eliminated. If the remaining itemset meets the minimum confidence requirement through the confidence calculation, then the strong association rule in the itemset is obtained, and the itemset that meets the strong association rule is the characteristic signal.

3 Time distribution of regional wildfires along the transmission corridor

The analysis of the signal generation mechanism of a transmission corridor wildfire in the previous section indicated that the characteristic signal can realize the remote and rapid early warning of a transmission corridor wildfire. However because the grid equipment operation fault or abnormality will also trigger a similar signal, relying only on the characteristic signal will trigger the early warning frequently and the accuracy of early warning is low. However, due to seasonal climate changes, vegetation changes, human activities, and other factors, the transmission corridor wildfires have the characteristics of time distribution and an auxiliary feature signal for judgment, which can improve the accuracy of the early warning.

A nonlinear regression model was established by polynomial (Lamich et al., 2017). The frequency of wildfires distribution in each month is expressed as follows:

$$X_{mi}\,(1 \le i \le 12,\ i \in N), \tag{2}$$

where i indicates a month in a year.

The monthly probability distribution of wildfires in this region is as follows:

$$P_{mi} = \frac{X_{mi}}{\sum_{i=1}^{12} X_{mi}}. \tag{3}$$

Similarly, i indicates an hour in a day, and the hourly probability distribution of wildfires in this region is

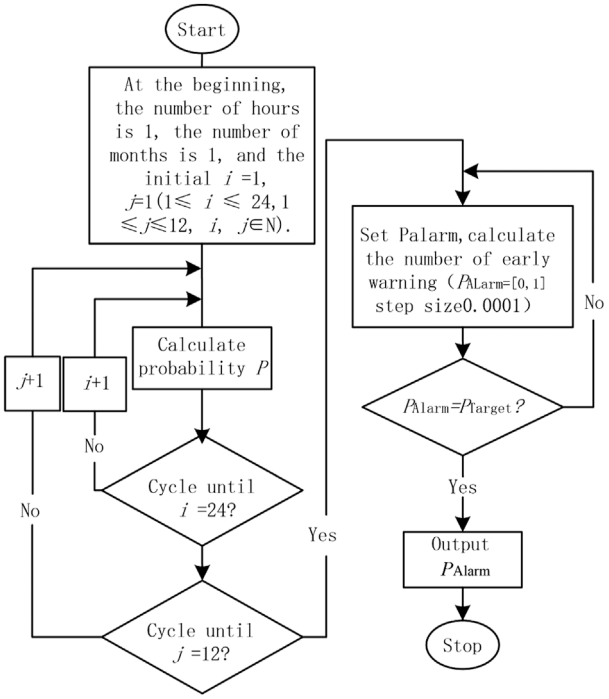


FIGURE 7
Optimization process of probability warning value.

$$P_{hi} = \frac{X_{hi}}{\sum_{i=1}^{24} X_{hi}}. \tag{4}$$

Equations 3, 4 were used to calculate the probability distribution, and then curve fitting was carried out to establish a nonlinear regression model, which was represented by polyfit (x,y,n). x represents the number of months or hours, y represents the probability value of the corresponding months or hours, and n represents the order of the fitting curve function. Since the occurrence probability of a wildfire near the transmission line is a low-probability event, to ensure that the accuracy of the model reaches 0.001, the order of fitting curve function is first determined by the successive increment method. The monthly distribution of wildfires in the transmission corridors can serve as an example; the calculation process of order determination is shown in Figure 6.

The expression of polyfit (x,y,n) for the fitting curve function is as follows:

$$f(x, y, n) = \text{polyfit}(x, y, n) \tag{5}$$

Initially, the order n of polyfit (x,y,n) is 1, Length (y) is the length of counting months, which is 12. J is the cyclic value of calculating the length, and j = 1 at the beginning.

The square of the difference between the probability fitting value and the historical probability value is (y(j) -y (j))^2. When

all the calculated results are less than the precision threshold, the order n that satisfies the precision of the fitting curve is output.

Combined with the calculation process, the calculation is cyclic from the time when the order n is 1, and finally the order n when the required precision is obtained.

Similarly, we take the monthly probability distribution of wildfires along the transmission corridor as an example; after determining the order n that meets the accuracy requirement, the sample point is (x_{mi}, y_{mi}), $mi = 1, 2, 3, \cdots, 12$, so the probability fitting value is $\hat{y}_{mi} = a_n x_{mi}^n + a_{n-1} x_{mi}^{n-1} + \cdots + a_0$. The loss function L is as follows:

$$L = \sum_{i=1}^{12} \left(Y_{mi} - \hat{y}_{mi}\right)^2 = \sum_{i=1}^{12} \left(Y_{mi} - a_n x_{mi}^n - a_{n-1} x_{mi}^{n-1} - \cdots - a_0\right)^2 \tag{6}$$

To minimize the loss function L, the partial derivative of each coefficient of the loss function is taken so that the value of the partial derivative is 0, which can be calculated using Eq. 7:

$$\begin{cases} \dfrac{\partial L}{\partial a_n} = 0 \\ \dfrac{\partial L}{\partial a_{n-1}} = 0 \\ \vdots \\ \dfrac{\partial L}{\partial a_0} = 0 \end{cases} \tag{7}$$

By solving these equations, we obtain the regression coefficient that minimizes the loss function L, that is, the regression coefficient of the fitting curve function polyfit (x,y,n), and finally have the nonlinear regression model with the required accuracy.

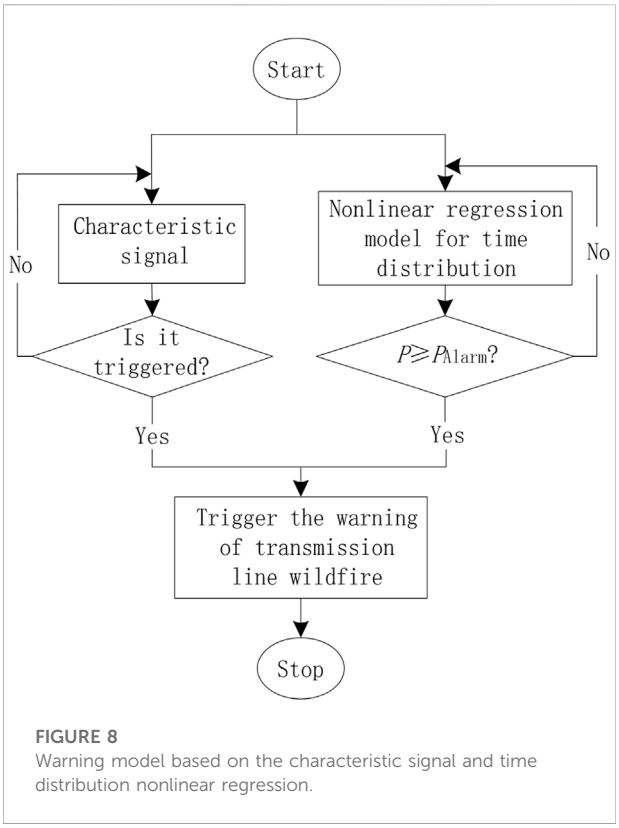


FIGURE 8
Warning model based on the characteristic signal and time distribution nonlinear regression.

4 An early warning method based on the characteristic signal and time distribution nonlinear regression model

Here, we elucidate the early warning method by combining the characteristic signals mined based on the Apriori association rule algorithm and the established nonlinear regression model of the time distribution of wildfires in the transmission corridor in the previous section.

4.1 Optimization of probability warning value

The probability warning value of the nonlinear regression model is set as P_{Alarm}. When the probability of the nonlinear regression model reaches the warning value, the warning about the time distribution is output. To optimize the probability early warning value, the increasing substitution method is used to optimize the early warning value. The final accuracy rate is set as the target value, and the purpose of optimization is to maximize it, as defined in the following formula:

$$P_{Target} = \left(\min\left(\text{number of warnings}\right) \middle| \max\left(\text{The number of correct warnings}\right)\right). \tag{8}$$

To ensure that no correct early warning is missed, the target value is the probability of the minimum number of early warning corresponding to the maximum number of correct early warning. The calculation process is shown in Figure 7, where 1 hour is taken as a step.

As shown in Figure 7, the probability warning value of each time point of the year P is calculated in hourly steps, and then the warning value P_{Alarm} is set. The occurrence of wildfires in transmission corridors is a small probability event. To ensure the accuracy of the model reaches 0.001, it is calculated in steps of 0.0001, incrementing from 0 to 1, and finally making P_{Target} reach the maximum, i.e., the output of the corresponding optimal warning value.

4.2 The early warning process

The probability of the early warning process of wildfire along transmission corridor based on the characteristic signals and

TABLE 2 Manual record of wildfires.

No.	Transmission line name	Start time	End time	Affects the power grid
1	220 kV Xilin	February 3 15:10	February 3 22:20	Yes
2	35 KV Lula	January 8 14:52	January 9 07:04	Yes
3	10 KV Kaiyuan	February 8 15:40	February 9 08:00	Yes

time distribution nonlinear regression model is shown in Figure 8.

Figure 8 and Eq. 9 show that, among the characteristic signals of transmission lines with different voltage levels, the output Boolean quantity is 1 when the characteristic signals are detected. Otherwise, the output Boolean quantity is 0:

$$\begin{aligned} \mathrm{T}_i &\rightarrow bool\colon 1 \\ \mathrm{T}_i &\rightarrow bool\colon 0 \end{aligned} \tag{9}$$

As shown in Eq. 10, the warning value of the nonlinear regression of the time distribution is set as P_{Alarm}. When the warning value is reached, the Boolean value of the output warning value is 1; otherwise, the Boolean value of the output is 0:

$$\begin{aligned} P &\geq \mathrm{P}_{\mathrm{Alarm}} \rightarrow bool\colon 1 \\ P &< \mathrm{P}_{\mathrm{Alarm}} \rightarrow bool\colon 0 \end{aligned} \tag{10}$$

When both the characteristic signal and the nonlinear regression warning of the time distribution meet the conditions, the warning will be output to realize the estimation of the wildfire occurrences along the power grid driven by the power system monitoring data.

5 Experimental verification

5.1 Examples of typical areas

5.1.1 Characteristic signal mining

We take the states with the most serious wildfire hazards in southwest China as an example: 277 wildfire incidents occurred in 2020, including 57 directly affecting the power grid. The records of the wildfire incidents along the transmission corridor are shown in Table 2.

The events include the fire start time, fire end time, and the name of the transmission line. The start time is the time when the fire is discovered and reported, and the end time is the time when the fire is confirmed to be over.

There were 31 events involving transmission lines of 110 kV and above. The recorded data of the wildfires in the transmission corridor with the substation or power plant on both sides of the line are matched, and the alarm data from 5 h before the transmission corridor fire starts to the end of the fire are associated to form a dataset. Based on the alarm signals in Table 1 and the process in Figure 5, the dataset was scanned to obtain the support of each candidate, and the minimum support threshold was set to 0.1. By comparing the candidate support count with the minimum support, Z_4 and Z_7 do not satisfy the condition of one-dimensional frequent set. $\{Z_2, Z_5\}$ and $\{Z_3, Z_5\}$ do not satisfy the two-dimensional frequent itemset. All items do not satisfy the three-dimensional frequent itemset. Therefore, the maximum frequent itemset is two dimensions. Then the confidence of each two-dimensional frequent itemset is calculated, and the calculation results are shown in Table 3.

Similarly, there are 26 events in 35 kV and below lines, and the dataset is scanned to obtain the support of each candidate item. Since the data are relatively dense, the minimum support threshold is set to 0.9. Compared to the candidate support count with the minimum support, Z_2, Z_3, Z_5, Z_6, and Z_7 do not meet the one-dimensional frequent set. $\{Z_1, Z_2\}$, $\{Z_1, Z_5\}$, and $\{Z_1, Z_7\}$ do not meet the two-dimensional frequent itemset. All are not satisfied with the three-dimensional frequent itemset. The maximum frequent itemset is two-dimensional. Then the confidence of each two-dimensional frequent itemset is calculated. The calculation results are shown in Table 4.

Using the calculation results for 35 kV and below lines, the grounding alarm signal Z_1 and voltage overrun signal Z_4 were selected as the characteristic signals.

5.1.2 Nonlinear regression model

The distribution probability of the hours and months was obtained through the processing of the historical record data, as shown in Tables 5, 6.

Using the calculation process of fitting the order of the curve function in Figure 6, the order is determined as 5 by substituting the data in Tables 5, 6, that is, $n = 5$.

Using the monthly distribution probability of wildfires in Table 5, the nonlinear regression model of the monthly distribution was established, as shown in Eq. 11:

$$\begin{aligned} f\left(x_m, y_m, 5\right) &= polyfit\left(x_m, y_m, 5\right) \\ &= a_5 x_m^5 + a_4 x_m^4 + a_3 x_m^3 + a_2 x_m^2 + a_1 x_m + a_0 \end{aligned} \tag{11}$$

By solving Eqs 6, 7, the regression coefficient of the nonlinear regression model of the monthly distribution was obtained, as shown in Eq. 12:

TABLE 3 Transaction dataset of 110 kV and above transmission lines.

Signal	P(Z_i)	Rule	P(Z_i Z_j)	P(Z_i|Z_j)
Z_1	54%	$Z_1 \rightarrow Z_2$	14%	26%
Z_2	46%	$Z_2 \rightarrow Z_1$	14%	31%
Z_3	46%	$Z_3 \rightarrow Z_1$	14%	31%
Z_4	3%	$Z_1 \rightarrow Z_3$	14%	26%
Z_5	28%	$Z_2 \rightarrow Z_3$	100%	100%
Z_6	0%	$Z_3 \rightarrow Z_2$	100%	100%
Z_7	3%	$Z_5 \rightarrow Z_1$	14%	50%
		$Z_1 \rightarrow Z_5$	14%	26%

TABLE 4 Transaction dataset of 35 kV and below transmission lines.

Signal	P(Z_i)	Rule	P(Z_i Z_j)	P(Z_i|Z_j)
Z_1	100%	$Z_1 \rightarrow Z_4$	96%	96%
Z_2	88%	$Z_4 \rightarrow Z_1$	96%	100%
Z_3	88%			
Z_4	96%			
Z_5	77%			
Z_6	0%			
Z_7	85%			

$$\begin{cases} a_5 = 0.0016 \\ a_4 = -0.0728 \\ a_3 = 1.2164 \\ a_2 = -8.8449 \\ a_1 = 22.8692 \\ a_0 = 4.4424 \end{cases} \tag{12}$$

Using the hourly distribution probability of line mountain fire in Table 6, the nonlinear regression model of the hourly distribution was established, as shown in Eq. 13:

$$f(x_h, y_h, 5) = polyfit(x_h, y_h, 5) = a_5 x_h^5 + a_4 x_h^4 + a_3 x_h^3 + a_2 x_h^2 + a_1 x_h + a_0 \tag{13}$$

By solving Eqs 6, 7, the regression coefficient of the nonlinear regression model with the hourly distribution was obtained, as shown in Eq. 14:

$$\begin{cases} a_5 = 0.0001 \\ a_4 = -0.0070 \\ a_3 = 0.1376 \\ a_2 = -1.0139 \\ a_1 = 2.7384 \\ a_0 = -1.5313 \end{cases} \tag{14}$$

Finally, the nonlinear regression model of time distribution is shown in Eq. 15:

$$P = polyfit(x_m, y_m, 5) \bullet polyfit(x_h, y_h, 5) \tag{15}$$

TABLE 5 Monthly distribution probability of wildfires.

Month	1	2	3	4	5	6
P_{mi}	0.21	0.18	0.27	0.13	0.10	0.023
Month	**7**	**8**	**9**	**10**	**11**	**12**
P_{mi}	0.002	0.014	0.009	0.008	0.02	0.047

TABLE 6 Hourly distribution probability of wildfires.

Hour	1	2	3	4	5	6
P_{hi}	0.006	0.008	0.002	0.002	0	0.003
Hour	**7**	**8**	**9**	**10**	**11**	**12**
P_{hi}	0.008	0.017	0.014	0.045	0.055	0.048
Hour	**13**	**14**	**15**	**16**	**17**	**18**
P_{hi}	0.067	0.12	0.11	0.10	0.13	0.08
Hour	**19**	**20**	**21**	**22**	**23**	**24**
P_{hi}	0.06	0.04	0.05	0.02	0.008	0.003

The time probability of wildfires along the transmission corridor of this region was obtained by inputting the monthly hours of the whole year into the nonlinear regression model of time distribution with an hour as the step, as shown in Figure 9.

It can be seen from the figure that the time distribution probability of wildfires in this area is very obvious, and it is impossible for wildfires to occur during certain times. In February at 16:00, the time probability of wildfires is the highest (0.025). From the 12-month distribution of the whole year, the frequency of wildfires is lower in the summer and autumn and higher in the winter and spring. During the Chinese New Year and Tomb-Sweeping Day, it reaches the peak (February). In the months with a high frequency of line wildfires, the probability of wildfires in 24 h presents a normal distribution.

5.1.3 Optimization of the probability warning value

After the time probability of wildfires is obtained, the probability warning value P_{Alarm} of the time distribution is optimized. Multiple mountain fires occurring at the same hour are assumed as one time for calculation, so that the probability warning value P_{Target} can be reached. The calculation results are shown in Figure 10.

As it can be seen in Figure 10 and according to the definition of Eq. 8, through the optimization of probability early warning value P_{Alarm}, $P_{\text{Target}} = 0.005$. At this time, the number of correct early warnings is 57, the total number of early warnings is 61, and the probability of an accurate early warning is 93%. When the probability warning value continues to increase, the probability of an accurate warning can be further improved, but the number of correct warnings is sacrificed.

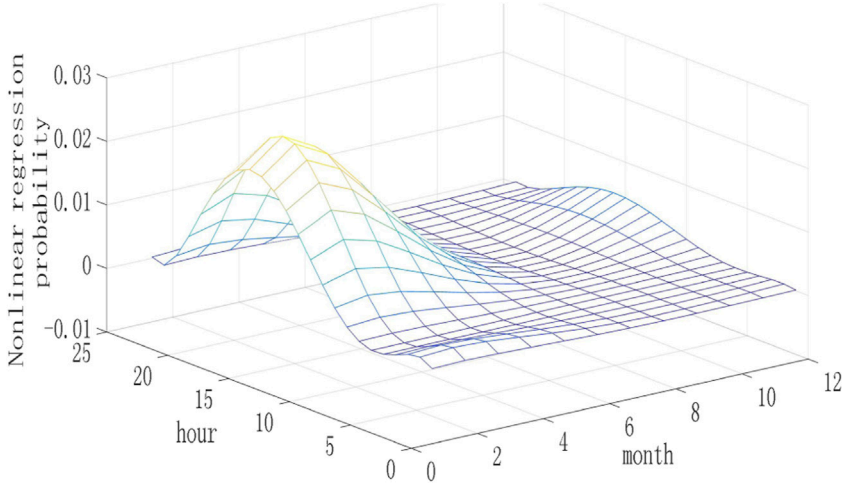


FIGURE 9
Time probability of wildfires.

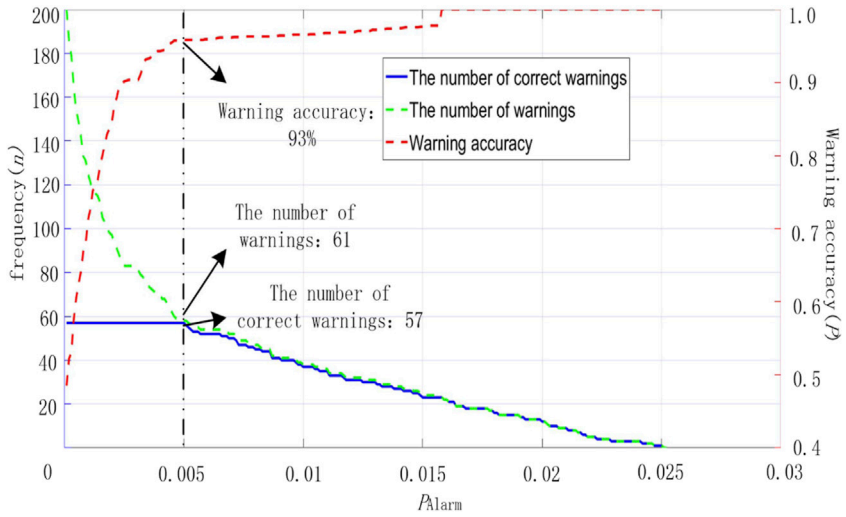


FIGURE 10
Optimization results of probability warning values.

5.1.4 Warnings of wildfires along transmission corridors

The monitoring and alert information of the power grid in January 2021 was extracted from the database of the main power grid regulation and control station of the city and state and calculated by substituting the alert process in Figure 8. In January 2021, there were three transmission corridor wildfire events affecting the grid operation in the region. To verify the accuracy of the time distribution regression model for the transmission corridor fires, the monitoring and warning information of the power grid was also substituted into the characteristic signal model without considering the time distribution regression warning. The prediction results using satellite image recognition technology for the same period were collected and compared to obtain the results shown in Figure 11.

Figure 11 shows that for the example of the transmission corridor wildfires affecting the power grid that occurred in the region in January 2021, the early warning model based on the

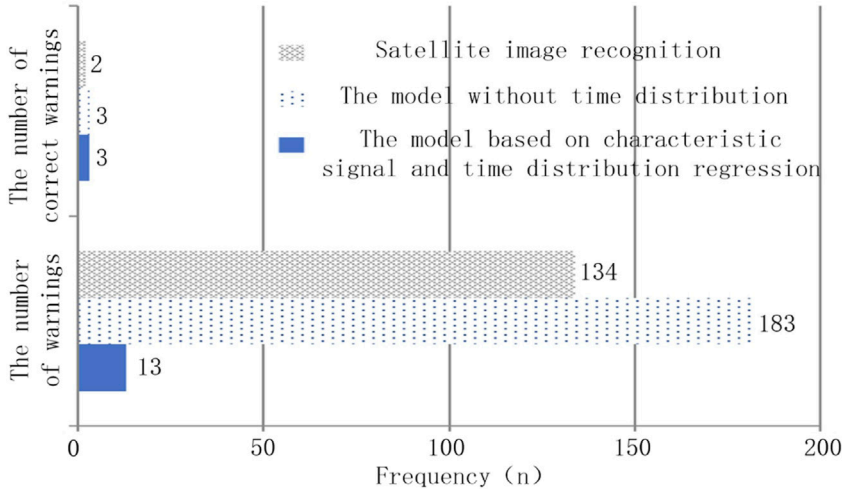


FIGURE 11
Comparison of prediction results.

TABLE 7 Warning times.

No	Model based on characteristic signal and time distribution regression	Satellite image recognition	Manual recording
1	January 3 15:05	January 3 15:25	January 3 15:05
2	January 7 13:40	January 7 13:50	January 7 13:57
3	January 3 13:15	None	January 3 13:30

TABLE 8 Records of wildfires affecting power grid in Liangshan district in March 2020.

No	Voltage level (kV)	Start time	End time	Impact and measures
1	220	March 30 17:12	March 30 21:09	Relay protection starts trip
2	10	March 30 17:00	March 30 17:41	Ground fault and no trip

TABLE 9 Warning times.

No.	Voltage Level (kV)	Warning condition	Early warning model in this paper	Lead time
1	220	The characteristic signal is triggered, and the time probability reaches the warning value	March 30 17:12	0 min
2	10	The characteristic signal is triggered, and the time probability reaches the warning value	March 30 15:33	87 min

characteristic signal and time distribution regression warned three times, while the satellite image recognition only warned correctly two times. There is a significant difference in the accuracy of these methods. With the proposed model, it warns 13 times with an accuracy of 23%. For the characteristic signal model without considering the time distribution regression warning, it warned 183 times with an accuracy rate of only 1.6% due to the numerous disturbing signals in the actual operation of the power grid. The satellite image recognition warned 134 times with an accuracy rate of 1.5%. Further analysis of the warning time is shown in Table 7.

The data in Table 7 indicate the satellite image identification technique is limited by the long transit time interval, the existence of blind areas in the scanning coverage, and the influence of weather. Two warning times are later than the method described in this paper. Because of the monitoring background alarm information update time of 1 minute, the early warning model based on the feature signal and time distribution regression was 12 and 15 min earlier than the manually determined start time in events 2 and 3, respectively.

5.2 Examples of "30 March" forest fire in liangshan, sichuan province

For the "30 March" forest fire in Liangshan, Sichuan in 2020, the fire was caused by a transmission line fault. Since the Liangshan region is adjacent to the above modeled area and has similar geographical and climatic conditions, the above model was used to extract the monitoring alarm information from the monitoring background for the Liangshan region in March 2020. In March 2020, there were two incidents of wildfires affecting the power grid in the Liangshan region, as shown in Table 8, among which the "3–30" forest fire is incident 2.

Our calculations show that the number of early warnings was 11, the correct number was two times, and the accuracy was 18%. The early warning results were analyzed (in March 2020, wildfire warnings based on satellite image recognition technology were not issued in this province), as shown in Table 9.

Table 9 indicates that the warning time of the proposed method is synchronized with the start time of the manual recording for 220 kV lines. For the "30 March" forest fire incident, since a single-phase ground fault does not directly trigger a trip in the transmission lines of 35 kV or below, the method in this paper has more obvious warning timeliness and can warn of a wildfire 87 min in advance.

6 Conclusion

This paper puts forward a rapid warning method for wildfires based on state grid monitoring information. First, the characteristic signals of wildfires were determined based on the mechanism analysis of the power grid alarm signals of a transmission line wildfire. To improve the warning accuracy, the time distribution characteristics of wildfires were analyzed, and the nonlinear regression model of the time distribution was established. We combined them to develop an early warning method based on the characteristic signals and the time distribution nonlinear regression model.

The effectiveness of the method was verified through typical regional arithmetic examples. Compared with the method based on satellite image recognition technology, the method can overcome the disadvantages of a long transit time interval, areas without adequate scanning coverage, weather, and other types of interference. The accuracy rate reaches 23%, which is much higher than that of satellite image recognition, and can provide earlier warnings compared with the start time of manual recordings. The advantage is obvious for the wildfires caused by transmission line faults. For the example case of the "30 March" wildfire in Liangshan, the method could have provided a warning 87 min earlier, further proving its effectiveness and rapidity, especially for low-voltage lines.

The method proposed in this paper does not require additional sensors and can avoid the interference of satellite image identification due to meteorological conditions. By combining with existing monitoring methods, such as satellite image identification technology and sensor detection, it can provide a complementary advantage for wildfire warning systems. Thus, it can be used as a powerful supplement to the existing wildfire prevention and control measures to guarantee the security, stability, and economic operation of large power grids. Furthermore, this can be effective in protecting people's lives and property from devastating wildfires.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

This work was supported by Power system wide area measurement and control of the Key Laboratory of Sichuan Province. XD found this problem in his daily work and proposed a solution, he also collected field data. QH, ZC, and JL gave instructions on how to solve the problem, XR conduct experiment.

Funding

This work was partially supported by Science and Technology Innovation Talent Program of Sichuan Province (grant number:22CXRC100).

Acknowledgments

This work was supported by the Power system wide area measurement and control of the Key Laboratory of Sichuan

Province. In addition, XD wants to thank QH, ZC, and JL for them affirmation and all those who love him.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Antonov, S. (2021). "Impact of the wildfires on the power supply networks," in 2021 IV International Conference on High Technology for Sustainable Development (HiTech), Sofia, Bulgaria, 07-08 October 2021 (IEEE). doi:10.1109/HiTech53072.2021.9614216

Beltramone, G., Alaniz, E., Ferral, A. E., Aleksinkó, A., Arijón, D. R., Bernasconi, I., et al. (2017). "Risk mapping of urban areas prone to flash floods in mountain basins using the analytic hierarchy process and geographical information systems," in 2017 XVII Workshop on Information Processing and Control (RPIC), Mar del Plata, Argentina, 20-22 September 2017 (IEEE), 1–6. doi:10.23919/RPIC.2017.8214379

Chen, W., Zhou, Y., Chen, J., Zhou, E., Zhou, W., and Sui, S. (2019). "Wildfire monitoring technology of transmission lines based on Himawari8 geostationary meteorological satellite," in Proc. IEEE 3rd Conf. Energy Internet Energy Syst. Integr. (EI), Changsha, China, 08-10 November 2019, 2408–2412. doi:10.13336/j.1003-6520.hve.20190498

Dian, S. Y., Cheng, P., Ye, Q., Wu, J., Luo, R., Wang, C., et al. (2019). Integrating wildfires propagation prediction into early warning of electrical transmission line outages. *IEEE Access* 7, 27586–27603. doi:10.1109/access.2019.2894141

Lamich, M., Balcells, J., Corbalán, M., and Griful, E. (2017). Nonlinear loads model for harmonics flow prediction, using multivariate regression. *IEEE Trans. Ind. Electron.* 64 (6), 4820–4827. doi:10.1109/tie.2017.2674596

Li, H., and Long, Z. (2021). Risk analysis of forest fire in Liangshan Prefecture, Sichuan Province based on Logistic model. *Journal of Safety and Environment* 21 (2), 498–505.

Liang, Y., Zhou, L., Chen, J., Huang, Y., Wei, R., and Zhou, E. (2020). Monitoring and risk assessment of wildfires in the corridors of high-voltage transmission lines. *IEEE Access* 8, 170057–170069. doi:10.1109/access.2020.3023024

Liu, H., Wang, S., Jiang, M., and Li, J. (2020). "Analysis of millimeter-wave radar for monitoring mountain fire in power grid," in 2020 IEEE 5th Information Technology and Mechatronics Engineering Conference (ITOEC), Chongqing, China, 12-14 June 2020 (IEEE), 805–809. doi:10.1109/ITOEC49072.2020.9141870

Liu, Y., Lu, J. Z., Luo, J., Zhang, G. Y., and He, L. F. (2018). Synchronous satellite wide area monitoring for overhead transmission line wildfire and tower location. *Power Syst. Technol.* 42 (4), 1322–1327. doi:10.13335/j.1000-3673.pst.2017.1812

Lu, J., Zhou, T., Wu, C., Li, B., Liu, Y., and Zhu, Y. (2017). Prediction and early warning technology of wildfire nearby overhead transmission lines. *High. Volt. Eng.* 43 (8), 2524–2532. doi:10.13336/j.1003-6520.hve.20161227041

Melnikov, V. M., Zrnic, D. S., Rabin, R. M., and Zhang, P. (2008). Radar polarimetric signatures of fire plumes in Oklahoma. *Geophys. Res. Lett.* 35 (14), 148155–L15443. doi:10.1029/2008gl034311

Muhs, J. W., Parvania, M., Nguyen, H. T., and Palmer, J. A. (2021). Characterizing probability of wildfire ignition caused by power distribution lines. *IEEE Trans. Power Deliv.* 36 (6), 3681–3688. doi:10.1109/tpwrd.2020.3047101

Muhs, J. W., Parvania, M., and Shahidehpour, M. (2020). Wildfire risk mitigation: A paradigm shift in power systems planning and operation. *IEEE Open J. Power Energy* 7, 366–375. doi:10.1109/oajpe.2020.3030023

Shu, S., Zhang, S., Xu, J., Xie, W., and Fang, C. (2020). Study on automatic identification algorithm of wild fire near transmission lines based on CINRAD-net monitoring. *Proc. CSEE* 40 (13), 4200–4209. doi:10.13334/j.0258-8013.pcsee.190516

Tian, M., Zhang, L., Guo, P., Zhang, H., Chen, Q., Li, Y., et al. (2020). Data dependence analysis for defects data of relay protection devices based on apriori algorithm. *IEEE Access* 8, 120647–120653. doi:10.1109/access.2020.3006345

Wu, T., Ruan, J., Hu, Y., Liu, B., and Chen, C. (2011). Study on forest fire induced breakdown of 500 kV transmission line in terms of characteristics and mechanism. *Proc. CSEE* 31 (34), 163–170. doi:10.13334/j.0258-8013.pcsee.2011.34.0

frontiers | Frontiers in Energy Research

TYPE Original Research
PUBLISHED 30 September 2022
DOI 10.3389/fenrg.2022.1023822

OPEN ACCESS

EDITED BY
Youbo Liu,
Sichuan University, China

REVIEWED BY
Xiandong Xu,
Tianjin University, China
Zao Tang,
Hangzhou Dianzi University, China
Yang Liu,
Sichuan University, China

*CORRESPONDENCE
Wei Yang,
yangwei_scu@126.com

SPECIALTY SECTION
This article was submitted to Smart Grids,
a section of the journal
Frontiers in Energy Research

RECEIVED 20 August 2022
ACCEPTED 02 September 2022
PUBLISHED 30 September 2022

CITATION
Yang W, Huang B, Zhang A, Li Q, Li J and Xue X (2022), Condition prediction of submarine cable based on CNN-BiGRU integrating attention mechanism. *Front. Energy Res.* 10:1023822. doi: 10.3389/fenrg.2022.1023822

Condition prediction of submarine cable based on CNN-BiGRU integrating attention mechanism

Wei Yang[1]*, Bo Huang[1], Anan Zhang[1], Qian Li[1], Jiaxing Li[1] and Xinghui Xue[2]

[1]School of Electrical and Information Engineering, Southwest Petroleum University, Chengdu, Sichuan, China, [2]Water and Power Supply Company, PetroChina Dushanzi Petrochemical Company, Karamay, Xinjiang, China

As the lifeline of energy supply for various offshore projects, accurately evaluating and predicting the operation status of submarine cables are the foundation for the reliable operation of energy systems. Based on fully mining the dynamic and static characteristics of submarine cable operation and maintenance data, this paper proposes a submarine cable operation status prediction method based on a convolutional neural network—bidirectional gated recurrent unit (CNN-BiGRU) integrating attention mechanism. Firstly, the evaluation index system of the submarine cable operation status is established by considering three key influencing factors including online monitoring, routine inspection, and static test. Then, the operation condition evaluation model for submarine cable is constructed based on the cooperative game theory and the multi-level variable weight evaluation. Finally, the CNN-BiGRU combined neural network model integrating the attention mechanism is established, and the historical operation data and condition quantification results (health value) are used as input characteristic parameters to predict the evolution trend of the operation status of the submarine cable. The case study shows that the proposed method can effectively predict the operation status of submarine cables, and the root mean square error of the prediction is as low as 1.36%, which demonstrates the superior performance compared with the back propagation (BP) neural network, CNN, long short-term memory (LSTM), CNN-LSTM, and other algorithms.

KEYWORDS

submarine cable, condition prediction, convolution neural network, bidirectional gate recurrent unit, attention mechanism, condition evaluation

1 Introduction

In order to achieve carbon neutrality and solve the global energy crisis, the utilization of marine resources is in a stage of rapid increase, which has driven the rapid development of various offshore wind farms, offshore oil and gas platforms, and other marine projects. Submarine cables are not only the main artery between offshore wind farms and onshore

power systems but also the lifeline of power and communication between offshore platforms and production systems, as well as the strategic foundation for building new offshore power systems based on distributed renewable energy (Bastien et al., 2018; Liu X.et al., 2019; Purvins et al., 2018). Until 2021, there are more than 450 submarine cables in use around the world, with a total length of over 1.2 km × 10^6 km (Xie et al., 2022). The reliable operation of submarine cables is essential for the normal operation of offshore projects. Due to the particularity of the environment where the submarine cable is located, it is inconvenient to carry out a routine manual inspection or UAV inspection. In recent years, thanks to the rapid development and application of optical fiber sensing technology, the relevant operating parameters of the submarine cable can be obtained in real-time (Fang et al., 2020; Antonio et al., 2021; Miguel et al., 2021; Cao et al., 2022), such as temperature, disturbance, strain, and other information of the submarine cable. These monitoring data can be used for operating condition evaluation and condition prediction (Lv et al., 2014; Xu et al., 2019), so as to understand the health status of the submarine cable, and the preventive maintenance can be carried out in a targeted manner, thereby improving the operation reliability of the offshore integrated energy system.

Condition monitoring and evaluation play an indispensable role in the electrical equipment area, many scholars have carried out related research on this issue. In terms of submarine cables, Wei et al. (2015) proposed an online monitoring method for 500 kV submarine oil-filled cable sheath insulation, and Liu Z. et al. (2019) established an evaluation model for the deterioration state of submarine cables considering factors such as morphology, chemical structure, mechanical properties, thermal properties, and dielectric properties, and the transformation of the sheath insulation current leakage ratio and the grounding loop circulating current inductance were used as the basis for sheath insulation judgment, and insulation condition of the submarine cable jacket was monitored by Zhou et al. (2018) and Nie et al. (2020). According to the requirements of field survey and data management in submarine cable inspection by remote-operated vehicle (Lei, 2021), analyzed the characteristics of submarine cable tracker and underwater video inspection data, designed reasonable system architecture, implemented real-time monitoring and data management of submarine cable inspection based on underwater navigation and positioning. Chen et al. (2021) proposed enhanced coherent optical time-domain reflectometry (E-COTDR) for monitoring submarine cable loss and vibration. Chen et al. (2020) established a temperature monitoring system based on Brillouin Optical Time Domain Analysis (BOTDA) technology, and the accuracy of the submarine cable temperature monitoring system based on optical fiber technology was verified. Jianfeng, (2019) proposed a submarine cable fault data monitoring method based on big data fusion and feature extraction. However, few studies have focused on the comprehensive evaluation and prediction of the submarine cable operating condition, which is essential for preventive maintenance.

With the rapid development of artificial intelligence, various intelligent algorithm models have been wildly used in various fields, such as load prediction and condition prediction. Pu et al. (2020) and Liu et al. (2019a) expounded on the application and development of big data in electrical equipment condition monitoring and analyzed the existing deficiencies and prospects. Xu and Jiang, (2022) developed a power load prediction method based on BiGRU-Attention-SENet, which combines the characteristics of different neural networks to improve accuracy. Tang and Yuan, (2022), Dai et al. (2018), and Song et al. (2018) used neural network models, such as CNN-BiGRU and LSTM, to predict the remaining useful life of lithium batteries and the condition of transformers. However, due to the special application scenarios and equipment properties of submarine cables, there are few studies dedicated to the evaluation and prediction of their operating condition.

Therefore, a submarine cable operation condition evaluation and prediction method based on CNN-BiGRU integrating attention mechanism is proposed in this paper. Firstly, a submarine cable status evaluation index system is established based on relative industry standards. Secondly, based on the cooperative game theory and the idea of multi-level variable weight, the comprehensive evaluation model of health status for submarine cables is developed. Thirdly, a submarine cable operation status prediction method based on CNN-BiGRU integrating attention mechanism is proposed, and the historical operation data and condition evaluation results are used as input characteristic parameters to predict the operation condition of submarine cables. Finally, the validity and feasibility of the proposed methodology are demonstrated through a submarine cable in Bohai, China. The contributions of this paper are as follows:

(1) Based on the industry guidelines and comprehensively considering the structural characteristics, operation history, maintenance and fault data of submarine cables, and the online monitoring data, a submarine cable status evaluation index system is established, which can accurately and comprehensively evaluate the operation condition of submarine cables.
(2) A multi-level variable weight submarine cable condition evaluation method is developed based on cooperative game theory, which can effectively change the evaluation weight adaptively according to different operating periods of the submarine cable, in order to obtain an appropriate comprehensive submarine cable operating condition.
(3) A CNN-BiGRU model integrating attention mechanism is proposed to predict the condition of submarine cables. CNN

is used to exploit the relationship between the various feature parameters, while BiGRU is used to capture long-term decencies from time series data. The attention mechanism is introduced to give different weights to the implied conditions of BiGRU through the mapping weighting and learning parameter matrix, so as to strengthen the influence of important information and further improve the forecasting efficiency.

(4) To verify the effectiveness of the proposed model, it is compared with different models, such as BP, LSTM, GRU, CNN-LSTM, and CNN-BiGRU. The maximum percentage of prediction error $\sigma_{E\text{-}max}$, the root mean square error R_{MSE}, and the mean absolute percentage error M_{APE} of different models are calculated to evaluate the performance. The results demonstrate that the forecasting accuracy of the proposed model is greatly improved.

The remainder of this paper is organized as follows: The operation condition evaluation model of submarine cable is described in Section 2, and Section 3 presents the CNN-BiGRU prediction method integrating the attention mechanism for submarine cable operation condition. Case studies are conducted and analyzed in Section 4. Finally, conclusion are drawn in Section 5.

2 Submarine cable operation condition evaluation model based on cooperative game theory and multi-level variable weight

2.1 Evaluation index system of submarine cable operation condition

Considering that many indicators can reflect the operating status of submarine cables, on the basis of fully considering the structural characteristics, operation history, maintenance, and fault data of submarine cables, combined with the "Operation Regulations for Submarine Power Cables (DL/T1278-2013)" issued by the National Energy Administration of China, this paper constructs an evaluation index system of submarine cable operation condition covering three aspects: online monitoring, routine inspection and static test, which integrates 21 important indicators for the submarine cable components, as shown in Figure 1.

2.2 Determination method of index layer weight based on cooperative game theory

2.2.1 Subjective weight based on improved analytical hierarchy process

AHP is the most commonly used weight determination method in the condition assessment research field. It has high practicability and simplicity, and its essence is to obtain its importance scale C (C = 1, 2, , 9) for subsequent calculation. Cause the value of C is relatively fixed, the result will deviate from the ideal expected value during the weight calculation. Besides, due to the ambiguity caused by scattered comparison, a consistency check is required (Cui et al., 2021; Ge et al., 2022). Therefore, an improved AHP method is proposed, and the process is as follows:

(1) Determine the importance of indicators.

In this paper, the score is used to replace the 1–9 scale method used by traditional AHP. First, the industry standards are comprehensively considered to obtain the importance score S of each indicator, and then the corresponding judgment matrix is obtained by comparing the importance scores of each indicator. The contrasted factor values are more accurate, and due to the certainty of the importance score, the proposed method does not need to be checked for consistency.

(2) Construct the judgment matrix.

Based on the importance value of each index obtained, a judgment matrix $\boldsymbol{C_1}$ is constructed by quantifying the comparison factor c_{ij} (i, j = 1, 2,...,n):

$$C_1 = \begin{bmatrix} 1 & c_{12} & c_{13} & \cdots & c_{1n} \\ c_{21} & 1 & c_{23} & \cdots & c_{2n} \\ c_{31} & c_{32} & 1 & \cdots & c_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & c_{n3} & \cdots & 1 \end{bmatrix} \tag{1}$$

The mathematical expression of c_{ij} is:

$$c_{ij} = S_i / S_j \tag{2}$$

where c_{ij} is the value of the i-th row and j-column of the constructed matrix, and S_i is the importance score of the i-th index.

(3) Determine the subjective weights.

The row means of the constructed judgment matrix after column normalization is the weight w, and its mathematical expression is:

$$c_{\mathrm{ij}}^{*} = c_{\mathrm{ij}} \Big/ \sum_{X=1}^{n} c_{\mathrm{ix}} \tag{3}$$

$$w_{j\_\mathrm{subjective}} = \frac{\sum_{X=1}^{n} c_{\mathrm{Xj}}^{*}}{n} \tag{4}$$

where c_{ij}* is the value after column normalization of the data in the i-th column and the j-th row, $w_{j\_subjective}$ is the subjective weight of the j-th index.

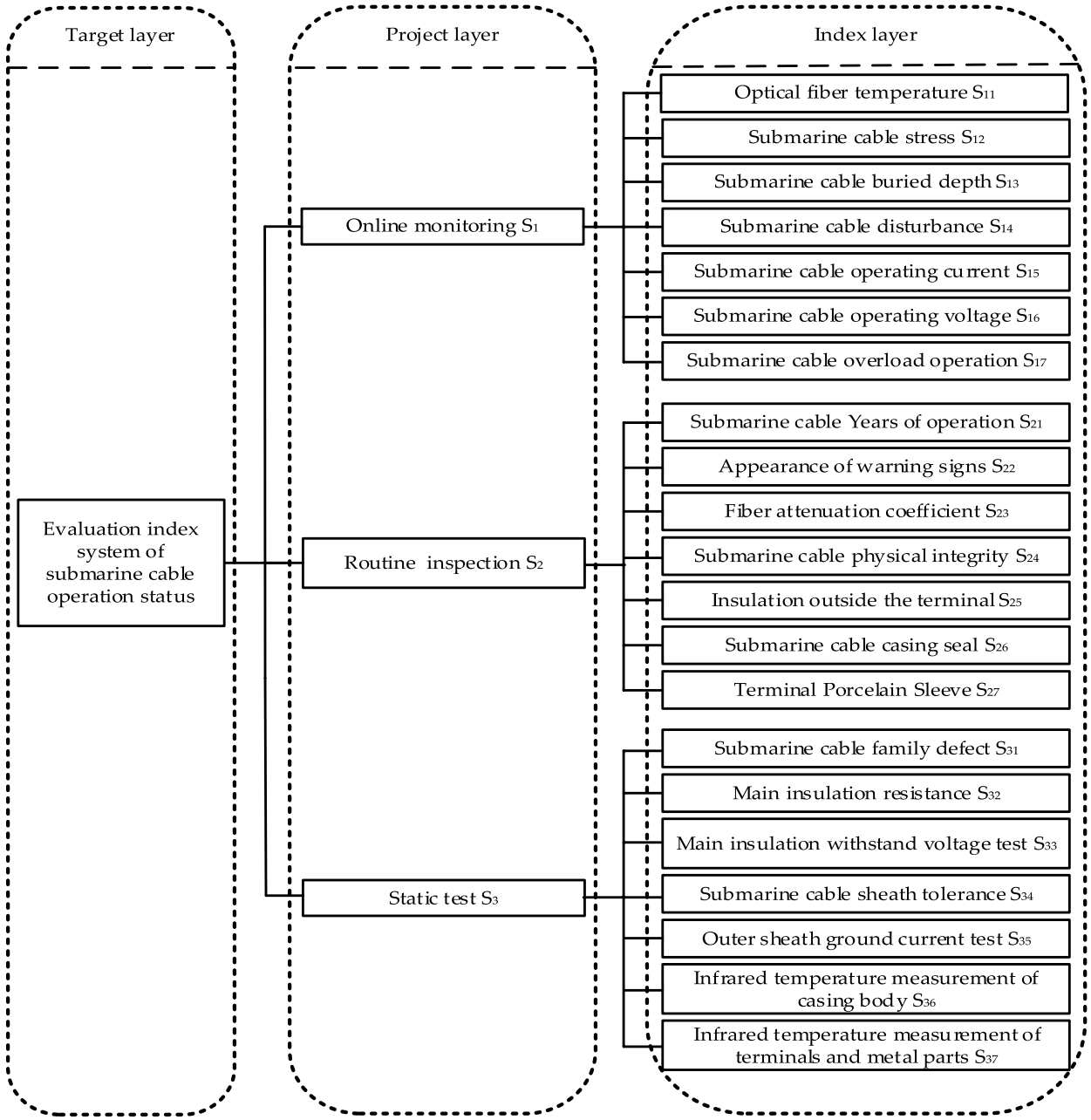


FIGURE 1
Evaluation index system of submarine cable operation condition.

2.2.2 Objective weight based on entropy weight method

Entropy is a measure for judging the degree of disorder of the system. The entropy weight measures the dispersion degree of each grass-roots index through the entropy value and then judges the degree of influence (weight) of the index on the comprehensive result (Zhu et al., 2020). The basic process is as follows: For the selected n indicators and m groups of evaluation data, the evaluation data is normalized to construct an evaluation matrix $\boldsymbol{C_2}$ $(m \times n)$:

$$C_2 = \begin{bmatrix} 1 & x_{12} & x_{13} & \cdots & x_{1n} \\ x_{21} & 1 & x_{23} & \cdots & x_{2n} \\ x_{31} & x_{32} & 1 & \cdots & x_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & x_{m3} & \cdots & 1 \end{bmatrix} \tag{5}$$

where x_{ij} is the i-th group of data of the j-th index.

The data of $\boldsymbol{C_2}$ is normalized by the range transformation method (that is, the worst value of each attribute after processing is 0, and the optimal value is 1) to calculate the data index ratio z_{ij}, and its mathematical expression is:

$$X^* = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{6}$$

$$X^* = \frac{X_{max} - X}{X_{max} - X_{min}} \tag{7}$$

$$z_{ij} = \frac{X_{ij}^*}{\sum_{i=1}^{n} X_{ij}^*} \left(i = 1, 2, ..., n; j = 1, 2, ..., m.\right) \tag{8}$$

where (7) and (8) are the range transformation formulas of benefits data and cost data, respectively. X^*, $X_{\max}$, and X_{min} are the value after range processing, the maximum value of the column, and the minimum value of the column, respectively. z_{ij} is the data index ratio of the j-th item evaluation index (when z_{ij} is 0, take $\ln z_{ij} = 0$).

According to the normalized data, the information entropy e_j of each index is obtained, and the entropy weight w of each evaluation index is finally calculated. Its mathematical expression is:

$$e_j = -\frac{1}{\ln n} \sum_{i=1}^{n} \left(z_{ij} \ln z_{ij}\right) \tag{9}$$

$$w_{j\_objective} = \frac{1 - e_j}{\sum_{j=1}^{m} \left(1 - e_j\right)} \tag{10}$$

where $w_{j\_objective}$ is the objective weight of the j-th index.

2.2.3 Determination of the weight of the indicator layer based on the cooperative game theory

The basic idea of the cooperative game theory is to combine several different types of weight methods and use the cooperative game idea as a whole to determine the combined weight, so as to ensure the rationality and accuracy of the weight determination (Li et al., 2021). Denote the weight by method i ($i = 1, 2, ..., n$) as $W(i)$, and the final combined weight W. The specific steps are as follows:

(1) Consistency correlation coefficient calculation.

Denote $W(m\_i)$ as the combined weight obtained by m-1 methods other than the i-th method, and the mathematical expression of the consistency correlation coefficient $L(i)$ is:

$$L(i) = \frac{\sum_{j=1}^{n} \left[W_j(i) - \overline{W(i)}\right]\left[W_j(m\_i) - \overline{W(m\_i)}\right]}{\left\{\sum_{j=1}^{n} \left[W_j(i) - \overline{W(i)}\right]^2\right\}^{1/2} \left\{\sum_{j=1}^{n} \left[W_j(m\_i) - \overline{W(m\_i)}\right]^2\right\}^{1/2}} \tag{11}$$

Where n is the number of evaluation indicators, and "—" means the average of the values.

(2) Calculate the combined weight $W'(i)$.

$$W' = \sum_{i=1}^{m} W(i) L(i) \tag{12}$$

(3) The combined weight is obtained recursively, that is, the number of weights is reduced by 1 after each calculation of the formula until the number of weights is 2.

(4) When the number of weights is 2, the mathematical expression of $W'(i)$ is:

$$W' = \frac{W(1) + W(2)}{2} \tag{13}$$

2.3 Project layer dynamic self-adaptive weight determination method

Regarding the weight coefficient of the project layer of the evaluation system is relatively fixed, due to the difficulty of inspection and testing for the submarine cable. A dynamic adaptive mechanism is used to determine the weight of the project layer: 1) when the submarine cable is in the inspection and static test stage, weights of monitoring, inspection indicators, and static tests are determined based on the guidelines and improved AHP. 2) when the inspection indicators and static tests are in a stable period, the weights of the indicators of each project layer are dynamically updated based on the expert evaluation results, so as to more accurately grasp the real-time operation status of submarine cables. Taking a certain period of submarine cable as an example, experts fully consider the historical condition and determine the importance score k ($k = 0, 1, \ldots, 5$) of each project layer according to factors such as the operating period of the submarine cable, and accumulate the importance score of the project layer. The mathematical expression of value K_i is as follows:

$$K_i = \sum_{j=1}^{n} K_{ij} \tag{14}$$

Where K_i is the sum of the expert scores at the i-th project level, and n is the number of experts participating in the scoring as shown in Table 1.

According to scoring results of the experts, the latest weight of each project layer is calculated by using the AHP method. The construction matrix and its results are shown in Table 2. The mathematical expression of the calculation process is as follows:

$$S_{ij}^* = S_{ij} \Big/ \sum_{X=1}^{n} S_{ix} \tag{15}$$

TABLE 1 Expert scoring table of project-level importance at a certain time.

Expert	A	B	C	D	K
Index					
Online monitoring	4	4	5	5	18
Routine inspection	2	2	3	4	11
Static test	2	2	3	4	11

$$w_{j\_project} = \frac{\sum_{X=1}^{n} S_{ij}^{*}}{n} \tag{16}$$

where $S_{ij}*$ is the value after column normalization of the data in the i-th column and the j-th row, and $w_{j\_project}$ is the row mean (weight).

2.3 Comprehensive condition evaluation of submarine cable

The mathematical expression of the final health value M of the submarine cable is as follows:

$$M = mK_T \tag{17}$$

where K_T is the life coefficient, m is the sum of the evaluation index of the submarine cable.

The mathematical expressions of K_T and m are:

$$K_T = \frac{100 - \lambda \times 0.5}{100} \tag{18}$$

$$m = \sum_{i=1}^{n} m_i w_i \tag{19}$$

where n is the number of index layers, λ is operation years, w_i is the weight coefficient of each project layer, m_i is the evaluation value, and the mathematical expression is:

$$m_i = \sum_{j=1}^{n_i} w_{ij} g_{ij} K_{ij} \tag{20}$$

where w_{ij} is the grass-roots index weight, K_{ij} is the family defect coefficient, it equals 0.95 if there is a family defect, and 1 if there is no family defect. g_{ij} is the evaluation value of a single indicator and its mathematical expression is:

$$g_{ij} = 100 \times \left(1 - \left|1 - \frac{y_{wr} - y_{now}}{y_{wr} - y_{st}}\right|\right) \tag{21}$$

Where y_{wr}, y_{now}, and y_{st} are the warning value, monitoring value, and standard value of the indicator, respectively.

The obtained comprehensive health assessment value M of the submarine cable is used as the predicted output label to calibrate the operation condition of the submarine cable. The numerical segmentation of M is as follows:

TABLE 2 AHP calculation of project layer weight.

	S_1	S_2	S_3	Weight (w_j)
S_1	1	18/11	18/11	0.45
S_2	11/18	1	1	0.275
S_3	11/18	1	1	0.275

$$\text{Assessment status} = \begin{cases} \text{normal} & (M \geq 90) \\ \text{notic} & (90 > M \geq 75) \\ \text{abnormal} & (75 > M \geq 60) \\ \text{serious} & (M < 60) \end{cases} \tag{22}$$

3 Convolutional neural network-Bidirectional gated recurrent unit prediction model integrating attention mechanism

3.1 Principle of deep learning

3.1.1 Convolutional neural network

The advantage of CNN is to mine the potential associations and local features of the data to improve the efficiency and accuracy of the model with higher fault tolerance. The core of its structure is the convolution layer, which obtains potential features through the convolution operation between the convolution kernel and the input data. The mathematical expression is:

$$x_{ji} = f\left(x_{j-1} * w_{ji} + b_{ji}\right) \tag{23}$$

where x_{ji} is the i-th feature of the output of the j-th layer, $x_{j\text{-}1}$ is the output of the j-1th layer, w_{ji} is the weight matrix of the i-th convolution kernel of the j-th layer, b_{ji} is the bias term, and $*$ is the convolution operator.

The relu activation function is used for all of the models in this paper, and its mathematical expression is:

$$f(z) = \begin{cases} 0 & (z < 0) \\ z & (z \geq 0) \end{cases} \tag{24}$$

The structure diagram of CNN is shown in Figure 2 (Li et al., 2013).

3.1.2 Bidirectional gated recurrent unit

LSTM neural network is the most commonly used network model for processing sequence data. It can achieve long-term memory through a unique gating unit, and then effectively solve the gradient problem existing in traditional recurrent neural networks. Compared with LSTM, GRU has a simpler structure, fewer parameters, higher computation efficiency,

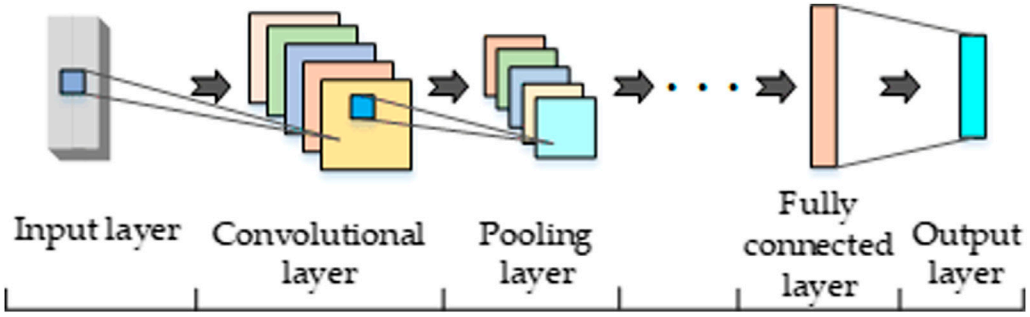


FIGURE 2
Structure diagram of CNN.

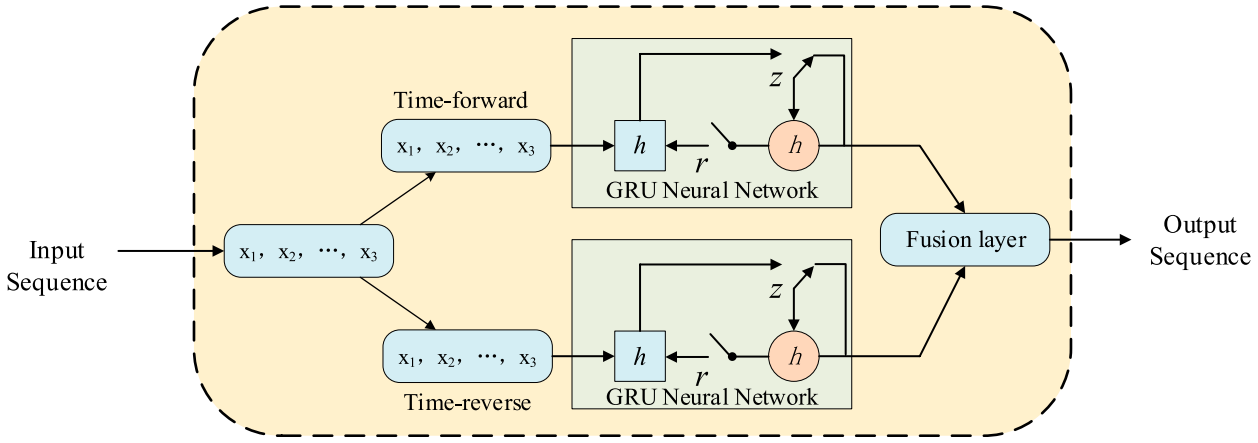


FIGURE 3
The structure of the BiGRU network.

and better generalization. It filters and transmits useful information through the update gate and reset gate in the unique neuron state for subsequent data processing, discards low-value information, and outputs the hidden state at each moment, so as to achieve the key information extraction and solve the gradient discrete problem. BiGRU network makes the data flow in two directions based on GRU, realizing deeper feature mining and more efficient use of data, and the mathematical expression of the GRU gating unit is as follows (Li et al., 2022; Yang et al., 2022):

$$z_t = \sigma\left(W^{(z)}x_t + U^{(z)}h_{t-1}\right) \tag{25}$$

$$r_t = \sigma\left(W^{(r)}x_t + U^{(r)}h_{t-1}\right) \tag{26}$$

$$h_{t2} = tanh\left(Wx_t + U\left(r_t * h_{t-1}\right)\right) \tag{27}$$

$$h_{t1} = (1 - z_t) * h_{t-1} + z_t * h_{t2} \tag{28}$$

where z_t and r_t represent the update gate and reset gate; σ is the Sigmoid function; tan$h$ is the hyperbolic tangent function; $W^{(z)}$, $W^{(r)}$, $U^{(z)}$, $U^{(r)}$, U, W are the matrix of training parameters, $*$ represents the product of the matrix, and h_{t2} is the summary of the input x_t and the output result of the previously hidden layer h_{t-1}.

The BiGRU structure is shown in Figure 3 (Zhang et al., 2022).

3.1.3 Attention mechanism

The essence of AM is to give higher weights to key features and reduce the weights of other parts accordingly, so as to extract key information from the feature data (Niu et al., 2022; Deng et al., 2021). In this paper, the features extracted from the CNN-BiGRU combined neural network model are fused with the attention mechanism: the CNN is used to extract the static features of the data, denoted as C, and the BiGRU is used to extract the time-series features of the data, denoted as G, and the formula of its fusion is (30). The correlation between the data feature and the i-th historical feature F_i is measured by the dot product operation, and then the attention value A is obtained. As shown in Eq. 31, the weighted sum of the historical features is calculated according to the attention value, namely the final data feature F_A (Lin et al., 2022). Through the fusion of feature extraction and attention mechanism performed by the CNN-BiGRU

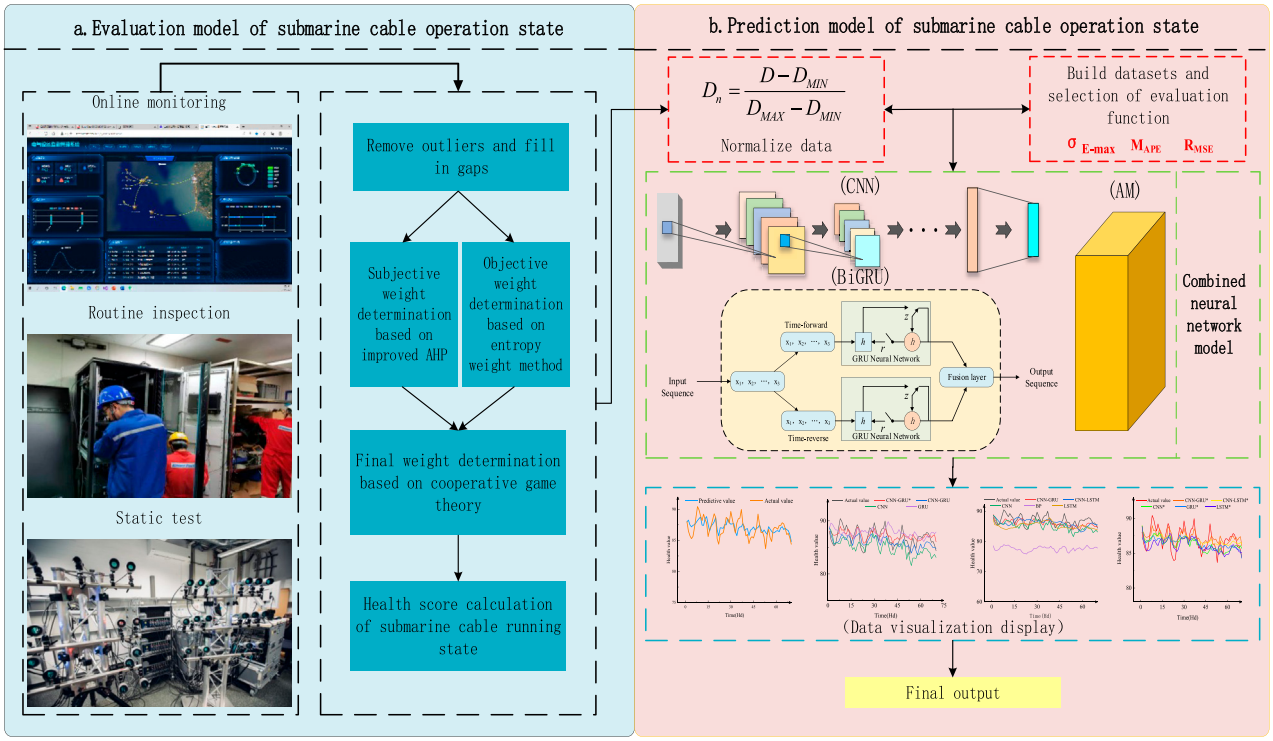


FIGURE 4
Prediction model of submarine cable operation condition considering condition evaluation.

TABLE 3 Experimental results of model network layer optimization.

BiGRU network layers	$\sigma_{E\text{-}max}$ (%)	M_{APE} (%)	R_{MSE} (%)
1	3.81	1.63	1.47
2	3.35	1.25	1.36
3	3.92	1.55	1.40
4	7.61	4.23	3.44

combined neural network, the model can fully extract the static and time-series features of historical data, and dynamically adjust them according to the importance of different features, so as to obtain better results.

$$F = tanh\left(GW_aC^T + b_a\right) \tag{29}$$

$$A = \frac{exp(F)}{\sum_{i=1}^{n} exp(F_i)} \tag{30}$$

$$F_A = \sum_{i=1}^{n} AF_i \tag{31}$$

where b_a is the bias term obtained by training, F is the fused data feature, F_A is the final data feature, W_a is an adaptive weight matrix obtained by training the attention mechanism.

3.2 Prediction model

Compared with the traditional neural network, CNN can efficiently and accurately extract the inherent features of the data set, but it cannot effectively use the time-series features existing in the data. Although a single BiGRU neural network can effectively extract and utilize the temporal features in the dataset, it is relatively inefficient for the use of correlation features in the data. When using a single CNN and BiGRU neural network to process long-term sequences or face multi-dimensional input data, problems such as insufficient data feature mining and loss of sequence or associated features will occur. Therefore, this paper complements the advantages of these methods, and a CNN-BiGRU combined neural network integrating the attention mechanism is proposed to improve the prediction performance.

The framework of the proposed model for the condition prediction of the submarine cable is shown in Figure 4. First, the historical operation data of the submarine cable and the comprehensive health value quantification are processed through the submarine cable operation condition evaluation model. Then, the historical operation data and condition quantification results are imported into the proposed combined neural network model to realize the operation condition prediction of the submarine cable.

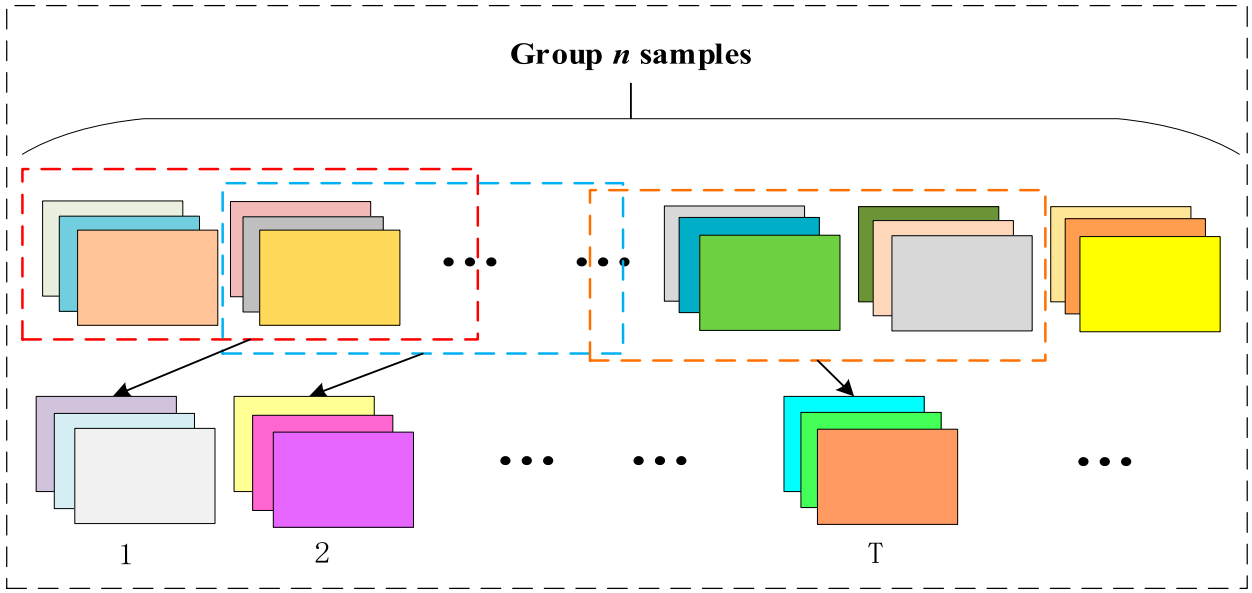
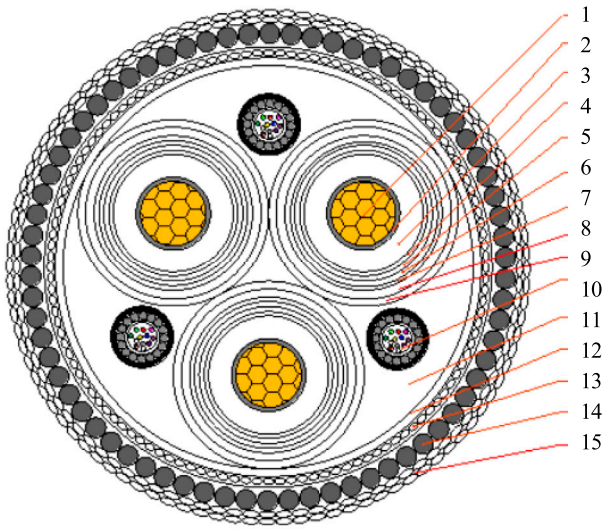


FIGURE 5
Sliding window data reading mode.



Structural description
1.Water blocking copper conductor
2.Conductor shield
3.XLPE insulation
4.Insulation shiel
5.Semiconducting resistance hose
6.Copper tape shield
7.Semiconducting resistance hose
8.Alloy lead sleeve
9.Asphalt+PE sheath
10.Fiber unit
11.Filling
12.Cable wrapping
13.PP inner cushion
14.Galvanized steel wire armor
15.Asphalt+PP outer quilt

FIGURE 6
Structure diagram of HYJQF41-F 35 KV submarine cable.

3.3 Prediction process

3.3.1 Data processing

Take each index value of the submarine cable operation status evaluation index system established in this paper as the original import data, and time step is half a day (Hd). In order to solve the problem of large error caused by abnormal data in the data set, the mean square method is adopted to process the data set, and its mathematical expression is as follows:

TABLE 4 Online monitoring index and evaluation results.

Evaluation index	Standard value	Warning value	Monitor value	Single item score	Weighted score
S_{11}	23.00	35.00	23.57	95.26	19.05
S_{12}	0.00	150.00	11.84	92.11	3.74
S_{13}	2.00	1.60	2.00	100.00	23.26
S_{14}	0.00	40,000.00	3,398.00	91.50	13.83
S_{15}	138.00	150.00	137.00	91.67	14.94
S_{16}	35,000.00	37,000.00	34,844.00	92.20	9.17
S_{17}	0.00	1.00	0.00	100.00	11.32
Health value	——	——	——	——	95.31
Evaluation status	——	——	——	——	Normal

$$\begin{cases} \delta^2 = \sum_{i=1}^{N} (x_i - \mu)^2 / (N-1) \\ \mu = \sum_{i=1}^{N} x_i / N \end{cases} ; \ i = 1, 2, ..., N \tag{32}$$

where N is the numbers of data of this type, and x_i is the value of the i-th evaluation. If $|x_i - u| > 3\delta$, it is judged that x_i is an abnormal point, and the point is eliminated.

In order to solve the negative impact caused by the different dimensions and the large differences of the input data, before the model training, the maximum and minimum values of various data in the data set are used as the benchmark to normalize the data set. The formula is:

$$D_n = \frac{D - D_{MIN}}{D_{MAX} - D_{MIN}} \tag{33}$$

where D_n is the data obtained after normalization, D is the data before processing, D_{MIN} and D_{MAX} are the minimum and maximum values of this type of data, respectively.

3.3.2 Predictive evaluation indicators

In order to evaluate the prediction accuracy of the proposed method, the maximum percentage of prediction error $\sigma_{E\text{-}max}$, the root mean square error R_{MSE}, and the mean absolute percentage error M_{APE} were selected. The mathematical expression is as follows (Liu et al., 2019b):

$$\sigma_{E-max} = max\left(\frac{|y_i - y_{pi}|}{y_i}\right) \times 100\% \tag{34}$$

$$R_{MSE} = \sqrt{\frac{\sum_{i=1}^{N} (y_i - y_{pi})^2}{N}} \times 100\% \tag{35}$$

$$M_{APE} = \sum_{i=1}^{N} \left(\left| \frac{y_{pi} - y_i}{y_i} \right| \times \frac{100\%}{N} \right) \tag{36}$$

where y_i is the actual value of the i-th sample point; y_{pi} is the predicted value of the i-th sample point; N is the number of sample points.

It is worth noting that in order to ensure the scientificity and unity of the model training and prediction process, the R_{MSE} and M_{APE} of the prediction results are the average values obtained by the models.

3.3.3 Model training

In this paper, the control variable method is used to optimize the structure of the proposed model. Regarding the importance of the number of BiGRU network layers, the prediction performance is tested by continuously increasing the number of BiGRU layers. The basic parameters of the CNN module and attention module remain unchanged, and the effect of the BiGRU layer on the prediction results is tested. Experiments show that when the number of BiGRU layers is 2, all of the evaluation indexes reach the optimal values, and then the error rate begins to rise, indicating that the model is over-learning. The results are shown in Table 3.

The number of convolution kernels of the CNN neural network is set as 10, the size of the convolution kernel is set as 2, and the step size is set as 1, the same convolution is selected for the convolution method, and valid maximum pooling is performed after continuous convolution. The BiGRU neural network adopts a sliding-window data reading mode, as shown in Figure 5, the time step is set to 10, and the batch size is set to 30 (Su et al., 2021; Jinah et al., 2022). The Adam algorithm is used to iteratively update the weights so that the weights and deviations of each neuron are continuously updated through the momentum and adaptive learning rate, so as to optimize the output value of the loss function. To deal with the possible overfitting problem, the Dropout method is used during the training. In the process of randomly discarding the network nodes from the network according to a certain probability, the attention mechanism is integrated into the training of CNN and BiGRU to improve the accuracy of the model. The mathematical expression of the loss function is:

$$F_{loss} = \frac{1}{n} \sum_{t=1}^{n} \left(\lambda_{act}(t) - \lambda_{pred}(t) \right)^2 \tag{37}$$

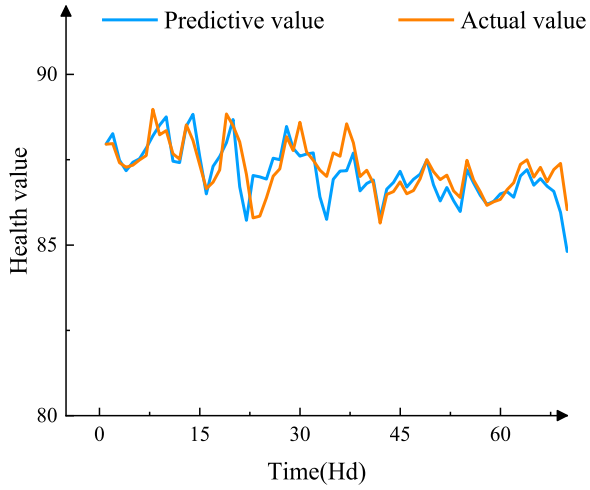


FIGURE 7
Prediction results.

where F_{loss}, $\lambda_{act}(t)$, $\lambda_{pred}(t)$, and n are the loss function of the model, the actual and predicted health values at time t, and the number of training samples, respectively.

4 Simulation verification

4.1 Data selection

In order to verify the feasibility and accuracy of the proposed method in this paper, the HYJQF41-F 35 KV three-core submarine cable of an offshore oil and gas platform in the Bohai Sea is taken as the research object. Its structure is shown in Figure 6. From 2021 to 2022, a total of 510 groups of monitoring and routine inspection data in 255 days are used. The training set is divided according to 88%, and the rest of the data is used as a test set, and the predicted results are compared with actual values.

4.2 Comprehensive condition evaluation and analysis

Taking a certain set of data selected in this paper as an example, the single-item evaluation index data of online monitoring is shown in the Table 4. Substitute the standard value, attention value, and monitoring value of each index into Eq. 21 to obtain the single-item evaluation score, and then substitute it into Eq. 20 to obtain a single weighted score, and the cumulative health value of the online monitoring module is 95.31, and its operating condition is normal, which is consistent with the actual situation on site, as shown in Table 4.

4.3 Condition prediction and analysis

4.3.1 Prediction results

The selected data set is evaluated and quantified by the evaluation model of the submarine cable operation condition proposed in this paper, and then the input feature set is formed and imported into the constructed CNN-BiGRU-AM prediction model. The prediction results are shown in Figure 7. As can be seen from Figure 7, the overall health status value of submarine cables fluctuated between 80 and 90 during this period. Due to the influence of various factors, it fluctuates within a certain range, but it is generally stable and shows a certain downward trend, which is in line with the actual condition of the submarine cable. Besides, the three evaluation indicators: the maximum percentage of prediction error $\sigma_{E\text{-}max}$, the root mean square error R_{MSE}, and the average percentage error M_{APE} are 3.35%, 1.36%, and 1.25%, respectively, which illustrates its prediction accuracy, and can provide effective support for the operation decision of submarine cables.

4.3.2 Comparative analysis

In order to verify the effectiveness and superiority of the proposed model in this paper, the same data set is input into CNN, LSTM, CNN-LSTM, and other neural network models. The training time required for a single round and single step of each model are recorded, and the results are shown in Table 5. It is noted that the right column of the table indicates that the model incorporates an attention mechanism.

From Table 5, it can be seen that the time spent by each neural network model is prolonged after combining with each other or integrating an attention mechanism. Among them, CNN has the fastest training rate, LSTM has the longest time due to its cumbersome gate structure, and the BiGRU network has certain advantages in terms of speed.

In order to verify the validity of the prediction model proposed above, the maximum percentage of prediction error $\sigma_{E\text{-}max}$, the root mean square error R_{MSE}, and the mean absolute percentage error M_{APE} are calculated respectively. The comparison results of various indicators are shown in Table 6 and Figures 8, 9.

According to the comparison information of different models in Table 6, compared with other similar neural network models, CNN-BiGRU-AM has the best prediction accuracy in each evaluation index, and further combined with the box plot shown in Figure 9, we can see that the average error and discrete error of the model are obviously dominant. From the perspective of whether the neural network model is combined or not, compared with the single neural network model, the maximum error percentage $\sigma_{E\text{-}max}$, the average absolute percentage M_{APE} and the root mean square error R_{MSE} of the combined neural network model are improved, and the comparison results are shown in Figure 10.

The attention mechanism is an important module of the proposed model optimization and feature fusion. From the

TABLE 5 Comparison of running speed of each model.

Prediction model	Time required for training (single round/single step)	
	Without AM	With AM
BP	-/4 ms	—
CNN	13 ms/1 ms	15 ms/2 ms
LSTM	41 ms/4 ms	42 ms/5 ms
GUR	39 ms/3 ms	40 ms/4 ms
CNN-LSTM	43 ms/5 ms	44 ms/5 ms
CNN-BiGRU	40 ms/4 ms	42 ms/5 ms

TABLE 6 Comparison of prediction results of various models.

Prediction model	$\sigma_{E\text{-}max}$		M_{APE}		R_{MSE}	
	Without AM (%)	With AM	Without AM (%)	With AM	Without AM (%)	With AM
BP	13.75	—	9.61	—	9.19	—
CNN	5.20	4.13%	2.39	1.37%	2.43	1.43%
LSTM	6.95	4.51%	2.32	1.52%	2.44	1.62%
GRU	5.86	3.95%	2.10	1.41%	2.14	1.46%
CNN-LSTM	5.31	3.55%	1.99	1.30%	2.08	1.38%
CNN-BiGRU	4.89	3.35%	1.79	1.25%	1.60	1.36%

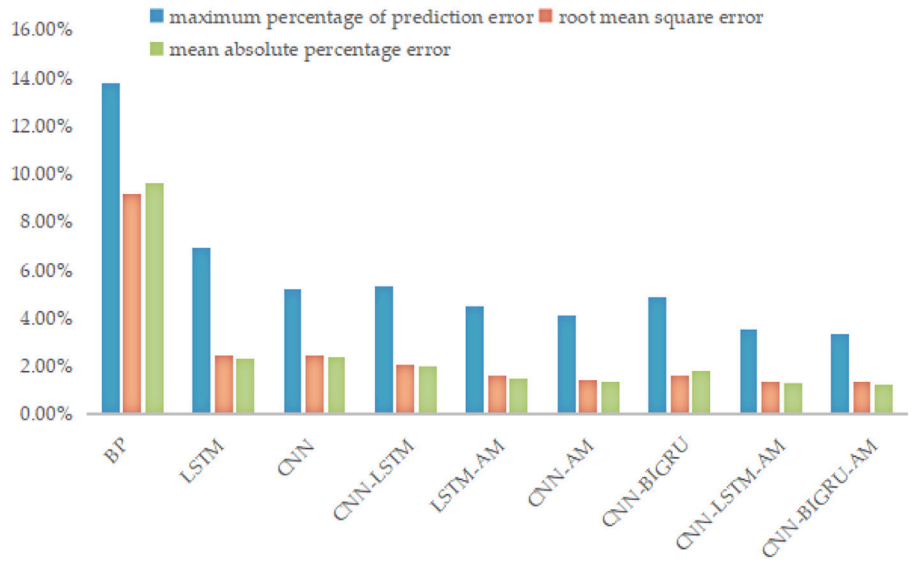


FIGURE 8
Comparison of various models.

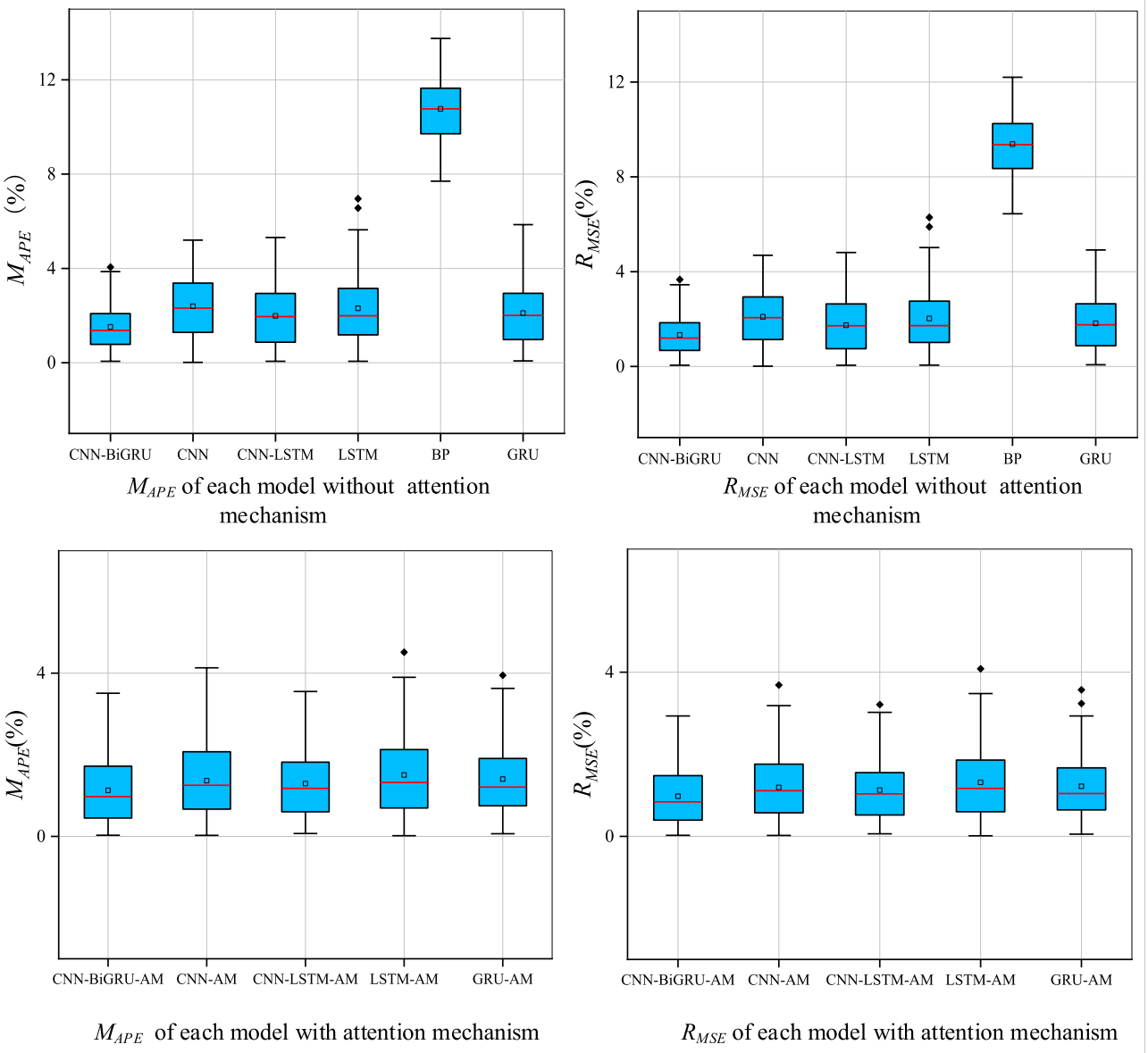


FIGURE 9
M_{APE} and R_{MSE} of different models with and without attention mechanism.

perspective of whether the model integrates the attention mechanism or not, the maximum percentage of prediction errors $\sigma_{E\text{-}max}$, the mean percentage error M_{APE} and root mean square error R_{MSE} increased by 1.54%, 0.3%, and 0.26%, respectively. Compared with other similar neural network models, each evaluation index has been improved by up to 9.03%, 7.72%, and 6.81%, reaching 3.35%, 1.25%, and 1.36%, respectively. The performance of each models with and without attention mechanism are shown in Figures 11, 12.

It can be seen from Figure 11 that the predicted value of the BP neural network is significantly lower than the actual value, but it has a certain trend consistency, which shows that its processing capability for long-sequence multi-feature samples (such as submarine cable operation data) is relatively general. CNN prediction results have a good fluctuation fit, but due to low numerical value, the accuracy is not ideal, which may be caused by the insufficient extraction of sample time series features. LSTM network has better learning efficiency for time series data, although the maximum error value is relatively high. But due to the lack of correlation features, there is room for further optimization. CNN-LSTM is better than the above models in terms of overall fitting, and remains relatively stable when the fluctuation is small. The fluctuations and trends of the CNN-BiGRU network are similar to the CNN-LSTM network, both

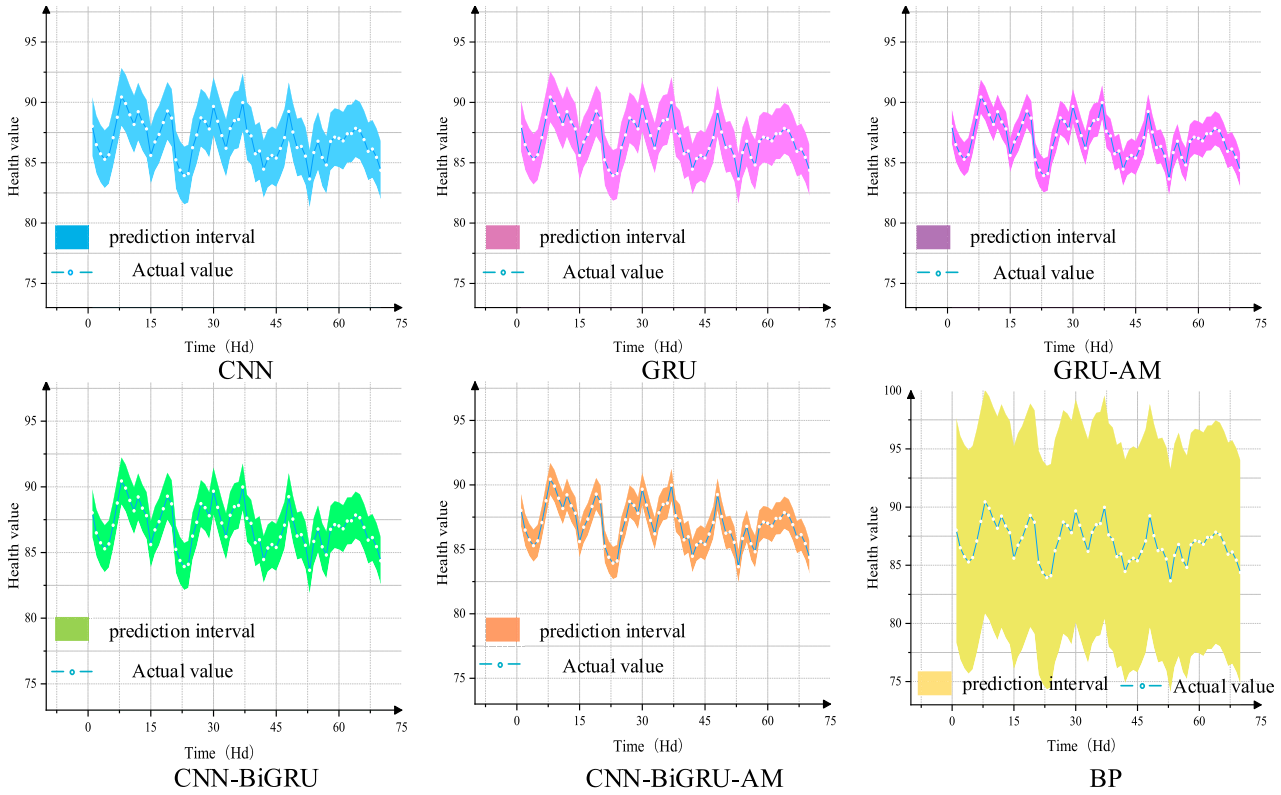


FIGURE 10
Effect comparison of important single and combined models.

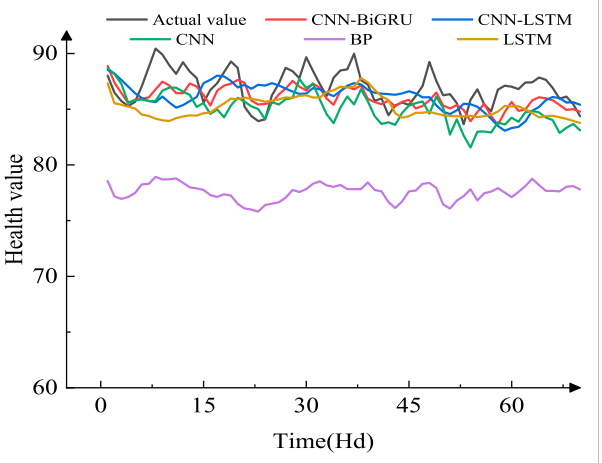


FIGURE 11
Comparison of prediction results of different models without attention of different models without attention.

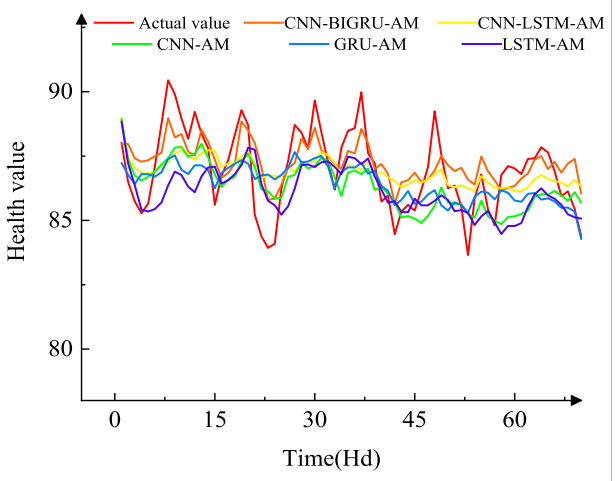


FIGURE 12
Comparison of prediction results of different models without attention of different models with attention.

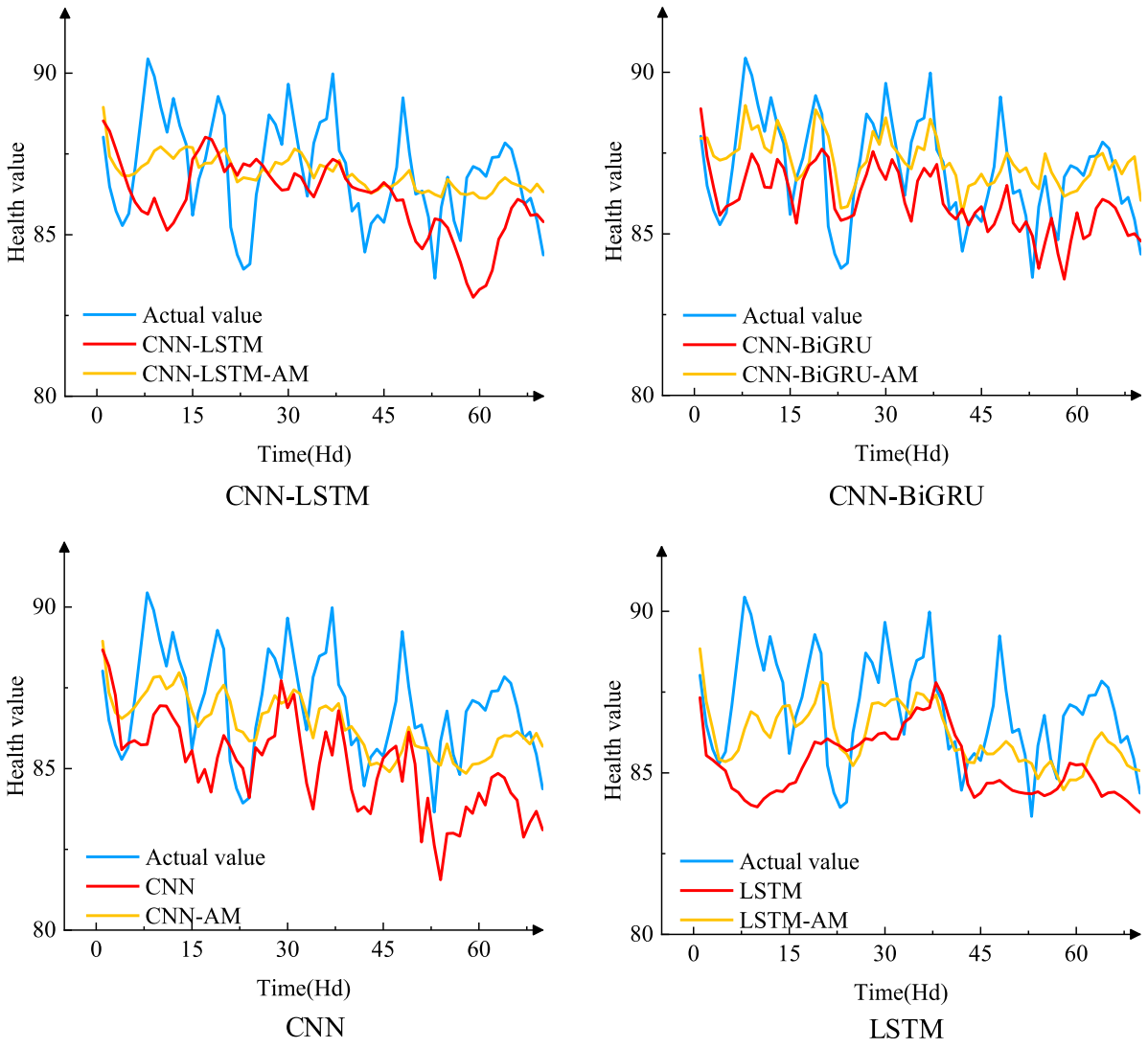


FIGURE 13
The effect of attention mechanism.

greater advantages and higher accuracy, but the former dominates in training efficiency (Table 5). From Figure 12, it can be found that after adding the attention mechanism, the curve fit of each neural network prediction model has been improved to varying degrees, which shows the effectiveness of the attention mechanism in neural network time series prediction. According to whether the unified network adds attention mechanism as a variable, the comparison effect is shown in Figure 13.

To sum up, the numerical example verifies the superiority of the model described in this paper, and the evaluation indexes (i.e. maximum error, root mean square error, and average percentage error) are better than other methods, and it can accurately predict the operation condition of submarine cables.

5 Conclusion

Aiming at the shortcomings of the existing research in the field of submarine cable operation condition evaluation and prediction, a submarine cable operation status evaluation index system is established in this paper, which includes online monitoring, routine inspection and static test. To obtain more scientific evaluation results, based on the improved AHP and cooperative game theory, a submarine

cable operation condition evaluation model is constructed, and numerical quantification is carried out in combination with the guidelines and engineering practice. Besides, the CNN-BiGRU-AM combined neural network model is established by synthesizing the advantages of correlation feature extraction of CNN, the time-series feature extraction of BiGRU and the feature fusion advantages of the attention mechanism. By comparing with other neural network models, the better prediction accuracy verifies the advantages of the proposed model in processing long-sequence, multi-characteristic submarine cable operation data samples.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

Conceptualization, WY, and BH; Methodology, WY, and BH; Validation, WY, and JL; Writing—original draft preparation, BH, and AZ; Writing—review and editing, BH and WY; Supervision, QL and XX.

Funding

This work was supported by the Sichuan (China) Science and Technology program (2020YFSY0037, 2022YFG0123), Central Government Funds for Guiding Local Scientific and Technological Development of China (2021ZYD0042, 2021ZYD0099).

Conflict of interest

Author XX was employed by PetroChina Dushanzi Petrochemical Company.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Antonio, M., Mattia, C., Jorge, C., Valey, K., Rafael, M., and Zhan, Z. (2021). Polarization sensing using submarine optical cables. *Optica* 8 (6), 788–795. doi:10.1364/OPTICA.424307

Bastien, T., Juan, B., Andrew, W., Gérard, T., Morgane, L., Nicolas, D., et al. (2018). A review of potential impacts of submarine power cables on the marine environment: Knowledge gaps, recommendations and future directions. *Renew. Sustain. Energy Rev.* 96, 380–391. doi:10.1016/j.rser.2018.07.026

Cao, C., Ge, Y., Ren, X., Peng, X., Chen, J., Lu, Z., et al. (2022). Experimental research on submarine cable monitoring method based on MEMS sensor. *Micro Nano Eng.* 15, 100130. doi:10.1016/j.mne.2022.100130

Chen, X., Zou, N., Liang, L., He, R., Liu, J., Zheng, Y., et al. (2021). Submarine cable monitoring system based on enhanced COTDR with simultaneous loss measurement and vibration monitoring ability. *Opt. Express* 29 (9), 13115–13128. doi:10.1364/OE.418920

Chen, Y., Wang, S., Hao, Y., Yao, K., Li, H., Jia, F., et al. (2020). The 500kV oil-filled submarine cable temperature monitoring system based on BOTDA distributed optical fiber sensing technology, in 2020 International Conference on Sensing, Measurement & Data Analytics in the era of Artificial Intelligence (ICSMD), 15-17 October 2020, Xi'an, China. (New Jersey, United States: IEEE).

Cui, X., Lu, F., He, Y., Ma, J., Liu, S., Xue, Q., et al. (2021). Two-dimensional evaluation model of electrical equipment based on combined weighting and rating algorithm. *Energy Rep.* 7, 443–448. doi:10.1016/j.egyr.2021.01.048

Dai, J., Song, H., Sheng, G., Jiang, X., Wang, J., and Chen, Y. (2018). Prediction method for power transformer running state based on LSTM network(article). *High. Volt. Eng.* 44 (4), 1099–1106. doi:10.13336/j.1003-6520.hve.20180329008

Deng, Y., Liu, X., Jia, R., Huang, Q., Xiao, G., and Wang, P. (2021). Sag source location and type recognition via attention-based independently recurrent neural network. *J. Mod. Power Syst. Clean Energy* 9, 1018–1031. doi:10.35833/MPCE.2020.000528

Fang, J., Peng, X., Liu, T., Chen, Y., Li, W., Wen, J., et al. (2020). Overview of big data development of power equipment condition monitoring. *Power Syst. Prot. Control* 48 (23), 176–186. doi:10.19783/j.cnki.pspc.200050

Ge, Q., Qiao, H., Li, C., Yang, Q., and Jiang, H. (2022). Real-time charging risk assessment for electric vehicles based on improved broad BP-AHP. *IEEE Trans. Ind. Electron.* 69 (9), 9472–9482. doi:10.1109/TIE.2021.3111591

Jianfeng, W. (2019). Monitoring method of submarine cable fault data. *J. Coast. Res.* 35, 735–740. doi:10.2112/SI93-102.1

Jinah, K., Minho, K., Sung-Dae, K., Jong-Seong, K., Joon-Gyu, R., and Jaeil, Kim. (2022). Spatiotemporal neural network with attention mechanism for El Niño forecasts. *Sci. Rep.* 12 (1), 7204–7215. doi:10.1038/s41598-022-10839-z

Lei, P. (2021). Design and implementation of ROV submarine cable inspection decision support system. *IOP Conf. Ser. Earth Environ. Sci.* 809, 012003. doi:10.1088/1755-1315/809/1/012003

Li, D., Sun, G., Miao, S., Gu, Y., Zhang, Y., and He, S. (2022). A short-term electric load forecast method based on improved sequence-to-sequence GRU with adaptive temporal dependence. *Int. J. Electr. Power & Energy Syst.* 137, 107627. doi:10.1016/j.ijepes.2021.107627

Li, L., Chen, K., Gao, J., Liu, J., Gao, Z., and Wang, M. (2021). Multiobjective optimization method and application of tolerance allocation for the steam turbine based on cooperative game theory. *Shock Vib.* 2021, 1–13. doi:10.1155/2021/9244389

Li, L., Zhang, D., Xie, L., Yu, B., and Lin, F. (2013). Power transformer condition assessment method using association rules comprehensive analysis and variable weight coefficients. *Chin. J. Electr. Eng.* 24, 152–159. doi:10.13334/j.0258-8013.pcsee.2013.24.019

Lin, J., Ma, J., Zhu, J., and Cui, Y. (2022). Short-term load forecasting based on LSTM networks considering attention mechanism. *Int. J. Electr. Power & Energy Syst.* 137, 107818. doi:10.1016/j.ijepes.2021.107818

Liu, X., Liu, Y., Liu, J., Xiang, Y., and Yuan, X. (2019). Optimal planning of AC-DC hybrid transmission and distributed energy resource system: Review and prospects. *CSEE J. Power Energy Syst.* 5 (3), 409–422. doi:10.17775/CSEEJPES.2019.00540

Liu, Y., Xu, Z., Li, G., Xia, Y., and Gao, S. (2019a). A review of the application of artificial intelligence-driven data analysis technology in power transformer condition maintenance. *High. Volt. Technol.* 45 (2), 337–348. doi:10.13336/j.1003-6520.hve.20190130001

Liu, Y., Zhao, J., Xu, L., Qiu, G., and Liu, J. (2019b). Online TTC estimation using nonparametric analytics considering wind power integration. *IEEE Trans. Power Syst.* 34 (1), 494–505. doi:10.1109/TPWRS.2018.2867953

Liu, Z., Hao, J., Liao, R., Li, J., Gao, Z., and Liang, Z. (2019). Morphological, structural, and dielectric properties of thermally aged AC 500 kV XLPE submarine cable insulation material and its deterioration condition assessment. *IEEE Access* 7, 165065–165075. doi:10.1109/ACCESS.2019.2953127

Lv, A., Li, Y., Li, J., and Wu, F. (2014). Strain/temperature monitoring of optical fiber composite submarine cable based on BOTDR. *High. Volt. Technol.* 2, 533–539. doi:10.13336/j.1003-6520.hve.2014.02.028

Miguel, Á., Bachiller-Soler, A., Cruz-Romero, P., Juan, C., and Rosendo-Macias, J. A. (2021). A thermal model for three-core armored submarine cables based on distributed temperature sensing. *Energies* 14 (13), 3897. doi:10.3390/en14133897

Nie, Y., Zhao, X., and Li, S. (2020). Research progress on condition monitoring and insulation diagnosis of XLPE cables. *High. Volt. Technol.* 46 (4), 1361–1371. doi:10.13336/j.1003-6520.hve.20200430028

Niu, D., Yu, M., Sun, L., Gao, T., and Wang, K. (2022). Short-term multi-energy load forecasting for integrated energy systems based on CNN-BiGRU optimized by attention mechanism. *Appl. Energy* 313, 118801. doi:10.1016/j.apenergy.2022.118801

Pu, T., Qiao, J., Han, X., Zhang, G., and Wang, X. (2020). Research and application of artificial intelligence technology in operation and maintenance of power equipment. *High. Volt. Technol.* 46 (2), 369–383. doi:10.1088/1755-1315/617/1/012001

Purvins, A., Sereno, L., Ardelean, M., Covrig, C., Efthimiadis, T., and Minnebo, P. (2018). Submarine power cable between europe and north America: A techno-economic analysis. *J. Clean. Prod.* 186, 131–145. doi:10.1016/j.jclepro.2018.03.095

Song, H., Dai, J., Luo, L., Sheng, G., and Jiang, X. (2018). Power transformer operating state prediction method based on an LSTM network. *ENERGIES* 11 (4), 914. doi:10.3390/en11040914

Su, T., Liu, Y., Zhao, J., and Liu, J. (2021). Probabilistic stacked denoising autoencoder for power system transient stability prediction with wind farms. *IEEE Trans. Power Syst.* 36 (4), 3786–3789. doi:10.1109/TPWRS.2020.3043620

Tang, T., and Yuan, H. (2022). A hybrid approach based on decomposition algorithm and neural network for remaining useful life prediction of lithium-ion battery. *Reliab. Eng. Syst. Saf.* 217, 108082. doi:10.1016/j.ress.2021.108082

Wei, X., Zhu, B., Pang, B., Chen, Q., Wang, S., and Li, R. (2015). Study on on-line jacket insulation monitoring for 500 kV submarine oil-filled cable. *Electr. Mach. Control* 19 (5), 21–29. doi:10.15938/j.emc.2015.05.004

Xie, Y., Wang, C., and Huang, J. (2022). Structure and evolution of the submarine cable network of Chinese mainland. *J. Geogr. Sci.* 32 (5), 932–956. doi:10.1007/s11442-022-1979-4

Xu, Y., and Jiang, X. (2022). Short-term power load forecasting based on BiGRU-Attention-SENet model. *Energy Sources, Part A Recovery, Util. Environ. Eff.* 44 (1), 973–985. doi:10.1080/15567036.2022.2053250

Xu, Z., Hu, Z., Zhao, L., Zhang, Y., Yang, Z., Hu, S., et al. (2019). Application of temperature field modeling in monitoring of optic-electric composite submarine cable with insulation degradation. *Measurement* 133, 479–494. doi:10.1016/j.measurement.2018.10.028

Yang, W., Pu, C., Yang, K., Zhang, A., and Qu, G. (2022). Transformer short-term fault prediction method based on cnn-gru combined neural network. *Power Syst. Prot. control* 50 (6), 107–116. doi:10.19783/j.cnki.pspc.210783

Zhang, M., Zhang, X., Mo, J., Xiang, Z., and Zheng, P. (2022). Brake uneven wear of high-speed train intelligent monitoring using an ensemble model based on multi-sensor feature fusion and deep learning. *Eng. Fail. Anal.* 137, 106219. doi:10.1016/j.engfailanal.2022.106219

Zhou, Z., Liu, X., Wang, S., Zhu, C., Liu, H., and Song, C. (2018). 500kV XLPE submarine cable insulation and jacket transient voltage simulation calculation. *High. Volt. Technol.* 44 (8), 2725–2731. doi:10.13336/j.1003-6520.hve.20180731032

Zhu, Y., Tian, D., and Yan, F. (2020). Effectiveness of entropy weight method in decision-making. *Math. Problems Eng.* 2020, 1–5. doi:10.1155/2020/3564835

frontiers | Frontiers in Energy Research

TYPE Original Research
PUBLISHED 30 September 2022
DOI 10.3389/fenrg.2022.1020607

OPEN ACCESS

EDITED BY
Youbo Liu,
Sichuan University, China

REVIEWED BY
Yuehui Ji,
Tianjin University of Technology, China
Ke Peng,
Shandong University of Technology, China
Shenxi Zhang,
Shanghai Jiao Tong University, China

*CORRESPONDENCE
Youjun Deng,
deng_youjun@hotmail.com

SPECIALTY SECTION
This article was submitted to Smart Grids,
a section of the journal
Frontiers in Energy Research

RECEIVED 16 August 2022
ACCEPTED 02 September 2022
PUBLISHED 30 September 2022

CITATION
Xu Y, Deng Y, Guo X, Wang J and Zhang J (2022), Economic-environmental dispatch for the integrated energy system considering off-design conditions.
Front. Energy Res. 10:1020607.
doi: 10.3389/fenrg.2022.1020607

Economic-environmental dispatch for the integrated energy system considering off-design conditions

Yurui Xu[1], Youjun Deng[1]*, Xiangwei Guo[1], Jiarui Wang[2] and Jiajun Zhang[2]

[1]Key Laboratory of Smart Grid of Ministry of Education, Tianjin University, Tianjin, China, [2]State Grid Jilin Electric Power Research Institute, Changchun, China

The integrated energy system (IES) is recognized as a promising energy utilization approach enabled to both improve the energy efficiency and reduce pollutant emissions. Although the economic-environmental dispatch (EED) problem of the IES has been widely studied, the fact that the IES is operated under off-design conditions, having a significant influence on the efficiency of energy devices, is neglected usually, resulting in the scheduled operations, for the IES could not be accurate enough in many actual situations. This study investigates the EED problem of the IES under off-design conditions. Technically, by integrating an efficiency correction process into the traditional energy hub (EH) model, a dynamic energy hub (DEH) model is first formulated for adapting itself to variable energy conversion efficiencies. Then, a deep neural network (DNN)-based efficiency correction method is proposed to predict and correct the time-varying efficiency of energy devices based on three main off-design conditions including the load rate, air temperature, and pressure. A multi-objective EED model is finally formulated for the IES, with the framework of the DEH model, aiming at establishing a trade-off between operational cost and emitted pollutants. Case studies show that the proposed approach helps in enhancing the accuracy of IES dispatch while having a good performance in both the operational cost and pollutant emission reduction.

KEYWORDS

integrated energy system (IES), energy hub (EH), off-design conditions, economic-environmental dispatch (EED), multi-objective optimization, deep neural network (DNN)

1 Introduction

With the increasing problem of environmental pollution and energy scarcity, both energy transformation and upgrading have received extensive attention worldwide (Jin et al., 2021). As one of the important energy utilization ways, the integrated energy system (IES) has been regarded as the development direction of the future energy field, which promotes the synergy and complementarity among multi-energy carriers (e.g., electricity,

gas, and heating) (Lin et al., 2021), having the advantages of enhancing the energy efficiency and reducing dependence on the fossil fuels. Based on this, aiming to promote effective coordination among multi-energy carriers, the optimal dispatch of the IES has become a prevalent research hotspot in academia and industries (Martínez Ceseña and Mancarella, 2019).

The optimal dispatch for the multi-energy systems has been studied extensively, and some representative works have also been reported. Bartnik et al. (2022) investigated the economic performances of the combined heating and power (CHP) systems, and an economic-effective dispatch model was developed. A dispatch model was proposed for the IES to reduce energy cost (Yu et al., 2019). On the other hand, the environmental consideration is another important concern of the multi-energy system dispatch. To this, the economic-environmental dispatch (EED) problem of the multi-energy systems has been introduced and is usually formulated as a multi-objective optimization problem to pursue a trade-off between the economic and environmental benefits. For instance, an EED model was proposed by Dougier et al. (2021), Sarfi and Livani (2018), and Nandimandalam et al. (2022) for microgrids including certain energy conversion and storage units, aiming to minimize the operational cost and pollutant emissions. Guo et al. (2022) studied the optimal economic operation problem for the IES, and the environmental consideration was modeled into a constraint. Msigwa et al. (2022) provided an EED framework that functioned as accommodating to the high penetration access of wind power. A multi-objective optimization model was developed by Xu et al. (2018), aiming to obtain the minimum fossil fuels and pollutant emissions of the IES, considering stochastic wind and solar power. Although the EED problem for multi-energy systems has been studied extensively in the aforementioned works, they all ignore the impact of the off-design conditions on the efficiency of energy devices and assume the latter is constant. In fact, the off-design conditions, such as the load rate (Zong et al., 2022), air temperature (Akpan and Fuls, 2019), and pressure (Gong et al., 2021), have a significant impact on the device efficiency and make it variable. As a result, the scheduled operations based on this constant device efficiency are essentially deviated from the true ones, and therefore, the operational results could also be inaccurate (Mu et al., 2022).

In the literature, several methods have been reported to deal with the impact of variable off-design characteristics on the efficiency of energy devices. In general, the existing methods can be classified into the following two streams: 1) the mathematical model-based method and 2) the data-driven-based method. In the first stream, the polynomial fitting method was used by Zheng et al. (2018) to model the energy device efficiency. However, the presented model was a nonlinear programming problem, which was easy to fall into the suboptimal or even infeasible solution. To address this issue, the nonlinear formulation of the device efficiency was piecewise linearized *via* approximate treatment (Almassalkhi and Towle, 2016). The piecewise linear (PWL) approach was also employed to modify the efficiency of energy devices, and a mixed-integer linear programming model was formulated by Huang et al. (2020). However, following Shao et al. (2020), the mathematical model-based methods are hard in terms of ensuring the accuracy of the calculated solutions and the computational efficiency. To this end, in the second stream, the data-driven-based method is studied further. In particular, the machine learning (ML) method is one of the most widely used methods. Based on the collected historical off-design conditions' data, many representative ML approaches can be used to predict the variable device efficiencies, such as polynomial regression (Li and Yao, 2021), support vector machines (Liu et al., 2020), and deep neural network (DNN) (Ghimire et al., 2019). The DNN has remarkable performance in both the computational accuracy and speed for nonlinear parameter forecast and has been applied in many machines' intelligence fields, such as image recognition (Chen S. et al., 2022), parameter forecast (Browell and Fasiolo, 2021), and system control (Li et al., 2021).

Based on the aforementioned discussion, this article focuses on the optimal dispatch of the IES and is intended to study the EED problem for the IES, considering the off-design conditions. Compared to the existing works (Sarfi and Livani, 2018; Xu et al., 2018; Yu et al., 2019; Dougier et al., 2021; Bartnik et al., 2022; Guo et al., 2022; Msigwa et al., 2022; Nandimandalam et al., 2022), the main novelty of this study is to take into account the impact of the off-design conditions on the energy device efficiency. The primary contribution of this article is to formulate an energy hub model, which enabled to operate with the time-varying energy device efficiency, named as the dynamic energy hub (DEH) model. The key idea of such a formulation is to incorporate an efficiency correction process into the traditional energy hub (EH) model. The second contribution of this article is to propose a DNN-based efficiency correction method, in which three main off-design conditions (i.e., load rate, air temperature, and pressure) are considered and taken as the input set of the designed DNN well, and the corrected energy device efficiency is the output. Moreover, to validate the benefits of considering the off-design conditions on the IES's dispatch, a multi-objective EED model with multi-scale operation considerations including operational economic and environmental impacts is developed for the IES under the framework of the DEH model.

The rest of this article is structured as follows. The overview of EED for the IES considering off-design conditions is described in Section 2. The DEH model is established in Section 3, and the DNN-based device efficiency correction method is developed in

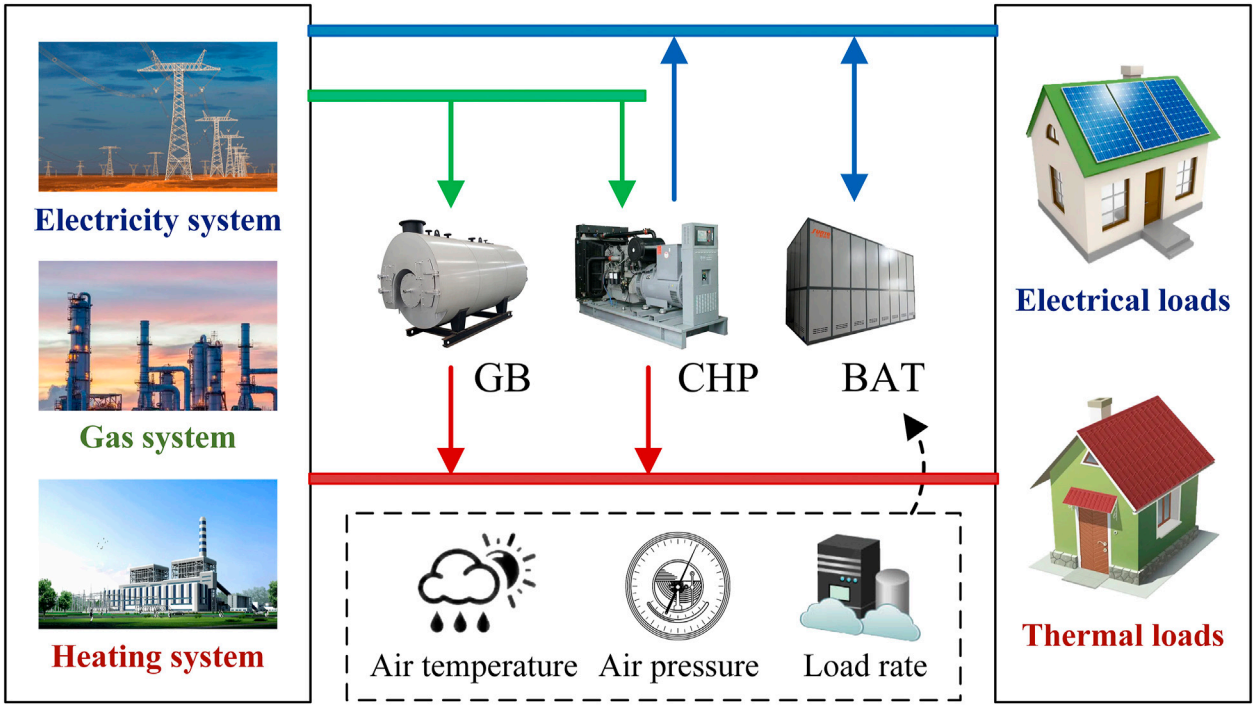


FIGURE 1
Typical structure of the IES.

Section 4. In Section 5, the multi-objective EED model is introduced. Case studies are applied to verify the proposed method in Section 6, and the conclusion is drawn in Section 7.

2 Overview of EED for the IES considering off-design conditions

Figure 1 illustrates a typical structure of an IES. Various energy devices (e.g., gas boiler (GB), CHP, and battery (BAT)) link the electricity system, gas system, and heating system together in the IES (Jin et al., 2016). Among them, the electrical loads are supplied by the electricity system, CHP, and BAT, and the CHP and GB supply the thermal loads. For the dispatch application, an IES is usually modeled into an energy hub framework based on the physical link topology of the IES.

The ideas of the proposed EED for the IES considering the off-design conditions are depicted in Figure 2. Taking into account the impact of off-design conditions on the efficiency of energy devices, the traditional EH model is remodeled as a DEH model that can operate with the time-varying device efficiency by integrating with an efficiency correction process. In the daily operation of the IES, the energy device efficiency is predicted and corrected in the efficiency correction process, based on the proposed DNN-based efficiency correction method, which is a data-driven method and will be introduced in detail in Section 3. Afterward, a multi-objective EED model is developed based on the obtained DEH model and is solved to realize a trade-off among multi-scale operational considerations over a finite operation horizon, that is, the operational cost and environmental impact reflected by the emitted pollutants.

3 Modeling of the dynamic energy hub

Based on this typical IES structure shown in Figure 1, the diagram of the traditional EH model can be presented in Figure 3A. With the consideration of variable off-design conditions, the energy device efficiency varies under their operation conditions, such as the load rate and other environmental-related factors (e.g., air temperature and pressure); thus, it is necessary to reformulate the traditional EH model to adapt it to the time-varying energy device efficiency. For this purpose, we plan to incorporate an

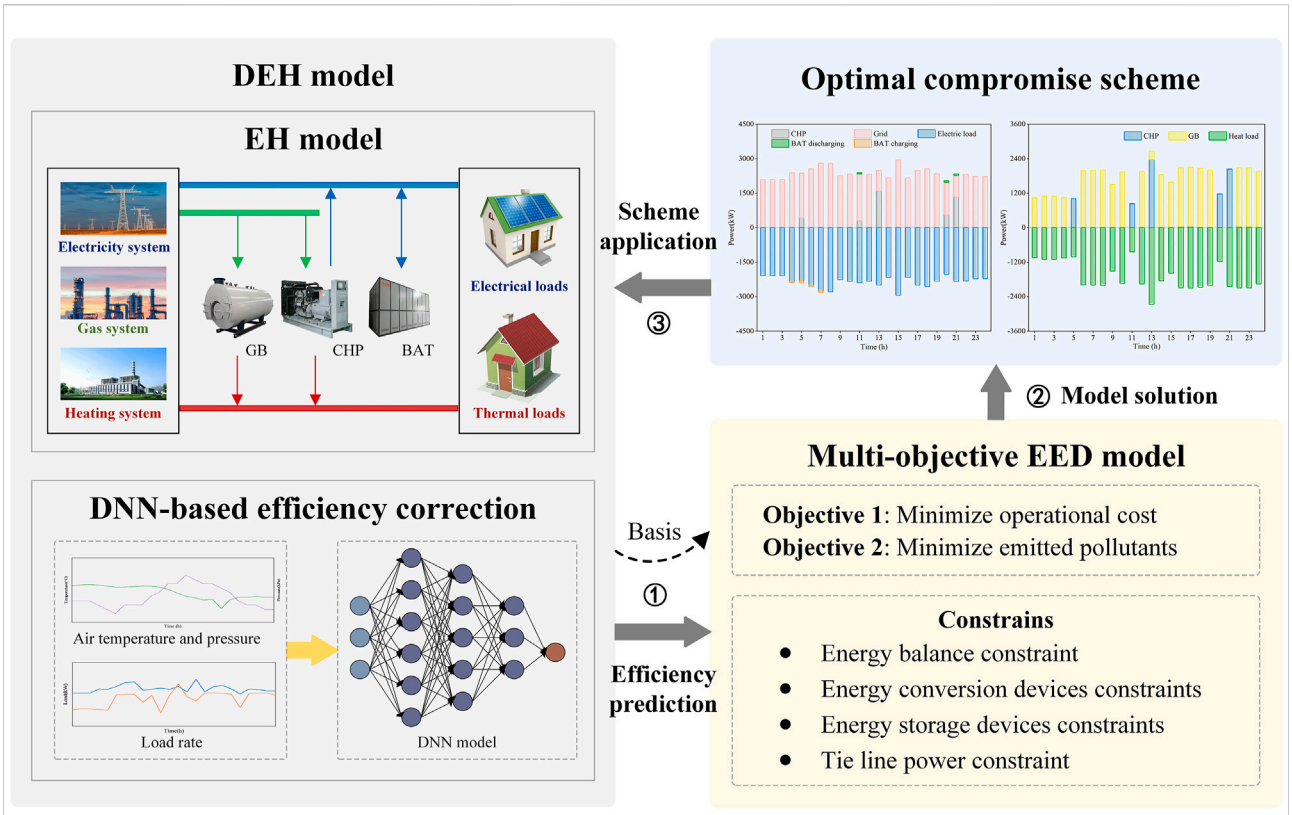


FIGURE 2
Framework of the proposed method.

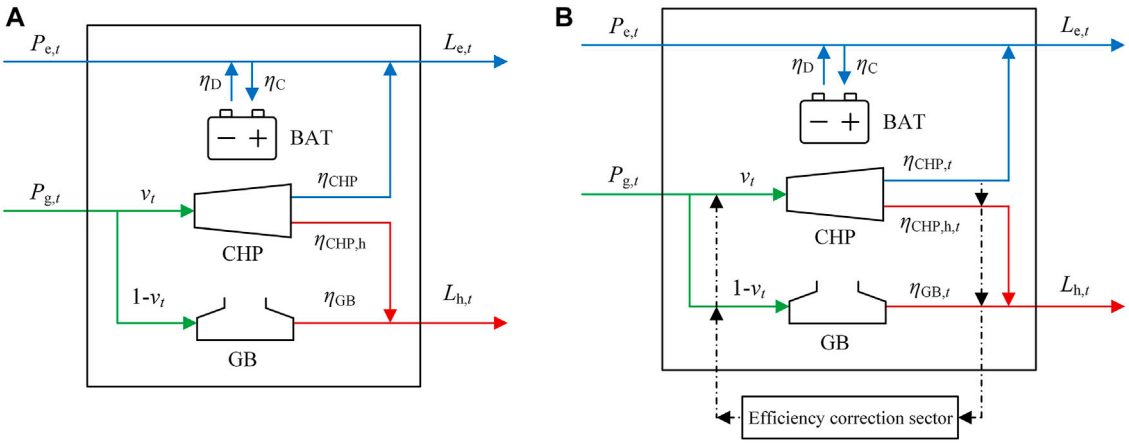


FIGURE 3
Typical topology model: **(A)** traditional EH model and **(B)** DEH model with the correction of the energy device efficiency.

efficiency correction process into the traditional EH model, as shown in Figure 3B, to develop the DEH model. With the added efficiency correction sector, the energy device efficiency can be adjusted with the variation of the off-design conditions. In this article, the efficiency correction process is functioned by the proposed DNN-based efficiency correction method.

The EH model is a linear multiple-input and multiple-output unit, where various energy sources are generated, converted, stored, and consumed, strengthening the integration of multiple energy carriers. The mathematical formulation for the traditional EH model, for a typical structure, as shown in Figure 3A, is presented in model (1).

$$\begin{bmatrix} L_{e,t} \\ L_{h,t} \end{bmatrix} = \begin{bmatrix} 1 & \nu_t \eta_{CHP} \\ 0 & \nu_t \eta_{CHP,h} + (1-\nu_t)\eta_{GB} \end{bmatrix} \begin{bmatrix} P_{e,t} \\ P_{g,t} \end{bmatrix} + \begin{bmatrix} \eta_C & \eta_D \\ 0 & 0 \end{bmatrix} \begin{bmatrix} W_{e,t} \\ 0 \end{bmatrix}. \tag{1}$$

In model (1), $P_{e,t}$ and $P_{g,t}$ denote the input electricity and gas powers, respectively, and $L_{e,t}$ and $L_{h,t}$ represent the electrical and thermal loads, respectively. η_{CHP}, $\eta_{CHP,h}$, and η_{GB} are the electrical efficiency of CHP, thermal efficiency of CHP, and the efficiency of GB, respectively. The dispatch factor is represented by $\nu_t \in [0,1]$, which defines the proportion of gas input to CHP and GB. $P_{BAT,t}$ is the charging/discharging power of BAT at time t.

As the IES is always operated under the off-design conditions, the device efficiency in the EH model is not constant but time-varying, which renders the EH model a multivariable nonlinear system (Chen M. et al., 2022). To address this challenge, the DEH model is modeled in this article, as shown in Figure 3B; the efficiency correction process is embedded in the EH model. $\eta_{CHP,t}$, $\eta_{CHP,h,t}$, and $\eta_{GB,t}$ are the electrical efficiency of CHP, thermal efficiency of CHP, and the efficiency of GB at time t, respectively. The coupling matrix of the DEH model is expressed as follows:

$$\begin{bmatrix} L_{e,t} \\ L_{h,t} \end{bmatrix} = \begin{bmatrix} 1 & \nu_t \eta_{CHP,t} \\ 0 & \nu_t \eta_{CHP,h,t} + (1-\nu_t)\eta_{GB,t} \end{bmatrix} \begin{bmatrix} P_{e,t} \\ P_{g,t} \end{bmatrix} + \begin{bmatrix} \eta_C & \eta_D \\ 0 & 0 \end{bmatrix} \begin{bmatrix} W_{e,t} \\ 0 \end{bmatrix}. \tag{2}$$

The efficiency correction process with execution on the basis of unit time interval dynamically calculates and corrects the energy device efficiency in the EH model based on three time-varying off-designs, including the actual load rate, air temperature, and pressure. Therefore, the formulated DEH model is capable of accurately simulating the operational environment of the IES under the variable off-design conditions.

4 DNN-based device efficiency correction

The off-design conditions have a significant impact on the efficiency of energy devices. Considering three main off-design conditions, namely, the load rate, air temperature, and pressure, the efficiency of energy devices can be modeled as a function of the aforementioned three elements as follows:

$$\eta_{i,t} = f(N_{i,t}, T_{i,t}, F_{i,t}), \tag{3}$$

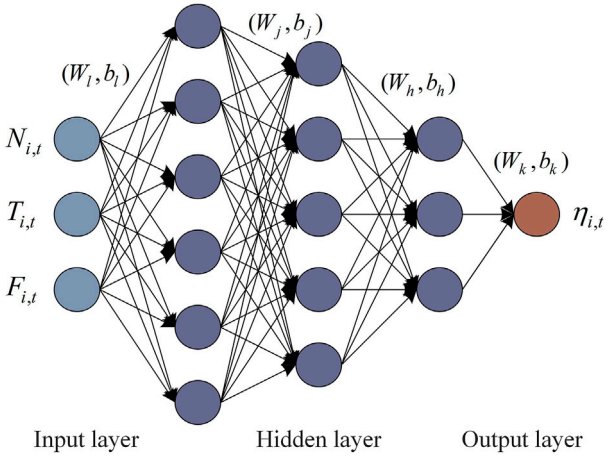


FIGURE 4
DNN model for device efficiency forecast.

where $\eta_{i,t}$ denotes the efficiency of the energy device i at time t. $N_{i,t}$ denotes the load rate of the energy device i at time t. $T_{i,t}$ and $F_{i,t}$ denote the air temperature and pressure at time t, respectively.

Due to the strong nonlinearity relationship between the energy device efficiency and the off-design conditions (Mu et al., 2022), the mathematical model-based method is difficult to achieve accurate predictions and is prone to overfitting and dimensionality curse. In this section, the DNN method (Ghimire et al., 2019) is employed to forecast the efficiency of energy devices.

A DNN framework is designed first in this article, which is a fully connected neural network with five layers, as presented in Figure 4. The designed DNN framework exhibits the best trade-off between the prediction speed and accuracy when it contains three hidden layers with six, five, and three neurons. The considered three off-design conditions are set as the input data set, and the device efficiency is the forecasted output. More specifically, the DNN model is formulated as follows:

$$x_{i,t} = [N_{i,t}, T_{i,t}, F_{i,t}], \tag{4}$$

$$h_l = g(W_l h_{l-1} + b_{l-1}), \forall l = 1, ..., 4, \tag{5}$$

$$y_{i,t} = \eta_{i,t}, \tag{6}$$

where h_l is the output vector of the l-th layer. $g(\ )$ is the activation function. W_l and b_{l-1} denote the connection weight matrixes. Model (4) represents the input sequence of the DNN. Model (5) describes the mathematical principles of forward propagation. Model (6) represents the output sequence of the DNN.

As shown in Figure 4, each layer is interconnected through weights (W_l) and biases (b_l). The activation function further introduces nonlinearity into the hidden layers. The sigmoid

function is selected as the activation function, given in Eq. 7, which can alleviate the gradient disappearance and improve the convergence rate (Ravnik and Hriberšek, 2019).

$$\text{sigmoid}(x) = \frac{1}{1+e^{-x}}. \tag{7}$$

After constructing the DNN model, the loss function is designed to guide the training. The mean square error between the obtained efficiency of energy devices and the actual value is defined as the loss function h_{loss}:

$$h_{\text{loss}} = \frac{1}{2}\sum\left(\hat{y}_{i,t} - y_{i,t}\right)^2. \tag{8}$$

The training process is performed by tuning the parameters of the DNN model to minimize the loss of the given training data, as given in model (9). Moreover, we apply the Levenberg–Marquardt algorithm (Long et al., 2020) to solve the problem in model (9), which is effective for the large-scale data set and makes the probability of falling into a local minimum much reduced.

$$\min_{W_l, b_l} \frac{1}{4}\sum_{l=1}^{4} f_{\text{loss},l}, \tag{9}$$

where $f_{\text{loss},k}$ is the loss of the lth layer in the training.

Moreover, the efficiency of energy storage devices primarily varies with long-term seasonal factors, and the effects on daily time-scale dispatch are ignorable (Mu et al., 2022). Thus, the variable off-design characteristics of energy storage devices are not taken into account in this article.

5 Multi-objective EED model

In this article, the IES is operated with multi-scale operational considerations, that is, the operational economic and environmental impact, under the off-design conditions. A multi-objective EED model is developed in this section for the IES, with two mutually exclusive objectives, that is, 1) to minimize the IES's operational costs over a finite dispatching horizon and 2) to minimize the pollutant emissions. The decision-making variables in the proposed multi-objective EED model include purchasing electricity $P_{\text{e},t}$, purchasing gas power $P_{\text{g},t}$, output power $P_{i,t}$ of the energy conversion device i, BAT charging power $P_{\text{C},t}$, and BAT discharging power $P_{\text{D},t}$.

5.1 Objective functions

5.1.1 Objective 1: Minimizing the operational cost

The first objective is referred to the operational economic consideration and is formulated to minimize the operational cost (f_{ope}) composed of the electricity purchase cost (C_{e}) and gas purchase cost (C_{g}) as follows:

$$f_{\text{ope}} = C_{\text{e}} + C_{\text{g}}, \tag{10}$$

$$C_{\text{e}} = \sum_{t}^{T} c_{\text{e},t} P_{\text{e},t} \Delta t, \tag{11}$$

$$C_{\text{g}} = \sum_{t}^{T} c_{\text{g}} P_{\text{g},t} \Delta t, \tag{12}$$

where T and Δt are the total dispatching period and unit dispatching period, respectively. $c_{\text{e},t}$ and c_{g} denote the unit prices of purchasing electricity and gas, respectively.

5.1.2 Objective 2: Minimizing the emitted pollutants

The second objective, that is, the environmental consideration, is to reduce the overall pollutant emissions (f_{emi}) caused by electricity usage (E_{e}) and gas usage (E_{g}) and is expressed as follows:

$$f_{\text{emi}} = E_{\text{e}} + E_{\text{g}}, \tag{13}$$

$$E_{\text{e}} = \sum_{t}^{T}\sum_{pg=1}^{N_{PG}} \chi_{\text{e},pg} P_{\text{e},t} \Delta t, \tag{14}$$

$$E_{\text{g}} = \sum_{t}^{T} \chi_{\text{g}} P_{\text{g},t} \Delta t, \tag{15}$$

where N_{PG} represents the types of pollutant emissions from electricity usage, including CO_2, CO, SO_2, and NO_x. $\chi_{\text{e},pg}$ represents the emission factor of the pg-type pollutant. χ_{g} denotes the emission factor of gas usage. Only CO_2 is taken into consideration in the pollutant emissions from gas usage due to the clean nature of gas (Lin et al., 2018).

5.2 Constraints

The proposed multi-objective EED model for the IES is subjected to the following constraints:

5.2.1 Energy balance constraint

The DEH model describes the coupling relationship between input and output power and adjusts the efficiency of energy devices dynamically by the DNN-based efficiency correction method. The energy balance constraint defined in the DEH model should be satisfied, as formulated in Eqs 2, 3.

5.2.2 Energy conversion devices' constraints

The energy conversion devices in this article include CHP and GB, and constraints are shown as follows.

$$P_{i,t} = \eta_{i,t} P_{\text{in},i,t}, \tag{16}$$

$$0 \le P_{i,t} \le P_{i,\text{cap}}, \tag{17}$$

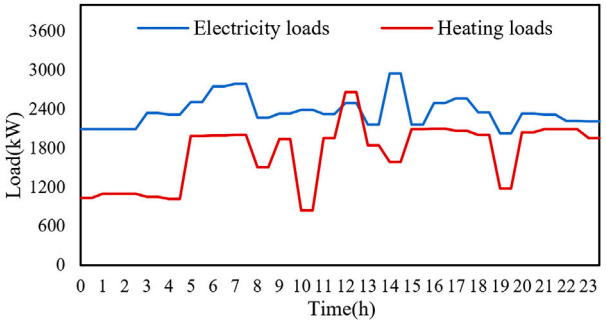


FIGURE 5
Electrical and thermal loads.

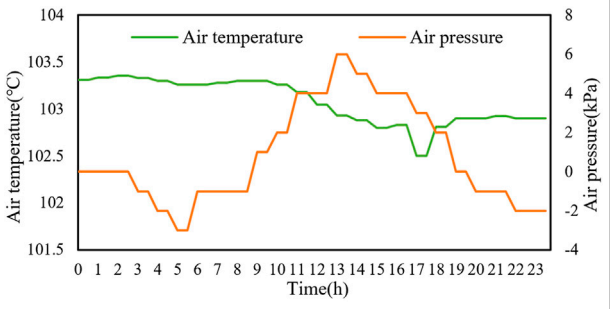


FIGURE 6
Air temperature and pressure.

where $P_{\text{in},i,t}$ represents the input power of the energy conversion device i at time t. $P_{i,\text{cap}}$ is the capacity of the energy conversion device i.

5.2.3 Energy storage devices' constraints

$$W_{\text{BAT},t+1} = W_{\text{BAT},t}(1-\sigma) + \left(P_{\text{C},t}\eta_{\text{C}} - P_{\text{D},t}/\eta_{\text{D}}\right)\Delta t, \quad (18)$$

$$0 \le W_{\text{BAT},t} \le W_{\text{BAT,cap}}, \quad (19)$$

$$0 \le P_{\text{C},t} \le P_{\text{C,max}}, \quad (20)$$

$$0 \le P_{\text{D},t} \le P_{\text{D,max}}, \quad (21)$$

$$W_{\text{start}} = W_{\text{end}}, \quad (22)$$

where $W_{\text{BAT},t}$ and $W_{\text{BAT},t+1}$ indicate the energy stored in BAT before and after charging and discharging, respectively. ηC and ηD are the charging and discharging efficiencies of BAT. σ and $W_{\text{BAT,cap}}$ are the self-discharge rate and capacity of BAT, respectively. $P_{\text{C,max}}$ and $P_{\text{D,max}}$ are the maximum charging and discharging power, respectively. W_{start} and W_{end} are the initial and final energy stored in BAT during 1 day (24 h), respectively.

5.2.4 Tie-line power constraint

The purchasing electrical power should not exceed the maximum tie-line power $P_{\text{grid,cap}}$:

$$\left|P_{\text{e},t}\right| \le P_{\text{grid,cap}}. \quad (23)$$

The non-dominated sorting genetical algorithm (NSGA-II) (Deb et al., 2002) is utilized in the MATLAB platform to solve the formulated multi-objective EED model in Section 5. The NSGA-II has strong performance in finding the Pareto frontier well of a multi-objective optimization problem and is one of the most recognized multi-objective evolutionary algorithms. With the obtained non-dominated solutions, the fuzzy decision-making criteria (Sakawa and Yano, 1985), with simplicity and capability in accounting for IES operators' preferences, are then employed to select and output the optimal compromise solution.

6 Case studies

6.1 System setup

The typical IES presented in Figure 1 is considered as an example. The electrical and thermal loads, air temperature, and pressure curves in a typical summer day are depicted in Figures 5, 6. The unit prices of purchasing electricity and gas are shown in Table 1. The emission factors of CO_2, CO, SO_2, and NO_x from electricity usage, the emission factors of CO_2 from gas usage, and the maximum tie-line power are shown in Table 2 (Lin et al., 2018). The technical parameters of devices are presented in Table 3 (Mu et al., 2022).

6.2 Simulation results and discussion

We set two benchmarks to comparatively validate the advantages of the proposed EED method as follows:

Case I. The EED method based on the constant efficiency EH model.

Case 2. The EED method based on the DEH model. In addition, different numbers of training iterations (*IT*) of DNNs are considered, and accordingly, multiple sub-cases are set. When *IT* is quite large enough (e.g., 500), the loss function values are approximately zero and DNNs converge completely, as shown in Figure 7. At this point, the dispatch results are considered accurate enough to serve as a reference for comparison with other cases.

The optimal compromise results in two cases are presented in Table 4. In case I, the operational cost is 7592.65$ and pollutant emissions are 12.632 t. The relative error reaches 3.726%. This demonstrates that the constant efficiency EH model cannot accurately simulate the off-design operating IES, resulting in a low precision of the dispatch scheme. In case II (*IT* = 200–500), the relative errors are significantly reduced, which indicates that the accuracy of the model is greatly improved. In addition, the relative errors in case II (*IT* = 200–500) are reduced as IT increases due to the increasing forecast precision of DNNs, as presented in Figure 7. Although the operating cost

TABLE 1 Unit prices of purchasing electricity and gas.

Types of energy	Time	Price ($/kWh)
Electricity	0:00–6:00 and 22:00–24:00	0.059
	6:00–9:00; 10:00–18:00; and 20:00–22:00	0.101
	9:00–10:00 and 18:00–20:00	0.142
Gas	0:00–24:00	0.065

TABLE 2 Emission factors.

Types of emission	Factor (t/MWh)
$\chi_{e,CO2}$	0.8647
$\chi_{e,CO}$	0.008
$\chi_{e,SO2}$	0.039
$\chi_{e,NOx}$	0.0309
χ_{g}	0.1940

TABLE 3 Rated parameters of devices.

Types of device	Parameter	Value
CHP	Rated capacity	1,600 kW
	Rated electrical efficiency	0.34
	Rated thermal efficiency	0.51
GB	Rated capacity	2050 kW
	Rated efficiency	0.94
BAT	Rated capacity	200 kWh
	Maximum charging/discharging power	60 kW
	Rated charging/discharging efficiency	0.9
	Self-discharge rate	0.0001

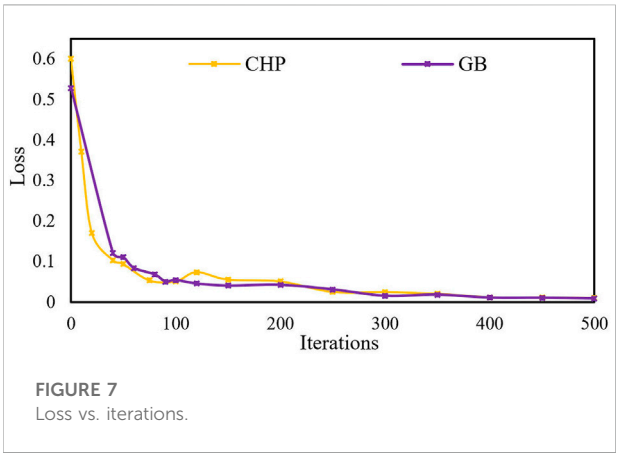


FIGURE 7
Loss vs. iterations.

and pollutant emissions in case II (IT = 500) are 293.84$ and 0.854 t, respectively, which are higher than those in case I, the dispatch scheme in case II (IT = 500) enjoys the highest precision and is most consistent under the actual off-design operating conditions of the IES.

The dispatch results of electrical power in case I and case II (IT = 500) are shown in Figure 8, which have great differences.

In case I, since devices are assumed to operate at the rated efficiency, CHP is more economical and is employed as the primary power supply. Electricity is purchased from the grid to satisfy the remainder of electrical loads. BAT is charged during the peak electricity prices and discharged during the low electricity prices.

In case II (IT = 500), electrical loads are mostly met by the grid. When the electricity prices are low (e.g., 0–6 h and 23 h), it is more economical to supply electrical loads from the grid. CHP supplies only at the peak electricity prices (10 h and 19–20 h). The charging and discharging states of BAT are mainly affected by electricity prices, which is similar to the state in case I.

The dispatch results of thermal power in case I and case II (IT = 500) are quite different, as shown in Figure 9.

In case I, although CHP operates at low load levels, it remains as the main heating device, and the variable off-design characteristics are ignored. GB supplies energy only during peak thermal loads to make up for the unmet thermal loads by CHP.

In case II (IT = 500), the thermal loads are mainly supplied by GB. To ensure the efficient operation of devices, GB operates alone when thermal loads are below the GB capacity, while most thermal loads are supplied by CHP when the thermal loads exceed the GB capacity (e.g., 4 h, 12 h, and 19–20 h). Moreover, thermal and electrical loads are simultaneously supplied by CHP during the peak electricity prices (10 h) to improve the operating economy of the IES.

By comparing the two cases, the conclusions can be drawn that the traditional constant efficiency EH model is difficult to accurately model the IES under off-design conditions, making the dispatch results unreasonable. The impact of the off-design conditions on the energy device's efficiency is supposed to be taken into consideration in the pursuit of the precise performance of the EED model.

7 Conclusion and future work

This article has proposed a multi-objective EED method for the IES, considering the off-design conditions. Technically, the

TABLE 4 Operational cost and pollutant emissions.

Case	Operational cost ($)	Pollutant emission (t)	Relative error
Case I	7592.65	12.632	3.7259%
Case II (*IT* = 200)	7825.49	13.448	0.7735%
Case II (*IT* = 300)	7867.36	13.473	0.0243%
Case II (*IT* = 400)	7879.85	13.481	0.0084%
Case II (*IT* = 500)	7886.49	13.486	-

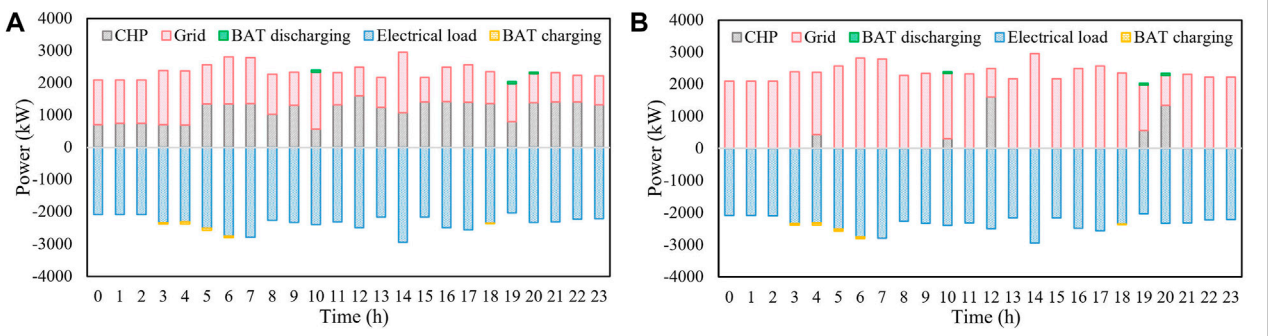


FIGURE 8
Dispatch results of electrical power: **(A)** case I and **(B)** case II (IT = 500).

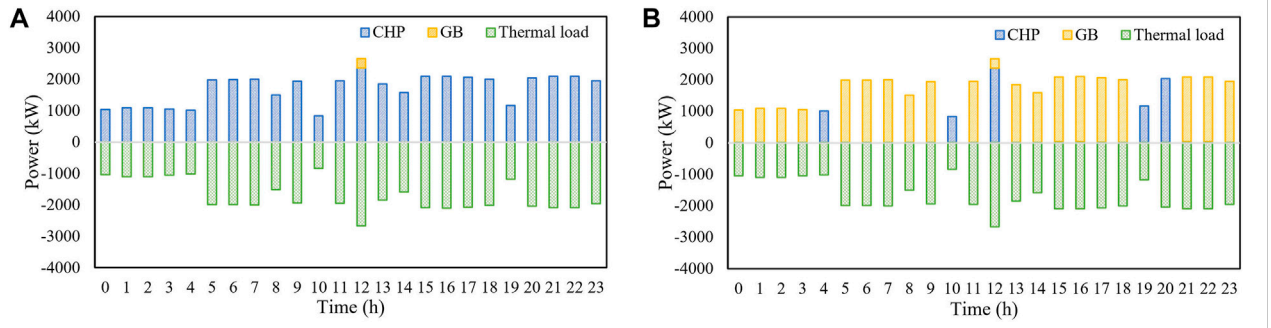


FIGURE 9
Dispatch results of thermal power: **(A)** case I and **(B)** case II (IT = 500).

traditional EH model is remodeled into the DEH model enabled to operate with the time-varying energy device efficiency, by integrating with an efficiency correction process. Here, this efficiency correction is functioned by the proposed DNN-based efficiency correction strategy, which is a data-driven approach and has good performance in terms of nonlinear data processing. A multi-objective EED model of the IES is proposed to evaluate the benefits of considering the off-design conditions on the IES dispatch. Case studies are conducted, and several key observations are obtained from the numerical simulation: 1) the load rate of devices exerts a significant impact on the device efficiency, while the device efficiency in turn affects its load rate; 2) compared to the traditional EH model, the proposed DEH model is more reasonable and appropriate for following the actual operational environment of the IES; thus, more accurate operation schedules can be obtained; and 3) the

proposed IES multi-objective EED scheme based on the DEH model can balance the IES's multi-scale operational considerations well after taking into account the off-design conditions.

Currently, the authors are studying the real-time operation of the IES with the consideration of the off-design conditions, in which a rolling horizon strategy is utilized for coping with the realization of variable off-design conditions.

Data availability statement

The original contributions presented in the study are included in the article; further inquiries can be directed to the corresponding author.

Author contributions

YX and YD conceptualized the study, performed the methodology, and validated the results. YX and YD prepared the original draft. GX, WJ, and ZJ reviewed and edited the draft. WJ and ZJ supervised the study.

Funding

This study was funded by the Science and Technology Project of State Grid Corporation (No. 5700-202229196A-1-1-ZN).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Akpan, P. U., and Fuls, W. F. (2019). Application and limits of a constant effectiveness model for predicting the pressure of steam condensers at off-design loads and cooling fluid temperatures. *Appl. Therm. Eng.* 158, 113779. doi:10.1016/j.applthermaleng.2019.113779

Almassalkhi, M. R., and Towle, A. (2016). Power systems computation conference. *PSCC*, 1–7. doi:10.1109/PSCC.2016.7540981enabling city-scale multi-energy optimal dispatch with energy hubs

Bartnik, R., Buryn, Z., Hnydiuk-Stefan, A., and Kowalczyk, T. (2022). Thermodynamic and economic comparative analyses of a hierarchic gas-gas combined heat and power (CHP) plant coupled with a compressor heat pump. *Energy* 244, 123116. doi:10.1016/j.apenergy.2016.07.080

Browell, J., and Fasiolo, M. (2021). Probabilistic forecasting of regional net-load with conditional extremes and gridded NWP. *IEEE Trans. Smart Grid* 12 (6), 5011–5019. doi:10.1109/TSG.2021.3107159

Chen, M., Zhou, D., Tayyebi, A., Prieto-Araujo, E., Dörfler, F., and Blaabjerg, F. (2022b). Generalized multivariable grid-forming control design for power converters. *IEEE Trans. Smart Grid* 13 (4), 2873–2885. doi:10.1109/TSG.2022.3161608

Chen, S., Chen, K., Zhu, X., Jin, X., and Du, Z. (2022a). Deep learning-based image recognition method for on-demand defrosting control to save energy in commercial energy systems. *Appl. Energy* 324, 119702. doi:10.1016/j.apenergy.2022.119702

Deb, K., Pratap, A., Agarwal, S., and Meyarivan, T. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Trans. Evol. Comput.* 6 (2), 182–197. doi:10.1109/4235.996017

Dougier, N., Garambois, P., Gomand, J., and Roucoules, L. (2021). Multi-objective non-weighted optimization to explore new efficient design of electrical microgrids. *Appl. Energy* 304, 117758. doi:10.1016/j.apenergy.2021.117758

Ghimire, S., Deo, R. C., Raj, N., and Mi, J. (2019). Deep solar radiation forecasting with convolutional neural network and long short-term memory network algorithms. *Appl. Energy* 253, 113541. doi:10.1016/j.apenergy.2019.113541

Gong, L., Zhang, Y., and Bai, Z. (2021). Geothermal-solar hybrid power with the double-pressure evaporation arrangement and the system off-design evaluation. *Energy Convers. Manag.* 244, 114501. doi:10.1016/j.enconman.2021.114501

Guo, S., Song, G., Li, M., Zhao, X., He, Y., Kurban, A., et al. (2022). Multi-objective Bi-level quantity regulation scheduling method for electric-thermal integrated energy system considering thermal and hydraulic transient characteristics. *Energy Convers. Manag.* 253, 115147. doi:10.1016/j.enconman.2021.115147

Huang, W., Zhang, N., Wang, Y., Capuder, T., Kuzle, I., and Kang, C. (2020). Matrix modeling of energy hub with variable energy efficiencies. *Int. J. Electr. Power Energy Syst.* 119, 105876. doi:10.1016/j.ijepes.2020.105876

Jin, X., Kgs, L., Wu, Q., Jia, H., and Hatziargyriou, N. (2021). Optimal integration of building heating loads in integrated heating/electricity community energy systems: A Bi-level mpc approach. *IEEE Trans. Sustain. Energy* 12 (3), 1741–1754. doi:10.1109/TSTE.2021.3064325

Jin, X., Mu, Y., Jia, H., Wu, J., Xu, X., and Yu, X. (2016). Optimal day-ahead scheduling of integrated urban energy systems. *Appl. Energy* 180, 1–13. doi:10.1016/j.apenergy.2016.07.071

Li, C., Zhang, Z., Li, J., Ma, Y., and Zou, J. (2021). Design of control strategy and effect evaluation for primary frequency regulation of wind storage system. *Front. Energy Res.* 9. doi:10.3389/fenrg.2021.739439

Li, X., and Yao, R. (2021). Modelling heating and cooling energy demand for building stock using a hybrid approach. *Energy Build.* 235, 110740. doi:10.1016/j.enbuild.2021.110740

Lin, W., Jin, X., Mu, Y., Jia, H., Xu, X., Yu, X., et al. (2018). A two-stage multi-objective scheduling method for integrated community energy system. *Appl. Energy* 216, 428–441. doi:10.1016/j.apenergy.2018.01.007

Lin, W., Jin, X., Jia, H., Mu, Y., Xu, T., Xu, X., et al. (2021). Decentralized optimal scheduling for integrated community energy system via consensus-based alternating direction method of multipliers. *Appl. Energy* 302, 117448. doi:10.1016/j.apenergy.2021.117448

Liu, X., Chen, X., Li, J., Zhou, X., and Chen, Y. (2020). Facies identification based on multikernel relevance vector machin. *IEEE Trans. Geosci. Remote Sens.* 58 (10), 7269–7282. doi:10.1109/TGRS.2020.2981687

Long, Q., Yu, H., Xie, F., Lu, N., and Lubkeman, D. (2020). Diesel generator model parameterization for microgrid simulation using hybrid box-constrained levenberg-marquardt algorithm. *IEEE Trans. Smart Grid* 12 (2), 943–952. doi:10.1109/TSG.2020.3026617

Martínez Ceseña, E. A., and Mancarella, P. (2019). Energy systems integration in Smart districts: Robust optimisation of multi-energy flows in integrated electricity, heat and gas networks. *IEEE Trans. Smart Grid* 10 (1), 1122–1131. doi:10.1109/TSG.2018.2828146

Msigwa, G., Ighalo, J. O., and Yap, P. (2022). Considerations on environmental, economic, and energy impacts of wind energy generation: Projections towards sustainability initiatives. *Sci. Total Environ.* 157755. doi:10.1016/j.scitotenv.2022.157755

Mu, Y., Xu, Y., Cao, Y., Chen, W., Jia, H., Yu, X., et al. (2022). A two-stage scheduling method for integrated community energy system based on A hybrid mechanism and data-driven model. *Appl. Energy* 323, 119683. doi:10.1016/j.apenergy.2022.119683

Nandimandalam, H., Aghalari, A., Gude, V. G., and Marufuzzaman, M. (2022). Multi-objective optimization model for regional renewable biomass supported electricity generation in rural regions. *Energy Convers. Manag.* 266, 115833. doi:10.1016/j.enconman.2022.115833

Ravnik, J., and Hriberšek, M. (2019). A method for natural gas forecasting and preliminary allocation based on unique standard natural gas consumption profiles. *Energy* 180, 149–162. doi:10.1016/j.energy.2019.05.084

Sakawa, M., and Yano, H. (1985). Interactive fuzzy decision-making for multiobjective nonlinear programming using reference membership intervals. *Int. J. Man-Mach. Study* 23 (4), 407–421. doi:10.1016/S0020-7373(85)80043-2

Sarfi, V., and Livani, H. (2018). An economic-reliability security-constrained optimal dispatch for microgrids. *IEEE Trans. Power Syst.* 33 (6), 6777–6786. doi:10.1109/TPWRS.2018.2835421

Shao, S., Zhang, H., You, S., Wang, Y., and Li, Y. (2020). Prediction of thermal potentials of multi-panel refrigerant heated radiator based on a mathematical model considering refrigerant flow distribution. *Energy Build.* 228, 110446. doi:10.1016/j.enbuild.2020.110446

Xu, J., Wang, F., Lv, C., Huang, Q., and Xie, H. (2018). Economic-environmental equilibrium based optimal scheduling strategy towards wind-solar-thermal power generation system under limited resources. *Appl. Energy* 231, 355–371. doi:10.1016/j.apenergy.2018.09.113

Yu, J., Li, Z., Guo, Y., and Sun, H. (2019). Decentralized chance-constrained economic dispatch for integrated transmission-district energy systems. *EEE Trans. Smart Grid* 10 (6), 6724–6734. doi:10.1109/TSG.2019.2910757

Zheng, J. H., Huang, J., Deng, S., Wu, Q. H., Li, M., Zhou, X., et al. (2018). Comparison and error analysis of off-design and design models of energy hubs. *CSEE J. Power Energy Syst.* 5 (4), 454–465. doi:10.17775/CSEEJPES.2018.00630

Zong, C., Ji, C., Cheng, J., Zhu, T., Guo, D., Li, C., et al. (2022). Toward off-design loads: Investigations on combustion and emissions characteristics of a micro gas turbine combustor by external combustion-air adjustments. *Energy* 253, 124194. doi:10.1016/j.energy.2022.124194

frontiers | Frontiers in Energy Research

TYPE Original Research
PUBLISHED 06 January 2023
DOI 10.3389/fenrg.2022.1030076

OPEN ACCESS

EDITED BY
Youbo Liu,
Sichuan University, China

REVIEWED BY
Jia Liu,
Hangzhou Dianzi University, China
Wei Yang,
Southwest Petroleum University, China
Lu Nan,
Sichuan University, China

*CORRESPONDENCE
Xiaodi Wang,
18280409667@163.com

SPECIALTY SECTION
This article was submitted to Smart Grids, a section of the journal Frontiers in Energy Research

RECEIVED 28 August 2022
ACCEPTED 20 September 2022
PUBLISHED 06 January 2023

CITATION
Wang X, Su Y, Xu W, Zhang S and Zhang Y (2023), A market decision-making model for load aggregators with flexible load.
Front. Energy Res. 10:1030076.
doi: 10.3389/fenrg.2022.1030076

A market decision-making model for load aggregators with flexible load

Xiaodi Wang*, Yunche Su, Weiting Xu, Shuai Zhang and Yongjie Zhang

State Grid Sichuan Economic Research Institute, Chengdu, China

The fast development of renewable energy has resulted in great challenges to the power system, which urgently needs more flexible resources to maintain a system supply/demand balance. This paper established a multi-stage electricity market framework in the presence of a load aggregator (LA) including a day-ahead energy/reserve market and a real-time balanced market. To actively participate in the day-ahead energy market and reserve market, a load profile perception model for LA is proposed to evaluate in detail the response performance of consumers. Meanwhile, a market-bidding model of LA and a market-clearing model of the system operator for the day-ahead market are also established. To actively join the real-time balance market, a market-bidding model of LA for the real-time balance market based on surplus flexible resources is established. The system operator further clears the real-time balance market and dispatches the collected flexible resources according to the system supply-demand state. A modified IEEE 30 bus system is tested and shows that the proposed market framework can effectively promote consumers to respond to system regulation requirements and lowers the system supply-demand imbalance risk.

KEYWORDS

demand Response, load aggregator, market strategy, electricity market, flexible load

Introduction

To build a low-carbon energy system, renewable energy units, such as wind turbines (WT) and photovoltaics (PV), have quickly developed. It is expected that the total installed capacity of PV and WT in China will reach more than 1.2 billion kW by 2030 (Yang et al., 2018). As the proportion of renewable energy in the power system exceeds a certain threshold, the system operation mode will go through great change. Additionally, the issue of insufficient flexible resources will become a key bottleneck for the future development of renewable energy.

Current flexible resources provided by traditional thermal units are far from enough to support the regulation demand of the power system. To protect system reliability and safety, a good solution is to encourage consumers to change their load patterns in response to system operation requests (Youbo et al., 2019). Demand response (DR) includes types of programs, which focus on modifying the load profiles of consumers to maintain system

generation and a consumption balance (Liang et al., 2021). One way is to fully exploit DR resources of the demand side, whose controllable devices can be directly dispatched by the system operator and can be a complementary solution to maintain a system supply-demand balance. Another way is to modify the electricity market framework. Reasonable market competition and incentive mechanisms can help coordinate distributed flexible resources on the demand side so that consumers together with the power system are able to all benefit from adjusting their power patterns.

The information and control technology (ICT) of the smart grid provides a foundation to support the integration of flexible demand resources. However, due to the stochastic response behaviors of distributed consumers, the system operator can hardly control large-scale distributed consumers (Du et al., 2021). At this stage, the interaction framework for the system operator and small consumers has not been fully formed. Regarding load scheduling and aggregation as a critical issue, middleman, such as the load aggregator (LA), retailers, and virtual power plants (VPP), can provide good way for small-scale consumers to participate in the electricity market (Youbo et al., 2018). Also, the liberalization of electricity markets helps incentive consumers to join the dispatching schedule of system operators. In terms of the market-bidding strategy, Vivekananthan et al. (2014) constructed a robust optimization strategy for LA to participate in bidding in the energy market based on controllable load resources. Bruninx et al. (2020) established multi-LA market-bidding and a scheduling model for frequency control service. Hu et al. (2017) proposed the regulation framework to LA to schedule the thermal (Fang et al., 2016) load such as air conditioning and water heaters. Li et al. (2018) provided a way for VPP to participate in the energy market and ancillary market by dispatching electric vehicles (EV) and WT. Pandžić et al. (2013) and Chen et al. (2018) established a VPP optimal bidding model that considers the uncertainties of multi-market prices including the long-term and day-ahead market. In terms of the load control method, Shao et al. (2013) found that retailers can determine the real-time electricity prices of customers to manage their demand portfolios. Liang et al. (2018) proposed a robust optimization algorithm for LA to find the optimal real-time electricity prices offered to consumers considering the uncertainties of market prices. Baharlouei et al. (2013) and the California Independent System Operator (2016) studied the integration of incentive DR programs for LA considering power flow constraints. Hu et al. (2017) and Sumaiti et al. (2020) provided a game theory-based structure for retailers and consumers to decide on incentive prices. Despite these studies in the field of DR, several research gaps still need to be filled. First, bidding strategies and load control methods of a middleman are highly dependent on the accurate evaluation and integration of response capacity of distributed consumers. Most studies ignore the process of perceiving the real-time response capacity of consumers. Second, in the studies mentioned, the middleman focuses on engaging in a single market. This is an urgent issue to reasonably arrange consumers' schedules and efficiently participate in a multi-stage market to meet the profit demands of LA and consumers.

With this in mind, this paper proposes a multi-stage electricity market framework for a real-time balanced market, day-ahead energy and a reserve market. By constructing a load state perception model, the real-time flexible resources of consumers can be perceived. In the day-ahead energy and reserve market, a bidding model for LA based on the consumers' flexible resources is proposed. The system operator is in charge of market clearing according to the bidding strategy of market players, and thus, the dispatching scheme of day-ahead units can be decided. In the real-time balancing market, day-ahead clearing results are considered. All LAs bid in the real-time balancing market based on the surplus of flexible resources of their signed consumers. Then, system operators clear real-time balancing resources and dispatch the spare flexible resources according to the actual system balance demand. A modified IEEE 30 bus system was tested to verify the proposed framework. The results show that the proposed model can stimulate consumers to effectively respond to system regulation demand to promote the consumption of renewable energy.

Market framework

Referring to the market framework of the Nordic electricity market, the market setting of this paper can be described as follows:

1) Market framework and market participants: The multi-stage market here includes the day-ahead energy market, day-ahead ancillary market and real-time balance market. The market operation framework is shown in Figure 1. Market participants include load aggregators, thermal units and the system operator. LA and thermal units take part in a centralized bidding transaction. The system operator clears the market and conducts a day-ahead schedule of units after the bidding of market participants.
2) Bidding in day-ahead energy and the day-ahead reserve market: LA will bid in the day-ahead energy market (24 points a day) and day-ahead reserve market (including upper and lower reserve capacity) based on predicted load profiles of signed consumers. These two markets shall be independently organized and jointly settled according to the uniform market-clearing price.
3) Bidding in the real-time balance market: Generally, not many balancing service varieties exist in Northern Europe, and part of balance regulation resources can be obtained through the real-time balancing market. The real-time balancing market

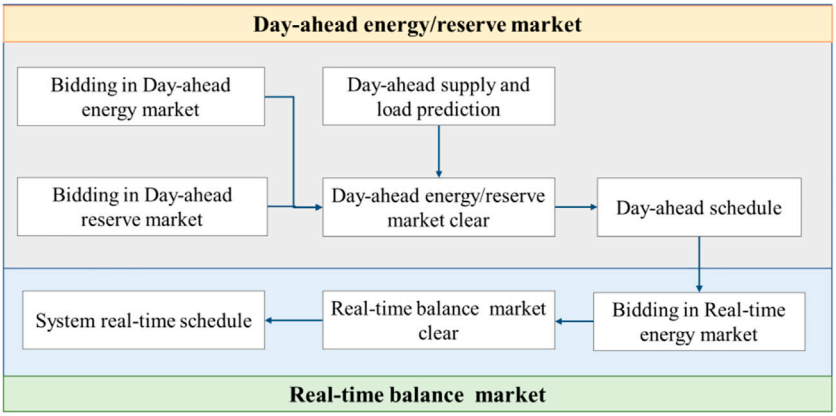


FIGURE 1
The trading framework of muti-stage markets.

will be carried out 1 h in advance. Market participants will bid to provide the up or down-regulation service in the real-time balancing market based on day-ahead market-clearing results. Bidding information includes aspects such as a market participation period, up-regulation capacity, and down-regulation capacity.

4) Market settling method: We adopt the market structure "main energy market + ancillary market + balanced market" multi-stage market operation process and different service varieties settle separately.
5) Load regulation and deviation assessment: LA optimizes the load profiles of signed distributed consumers according to multi-stage market-clearing results to avoid a market deviation check.

LA bid strategy

Load profiles perception model

1) Optimal control strategy for LA to regulate signed consumers

Before bidding in the markets, LAs need to accurately evaluate the load profiles and flexible resources of signed consumers whose controllable devices include washing machines or EVs. By regulating the controllable devices of its signed consumers, LA aims to minimize the total energy cost of consumers as long as the aggregation resources can satisfy LA's regulation demand. For scenario $\omega \in \{1, \cdots, \Omega\}$ time $t \in \{1, \cdots, T\}$, the objective of LA $n \in N$ can be described as follows:

$$\min_{u_{i,h,t,\omega}} \sum_{t}^{T} \lambda^{S}_{n,t,w} \left(P^{D}_{n,h,t,\omega} + \sum_{i}^{I} P_{n,i,h,w} u_{n,i,h,t,\omega} \right) \forall t, \omega, n \qquad (1)$$

$$\sum_{t=1}^{T} u_{n,i,h,t,\omega} \geq t^{R}_{n,i,h,\omega} \; \forall i, h, t, \omega, n \qquad (2)$$

$$\sum_{t=1}^{T} \left| u_{n,i,h,t,\omega} - u_{n,i,h,t-1,\omega} \right| - 1 = K_i \; \forall i, h, t, \omega, n \qquad (3)$$

$$u_{n,i,h,t,\omega} \leq t^{*}_{n,i,h,t,\omega} \forall i, h, t, \omega, n \qquad (4)$$

$$t^{*}_{n,i,h,t,\omega} = \begin{cases} 1, t \in \left[t^{s}_{n,i,h,\omega}, t^{e}_{n,i,h,\omega} \right] \\ 0, \quad \text{else} \end{cases} \forall i, h, t, \omega, n \qquad (5)$$

where $N, H,$ and I is the set of LAs, signed consumers, and controllable devices. $P_{n,i,j,h,w}$ is the original load of the controllable device $i \in I$ of the signed consumer $h \in \{1, \cdots, H\}$ of LAn. $P^{D}_{n,h,t,\omega}$ is the fixed load of the signed consumer $h \in \{1, \cdots, H\}$ of LAn. $j \in J$ is the energy consumption duration of the device and $\lambda^{S}_{n,t,\omega}$ is the offered retail price of LAn. $u_{n,i,h,t,\omega}$ is the 0–1 variable, which denotes the dispatch schedule for the device $i \in I$. $u_{n,i,h,t,\omega} = 1$ means the device $i \in I$ is in a working state. $t^{R}_{n,i,h,\omega}$ is the minimum power usage duration. K_i refers to the total time of starts and stops that can be controlled during the scheduling period. $t^{s}_{n,i,h,\omega}, t^{e}_{n,i,h,\omega}$ is the earliest and latest device usage time specified by consumer h for device I, respectively. $t^{*}_{n,i,h,\omega}$ means that the controllable device can participate in regulation during the energy consumption period specified by the consumer h. Constraint 2 ensures that the basic needs of each consumer for washing machines or EV charging can still be satisfied within the specific period; Constraint 3 ensures that the start and stop times of the device are in the restricted range. Constraints 4, 5 ensure that the LA's schedule plan satisfies the consumers' basic demands.

2) The initial load profile prediction of LA signed consumers

The initial load profiles of LA signed consumers will be evaluated and predicted, which can be expressed as follows:

$$P^{F}_{n,t,\omega} = \sum_{h}^{H}\left(P^{D}_{n,h,t,\omega} + \sum_{i}^{I} P_{n,i,h,w} u^{R}_{n,i,h,t,\omega}\right) \forall t, \omega, n \tag{6}$$

$$u^{R}_{n,i,h,t,\omega} = \begin{cases} 1, t = t^{R}_{n,i,h,\omega} \\ 0, \quad else \end{cases} \forall i, h, t, \omega, n \tag{7}$$

where $u^{R}_{n,i,h,t,\omega}$ is the 0–1 variable. $u^{R}_{n,i,h,t,\omega} = 1$ represents the initial state for device i at time plot t in an open state. $t^{R}_{n,i,h,\omega}$ is the initial open time for device i.

3) Flexible DR resource evaluation

In the typical scenario ω, assuming the response ratio of signed consumers of LAn is $d_{n,w}$. Thus, the number of responsive consumers is $d_{n,w}H$ and the number of unresponsive consumers is $d_{n,w}(1-H)$. Thus, the aggregation load of LAn can be expressed as:

$$\begin{aligned} P^{T}_{n,t,\omega} &= P^{NN}_{n,t,\omega} + P^{FN}_{n,t,\omega} \\ &= \sum_{h \in H^{N}_{n,w}}\left(P^{D}_{n,h,t,\omega} + \sum_{i}^{I} P_{n,i,h,w} u_{n,i,h,t,\omega}\right) + \sum_{h \in H^{C}_{n,w}}\left(P^{D}_{n,h,t,\omega} + \sum_{i}^{I} P_{n,i,h,w} u^{R}_{n,i,h,t,\omega}\right) \forall t, \omega, n \end{aligned} \tag{8}$$

where $P^{T}_{n,t,\omega}$ is the total load of signed consumers of LAn. $H^{N}_{n,w}$ is the unresponsive consumers set of signed consumers. $P^{NN}_{n,t,\omega}$ is the unresponsive load of signed consumers. $H^{C}_{n,w}$ is the responsive consumers set of signed consumers. $P^{FN}_{n,t,\omega}$ is the responsive load of signed consumers.

Bidding model of LA in the day-ahead energy market and reserve market

After evaluating the load profiles of signed consumers, LAn aggregate the load resources of signed consumers to join the day-ahead energy market, the day-head reserve market and the real-time balance market to gain profits. The bidding strategy of LAn in the day-ahead market can be expressed as follows:

$$\begin{aligned} \max_{P^{bid}_{n,t,c}, r^{up}_{n,t,w}, r^{down}_{n,t,w}} & \sum_{\omega=1}^{\Omega} \pi_{\omega} \sum_{t}^{T} \left(\lambda^{S}_{n,t,w} P^{T}_{n,t,\omega} - \lambda^{A}_{n,t,w} P^{bid}_{n,t,w} + \lambda^{LAC}_{n,t,w}\left(r^{up}_{n,t,w} + r^{down}_{n,t,w}\right)\right. \\ & \left. - \lambda^{P}_{n,t,w}\left(P^{Im^{+}}_{n,t,w} + P^{Im^{-}}_{n,t,w}\right)\right) \end{aligned} \tag{9}$$

$$P^{Im}_{n,t,w} = P^{T}_{n,t,\omega} - P^{bid}_{n,t,w} \forall t, \omega, n \tag{10}$$

$$P^{Im}_{n,t,w} = P^{Im^{+}}_{n,t,w} - P^{Im^{-}}_{n,t,w} \forall t, \omega, n \tag{11}$$

$$0 \le r^{up}_{n,t,w} \le r^{up,max}_{n,t,w} \forall t, \omega, n \tag{12}$$

$$0 \le r^{down}_{n,t,w} \le r^{down,max}_{n,t,w} \forall t, \omega, n \tag{13}$$

$$P^{NN}_{n,t,\omega} + \sum_{h \in H^{C}_{n,w}} P^{D}_{n,h,t,\omega} \le P^{bid}_{n,t,w} + r^{up}_{n,t,w} \le P^{T}_{n,t,\omega} \forall t, \omega, n \tag{14}$$

$$0 \le P^{NN}_{n,t,\omega} + \sum_{h \in H^{C}_{n,w}} P^{D}_{n,h,t,\omega} \le P^{bid}_{n,t,w} - r^{down}_{n,t,w} \forall t, \omega, n \tag{15}$$

where π_{ω} is the probability of scenario ω. $\lambda^{A}_{n,t,w}, \lambda^{LAC}_{n,t,w}, \lambda^{P}_{n,t,w}$ are the predicted prices of the day-ahead energy market, the day-ahead reserve market, and the real-time unbalance penalty. $P^{T}_{n,t,\omega}, P^{bid}_{n,t,w}, P^{Im}_{n,t,w}$ are the day-ahead forecast load demand, LA bid quantity in the day-ahead energy market, and LA bid quantity in the real-time balance market. $r^{up}_{n,t,w}, r^{down}_{n,t,w}$ are the up and down reserve capacity that LA bid in the reserve market. $r^{up,max}_{n,t,w}, r^{down,max}_{n,t,w}$ is the maximum up and down reserve capacity. $P^{Im^{+}}_{n,t,w}, P^{Im^{-}}_{n,t,w} > 0$; negative $P^{Im}_{n,t,w}$ indicates that LA buys too much power in the day-ahead energy market, and abundant power can be used to bid in the real-time balance market for down-regulation. Positive $P^{Im}_{n,t,w}$ indicates LAn needs to bid for up-regulation in the real-time balance market. $\lambda^{P}_{n,t,w}(P^{Im^{+}}_{n,t,w} + P^{Im^{-}}_{n,t,w})$ is a penalty coefficient to ensure that the bidding quantity of LAn in the day-ahead market can meet the majority of its load demand. Constraints 14 and 15 ensure that LA bidding quantity can satisfy the base load demand and LA bidding quantity is within the controllable range.

The bidding curve of LA in the day-ahead energy market and reserve market can be expressed as a piecewise decreasing function $C^{LA}_{n,t,w}(P^{bid}_{n,t,w})$ related to the demand interval and the energy price as shown below:

$$C^{LA}_{n,t,w}\left(P^{bid}_{t,w}\right) = P^{bid}_{n,t,c+1} + \left(P^{bid}_{n,t,c} - P^{bid}_{n,t,c+1}\right)\frac{\lambda^{bid}_{n,t,c+1} - \lambda^{A}_{n,t,w}}{\lambda^{bid}_{n,t,c+1} - \lambda^{bid}_{n,t,c}} \forall t, \omega, n \tag{16}$$

where the bidding curve includes the C segment $(P^{bid}_{n,t,1}, \lambda^{bid}_{n,t,1}), (P^{bid}_{n,t,2}, \lambda^{bid}_{n,t,2}), \cdots, (P^{bid}_{n,t,C}, \lambda^{bid}_{n,t,C})$. We have $\lambda^{bid}_{n,t,1} \le \lambda^{bid}_{n,t,2} \le \cdots \le \lambda^{bid}_{n,t,C}$, $P^{bid}_{n,t,1} \ge P^{bid}_{n,t,2} \ge \cdots \ge P^{bid}_{n,t,C}$. For scenario ω, the final market clear price $\lambda^{A}_{n,t,w}$ will be settled between $(\lambda^{bid}_{n,t,c}, \lambda^{bid}_{n,t,c+1})$ and shown in Figure 2.

After evaluating the load profile of the signed consumers, LAn aggregates the load resources of signed consumers to join the day-ahead energy market, the day-head reserve market and the real-time balance market to gain profits. The bidding strategy of LAn in the day-ahead market can be expressed as follows:

After the day-head energy market and reserve market are cleared, LA can participate in the real-time balance market to

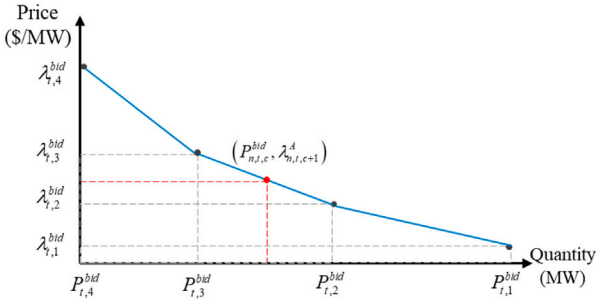


FIGURE 2
The bidding curve of the load aggregator for the day-ahead energy market. The bidding model of LA in the real-time balance market.

provide an up or down power regulation service based on its abundant flexible resources of signed consumers. LA's objective function can be expressed as follows:

$$\max_{\substack{P^{Im^+}_{bw},P^{Im^-}_{t,w}\\P^{Tr}_{n,t,\omega},u_{n,i,h,t,\omega}}} \sum_{\omega=1}^{\Omega}\pi_\omega\sum_{t}^{T}\left(\lambda^{S}_{n,t,w}\left(P^{Im^+}_{n,t,w}-P^{Im^-}_{n,t,w}\right)+\lambda^{B}_{t,w}\left(P^{Im^+}_{n,t,w}+P^{Im^-}_{n,t,w}\right)-\lambda^{div}_{n,t,w}\left|P^{div}_{n,t,w}\right|\right) \tag{17}$$

$$P^{div}_{n,t,w}=P^{bid^*}_{n,t,w}+P^{Im^+}_{n,t,w}-P^{Im^-}_{n,t,w}+r^{up^*}_{n,t,w}-r^{down^*}_{n,t,w}-P^{T}_{n,t,\omega}\ \forall t,\omega,n \tag{18}$$

$$P^{Im^+}_{n,t,w}\le\max\left\{\left(P^{T}_{n,t,\omega}-P^{bid^*}_{n,t,w}-r^{up^*}_{n,t,w}\right),0\right\}\forall t,\omega,n \tag{19}$$

$$P^{Im^-}_{n,t,w}\le\max\left\{\left(P^{bid^*}_{n,t,w}-r^{down^*}_{n,t,w}-P^{NN}_{n,t,\omega}-\sum_{h\in H^{C}_{n,w}}P^{D}_{n,h,t,\omega}\right),0\right\}\forall t,\omega,n \tag{20}$$

where $\lambda^{div}_{n,t,w}$ is the penalty price for energy deviation of LAn. $P^{bid^*}_{n,t,w}$ is the energy obtained by bidding in the day-ahead energy market. $\lambda^{B}_{t,w}$ is the up/down-regulation price in the real-time balance market. Constraints 19, 20 ensure that the up/down-regulation power is within the controllable range.

Electricity market model

Day-ahead energy market and reserve market model

After each market participant completes the bidding in the day-ahead market, the system operator will jointly clear the day-ahead energy market and the day-ahead reserve market, which aims at minimizing the system operation cost (including the total cost of generation and the cost of the system ancillary service). Thus, the day-ahead scheduling plan for thermal units can be obtained as follows:

$$\min_{\substack{P^{G}_{g,t,w},r^{up}_{g,t,w},r^{down}_{g,t,w},\\P^{bid}_{n,t,w},r^{up}_{n,t,w},r^{down}_{n,t,w}}} \sum_{\omega=1}^{\Omega}\pi_\omega\sum_{t}^{T}\left(\sum_{g=1}^{G}\left(\begin{array}{l}C^{GS}_{g,t,w}\left(P^{G}_{g,t,w}\right)\\+\lambda^{GC}_{g,t,w}\left(r^{up}_{g,t,w}+r^{down}_{g,t,w}\right)\end{array}\right)-\sum_{n=1}^{N}\left(\begin{array}{l}C^{LA}_{n,t,w}\left(P^{bid}_{n,t,w}\right)\\-\lambda^{LAC}_{n,t,w}\left(r^{up}_{n,t,w}+r^{down}_{n,t,w}\right)\end{array}\right)\right)\forall t,\omega \tag{21}$$

where we have G thermal units to join the day-ahead market. Like the LA bidding curve $C^{LA}_{n,t,w}(P^{bid}_{n,t,w})$, the bidding curve of thermal units $C^{GS}_{g,t,w}(P^{G}_{g,t,w})$ is a multi-segment linear increasing function related to the bidding price and output. $P^{G}_{g,t,w}$ is a bidding quantity of thermal unit $g$in the day-ahead energy market and the bidding up/down reserve capacity in day-ahead reserve market. $r^{up}_{g,t,w}, r^{down}_{g,t,w}\lambda^{GC}_{g,t,w}$ is the bidding reserve price. The optimization model (Eq. 21) will subject to the following constraints.

1) System balance constraint:

$$P^{G}_{b,t,w}+P^{re}_{b,t,w}-\sum_{l|b\in o(l)}f_{l,t,w}+\sum_{l|b\in r(l)}f_{l,t,w}=P^{bid}_{b,t,w}+P^{D0}_{b,t,w}\ \forall t,\omega,b \tag{22}$$

where $P^{D0}_{t,w}$ is the load demand of none-market consumers, i.e., unsigned consumers by LA. $P^{bid}_{b,t,w}$ is the day-ahead bidding quantity of LA at bus b. $P^{G}_{b,t,w}$ is the day-ahead bidding quantity of thermal units at bus b. $P^{re}_{b,t,w}$ is the predicted renewable energy output. $f_{l,t,w}$ is the power flow of the transmission line $l\in L$. $b\in o(l)$ is the line set where bus b is used to send power. $b\in r(l)$ is the line set where bus b is used to receive power.

2) Thermal unit constraint:

$$G_g{}^{min}+r^{down}_{g,t,w}\le P^{G}_{g,t,w}\le G_g{}^{max}-r^{up}_{g,t,w}\ \forall t,\omega,g \tag{23}$$

$$0\le r^{up}_{g,t,w}\le R^{RU}_{g,w}\ \forall t,\omega,g \tag{24}$$

$$0\le r^{down}_{g,t,w}\le R^{RD}_{g,w}\ \forall t,\omega,g \tag{25}$$

where $G_g{}^{max}, G_g{}^{min}$ is the maximum/minimum output of thermal unit g. $R^{RU}_{g,w}, R^{RD}_{g,w}$ is the maximum/minimum reserve capacity of thermal unit g.

3) Reserve capacity constraint:

$$\sum_{g\in G}r^{up}_{g,t,w}+\sum_{n\in N}r^{down}_{n,t,w}\ge\phi^{G}\sum_{b\in B}P^{re}_{b,t,w}+\phi^{D}\sum_{b\in B}P^{D0}_{b,t,w}\ \forall t,\omega,b \tag{26}$$

$$\sum_{g\in G}r^{down}_{g,t,w}+\sum_{n\in N}r^{up}_{n,t,w}\ge\phi^{G}\sum_{b\in B}P^{re}_{b,t,w}+\phi^{D}\sum_{b\in B}P^{D0}_{b,t,w}\ \forall t,\omega,b \tag{27}$$

where Constraints 26, 27 ensure that the system up/down reserve capacity can satisfy the basic reserve demand. ϕ^{G},ϕ^{D} denote the reserve ratio of renewable energy and load, respectively.

4) System operation constraint:

The system operation constraints include the voltage constraint (29) and line capacity constraint (30) as follows:

$$f_{l,t,w}=B_l\left(\theta_{o(l),t,w}-\theta_{r(l),t,w}\right)\forall t,\omega,l \tag{28}$$

$$\theta^{min}\le\theta_{b,t,w}\le\theta^{max}\ \forall t,\omega,l \tag{29}$$

$$F_l{}^{max}\le f_{l,t,w}\le F_l{}^{max}\ \forall t,\omega,l \tag{30}$$

where $F_l{}^{max}$ is the maximum capacity of line l. $\theta_{j,t,b}$ is the bus voltage. $\theta^{min},\theta^{max}$ is the minimum/maximum bus voltage.

Real-time balance market model

In the real-time balance market and to eliminate the net deviation caused by renewable energy, the system operator dispatches the up/down flexible resources provided by the

day-ahead reserve market and real-time balance market. Thus, the real-time balance market can be modeled as follows:

$$\min_{\substack{r_{g,t,w}^{up,rt}, r_{g,t,w}^{down,rt}, r_{n,t,w}^{up,rt} \\ r_{n,t,w}^{down,rt}, d_{t,w}^{load}, d_{t,w}^{re} \\ P_{n,t,w}^{Im^+}, P_{n,t,w}^{Im^-}}} \sum_{\omega=1}^{\Omega} \pi_\omega \left(\sum_{g=1}^{G} \begin{pmatrix} \lambda_{g,t,w}^{GC} \left(r_{g,t,w}^{up,rt} + r_{g,t,w}^{down,rt} \right) \\ +\delta_{t,w} \left(d_{t,w}^{load} + d_{t,w}^{re} \right) \end{pmatrix} \right.$$

$$\left. + \sum_{n=1}^{N} \begin{pmatrix} \lambda_{n,t,w}^{LAC} \left(r_{n,t,w}^{up,rt} + r_{n,t,w}^{down,rt} \right) \\ +\lambda_{t,w}^{B} \left(P_{n,t,w}^{Im^+,rt} + P_{n,t,w}^{Im^-,rt} \right) \end{pmatrix} \right) \forall t, \omega \quad (31)$$

$$\sum_{g=1}^{G} r_{g,t,w}^{down,rt} + \sum_{n=1}^{N} r_{n,t,w}^{up,rt} + \sum_{n=1}^{N} P_{n,t,w}^{Im^+} + d_{t,w}^{re} = L_{t,w}^{De^+} \forall t, \omega \quad (32)$$

$$\sum_{g=1}^{G} r_{g,t,w}^{up,rt} + \sum_{n=1}^{N} r_{n,t,w}^{down,rt} + \sum_{n=1}^{N} P_{n,t,w}^{Im^-} + d_{t,w}^{load} = L_{t,w}^{De^-} \forall t, \omega \quad (33)$$

$$0 \le r_{g,t,w}^{up,rt} \le r_{g,t,w}^{up}, 0 \le r_{n,t,w}^{up,rt} \le r_{n,t,w}^{up} \forall t, \omega, g \quad (34)$$

$$0 \le r_{g,t,w}^{down,rt} \le r_{g,t,w}^{down}, 0 \le r_{n,t,w}^{down,rt} \le r_{n,t,w}^{down} \forall t, \omega, n \quad (35)$$

$$0 \le P_{n,t,w}^{Im^-,rt} \le P_{n,t,w}^{Im^-}, 0 \le P_{n,t,w}^{Im^+,rt} \le P_{n,t,w}^{Im^+} \forall t, \omega, n \quad (36)$$

where $L_{t,w}^{De^+}$, $L_{t,w}^{De^-}$ is the net load deviation. $L_{t,w}^{De^+}$ indicates that the renewable energy output is more than the prediction output, so it is necessary to dispatch up-regulation power, down reserve capacity provided by thermal units, up reserve capacity provided by LAs, and renewable energy curtailment to maintain the system balance. $L_{t,w}^{De^-}$ indicates that the renewable energy output is less than the prediction output, so it is necessary to dispatch the down-regulation power, up reserve capacity provided by the thermal units, and down reserve capacity provided by LAs and load curtailment. $d_{t,w}^{\text{load}}$ is the system load curtailment. $d_{t,w}^{re}$ is the system renewable energy curtailment. $r_{g,t,w}^{up,rt}$,$r_{g,t,w}^{\text{down},rt}$,$r_{n,t,w}^{up,rt}$, $r_{n,t,w}^{\text{down},rt}$ is the up/down reserve capacity provided by thermal unit g and LAn. $P_{n,t,w}^{Im^+,rt}$, $P_{n,t,w}^{Im^-,rt}$ is the up/down-regulation power dispatched by the system operator. $\delta_{t,w}$ is the price to curtail renewable energy and load. Constraints 34–36 guarantee that the flexible resources provided by each market participant are within the market-clearing range.

The bidding process of LAs and the clearing process of the system operator in the multi-stage market is shown in Figure 3.

Case study

Basic data

A modified IEEE 30 bus system is used to verify the proposed framework as shown in Figure 4. We have three LAs = {LA1, LA2, LA3} in charge of load at the corresponding bus, where each LA will aggregate 1,000 consumers with flexible resources. The system has 3 thermal units G = {G1, G2, G3} and 3 renewable energy generators, i.e., WT. Renewable energy generators are not considered participators in the market. The detailed parameters of LAs, thermal units and WT are shown in Table A1, Table A2 and Table A3 in the Appendix. Suppose that the real-time balance market will be open 1 h in advance, then the predicted load profile of day-ahead data and the predicted electricity prices of day-ahead and real-time data are shown in Figure 5. Other price parameters are shown in Table 1. In this paper, we assume that the reserve rates for renewable energy and the load are $\phi^G = 0.02$ and $\phi^D = 0.01$.

Considering the scenario below:

Scenario one: the real-time renewable energy generation is much more than the day-ahead prediction as shown in Figure 6A. Scenario two: the real-time load is much more than the day-ahead prediction as shown in Figure 6B.

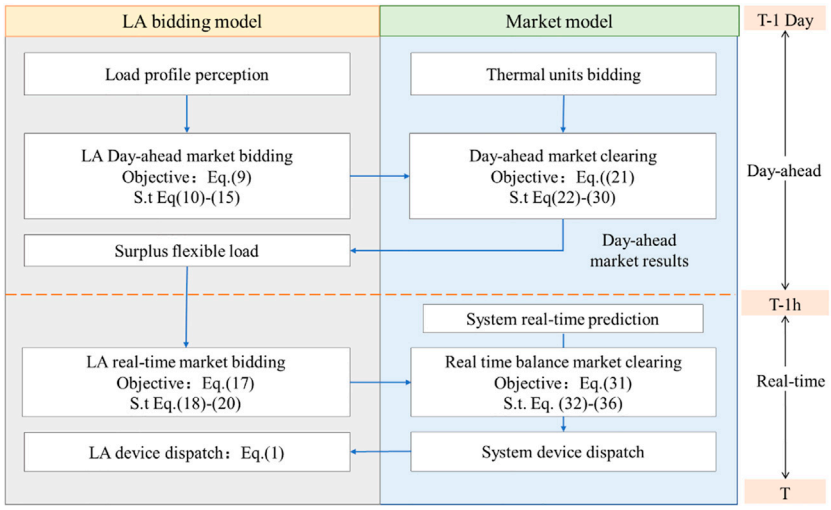


FIGURE 3
The flowchart for market-bidding and clearing.

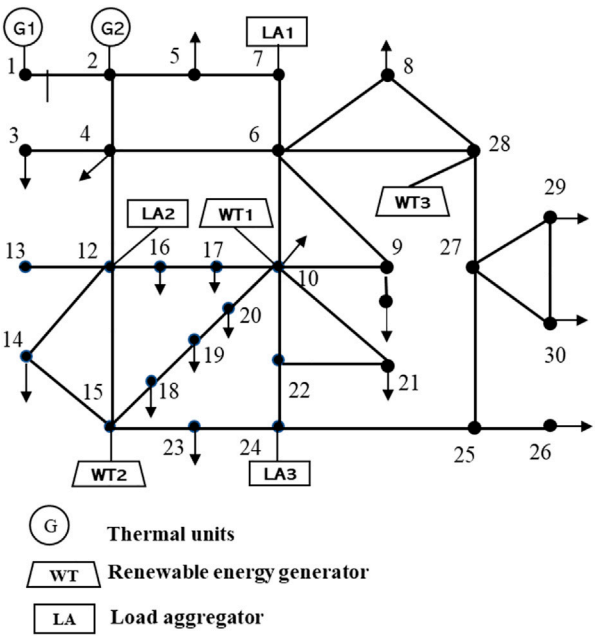


FIGURE 4
The modified IEEE 30 bus system.

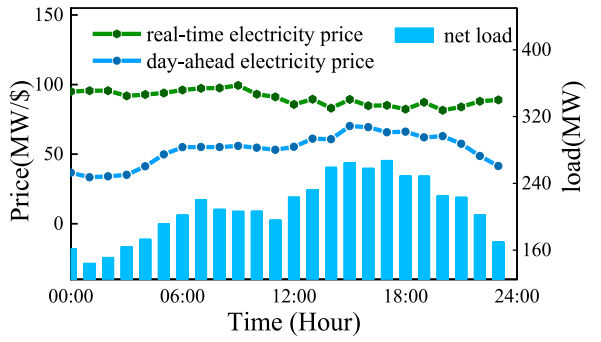


FIGURE 5
Day-ahead prediction for the net load and electricity price.

TABLE 1 Parameter setting of prices.

Parameters	$\lambda^{P}_{n,t,w}$	$\lambda^{div}_{n,t,w}$	$\delta_{t,w}$	$\lambda^{S}_{n,t,w}$
Price ($/MWh)	100	300	500	$1.5\lambda^{A}_{n,t,w}$

Market-bidding results

1) Day-ahead market-clearing result

The bidding results of LAs and thermal units in the day-ahead energy market and reserve market are shown in Figure 7. In the day-ahead energy market and due to the differential bidding strategies, LAs and thermal units have different bidding quantity.

In the day-ahead reserve market, peak load and peak electricity prices usually appear during the afternoon and night, and thus, consumers are more inclined to take part in the DR program. LA can aggregate a more price-sensitive load to join the reserve market as shown in Figure 7B.

2) Real-time balance market-clearing result

Once the day-ahead dispatching schedule for thermal units is decided, more flexible regulation resources may be required from the real-time market to balance the system supply-demand status under certain extreme scenarios.

The system has a different regulation demand under different scenarios. For example, under scenario one where the real-time renewable energy output is much more than the day-ahead predicted output, to consume more renewable energy, the system operator needs to dispatch the up-regulation resources of LA from the real-time balance market and day-ahead reserve market, and the down-regulation resource of thermal units from day-ahead reserve market is dispatched. Thus, all LAs only win up-regulation resources from the real-time balance market without winning down-regulation resources as shown in Figure 8A.

Moreover, under scenario two where real-time load is much more than the day-ahead predicted load, to provide more power, the system operator needs to dispatch the down-regulation resources of LA from the real-time balance market and day-ahead reserve market, and to up-regulation resource of thermal units from day-ahead reserve market are dispatched. Thus, all LAs only win down-regulation resources from the real-time balance market without winning up-regulation resources as shown in Figure 9B.

Market participants' revenue

The benefits and costs of three LAs and their signed consumers are shown in Table 2. According to Table 2, the economic benefits of each LA with their signed consumers have been greatly improved after participating in the market. Under scenario one, the benefits of LAs have increased by 3.0%, 6.5% and 11.9% and the costs of signed consumers of each LA are reduced by 10.6%, 2.6% and 4.7%. Under scenario two, the benefits of LAs have increased by 16.5%, 13.4% and 25.1% and the costs of signed consumers of each LA are reduced by 12.7%, 4.5% and 7.6%.

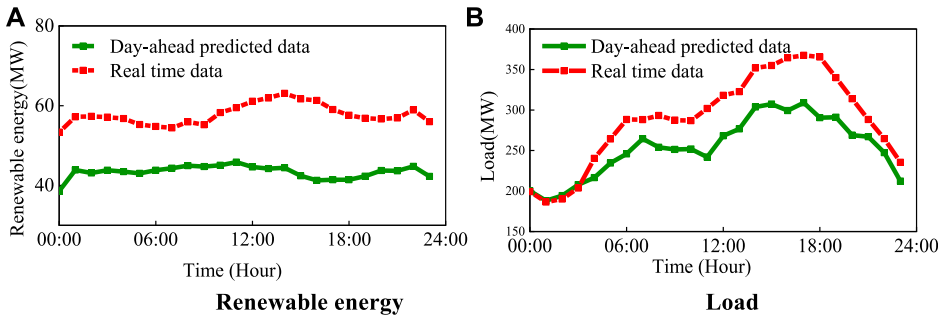


FIGURE 6
Transaction process of the slot-ahead ancillary market for peak-regulation. **(A)** Renewable energy and **(B)** load.

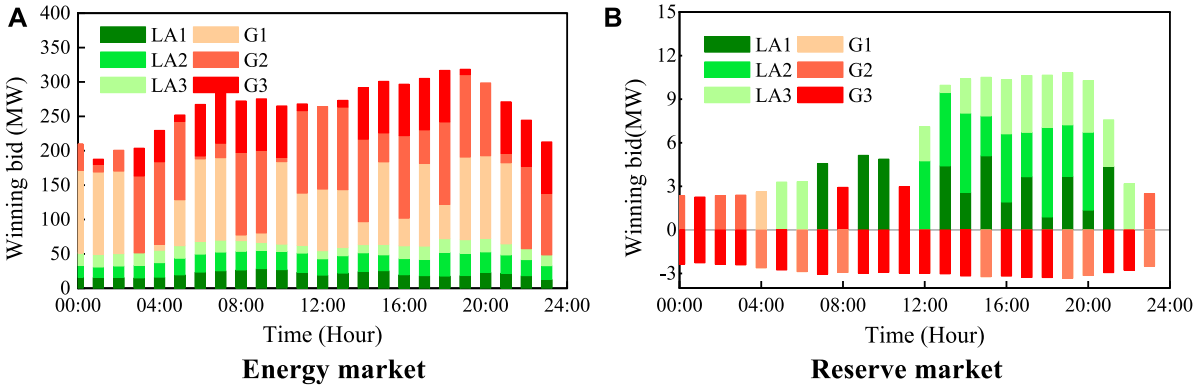


FIGURE 7
The market-clearing results for the day-ahead energy market. **(A)** Energy market and **(B)** reserve market.

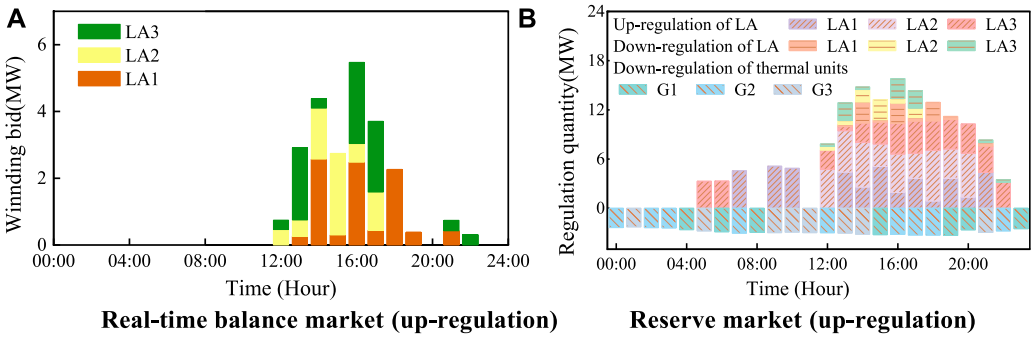


FIGURE 8
The regulation status for balance and reserve resources under scenario one. **(A)** Real-time balance market (up-regulation) and **(B)** reserve market (up-regulation).

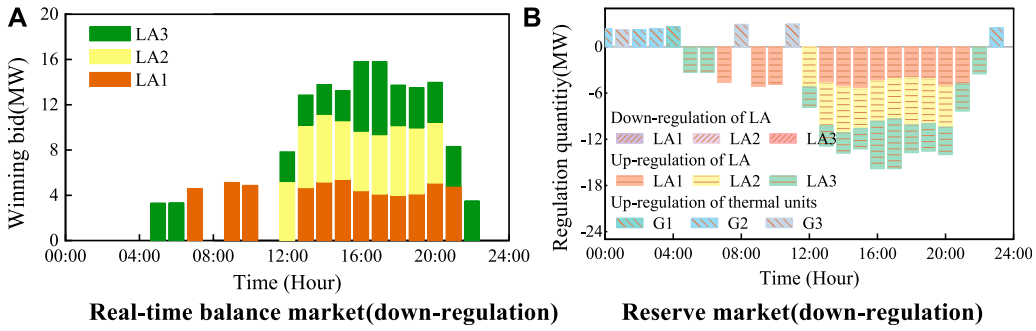


FIGURE 9
The regulation status for balance and reserve resources under scenario two. **(A)** Real-time balance market (down-regulation) and **(B)** reserve market (down-regulation).

TABLE 2 Revenue of market participants.

Scenario	Market participants	The revenue without joining the market ($)	The revenue from joining the day-ahead market ($)	The revenue from joining the day-ahead and real-time market ($)
Scenario one	LA1	15,675.9	15,457.9	16,147.3
	LA2	17,774.3	17,988.6	18,934.2
	LA3	11,568.4	12,152.3	12,948.6
	Signed consumers of LA1	−47028.3	−47341.8	−42054.7
	Signed consumers of LA2	−53321.1	−53008.8	−51915.34
	Signed consumers of LA3	−34703.1	−35492.9	−33080.5
Scenario two	LA1	15,675.9	15,457.9	18,261.5
	LA2	17,774.3	17,988.6	20,162.3
	LA3	11,568.4	12,152.3	14,476.3
	Signed consumers of LA1	−47028.3	−47341.8	−41034.6
	Signed consumers of LA2	−53321.1	−53008.8	−50914.7
	Signed consumers of LA3	−34703.1	−35492.9	−32070.8

TABLE 3 System operation results.

Scenario	Renewable energy/load curtailment (MWh)		System operation fee ($)	
	Without the real-time balance market	With the real-time balance market	Without the real-time balance market	With the real-time balance market
Scenario one	153.58	130.01	169,117.5	158,620.9
Scenario two	528.27	470.40	414,955.9	335,060.2

The system operation result is shown in Table 3. The operation of the real-time balance market can effectively aggregate surplus flexible resources of LAs to participate in market regulation, which is helpful in reducing renewable/load curtailment and system operating cost. Under scenario one, the power curtailment of renewable energy is reduced by 15.3% and the system operating cost is reduced by 6.2% with the operation of real-time balance market. Under scenario

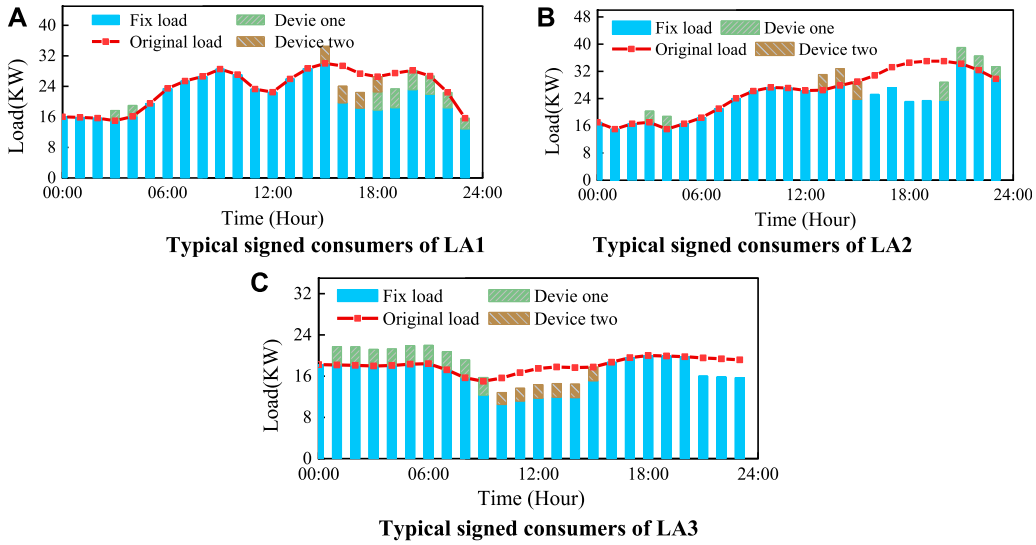


FIGURE 10
The control scheme for end-users' devices. **(A)** Typical signed consumers of LA1, **(B)** typical signed consumers of LA2, and **(C)** typical signed consumers of LA3.

two, the load curtailment is reduced by 11.0% and the system operating cost is reduced by 19.3% with the operation of the real-time balance market.

According to Tables 2 and 3, the proposed multi-stage market framework can provide a win-win market platform for LAs, consumers and the system operator, which helps to guarantee the economic benefits of each market participant and reduces the operating cost of the system.

LA dispatch results

Under scenario two, the device dispatch scheme of typical signed consumers of each LA is shown in Figure 10. Figure 7 shows that the winning bid of LAs in the day-ahead reserve market is up-regulation capacity without down-regulation capacity. As shown in Figure 9, the winning bid of LAs in the real-time balance market is the down-regulation capacity, who will reduce their power usage in the afternoon and night. Each LA will dispatch the controllable devices of their signed consumers in response to system regulation demand. Type one of controllable device is the washing machine and type two is the EV. All controllable devices are forbidden from being used during the peak load period and are to be used during other periods as shown in Figure 10.

Conclusion

This paper establishes a multi-stage electricity market-bidding and clearing framework including the day-ahead energy market, the day-ahead reserve market and the real-time balance market. A modified IEEE 30 bus system is utilized to verify the effectiveness of the proposed market framework, and the simulation results show that a win-win market trading platform for LAs, consumers and the system operator is proposed, which can effectively coordinate the interests of multiple market participants.

1) A load profile perception model for LA is established to achieve accurate load prediction and optimal device control. Moreover, the bidding strategies for LAs are also proposed to assist LA to participate efficiently in markets.

2) By establishing a real-time balance market, LAs can effectively aggregate flexible resources based on signed consumers, and encourage signed consumers to adjust load profiles of controllable devices to respond to system regulation demand. As system renewable energy is abundant, consumers can provide

more up-regulation capacity to consume renewable energy, whereas when the system is short of power, consumers can provide more down-regulation capacity to maintain the system supply-demand balance.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, and further inquiries can be directed to the corresponding author.

Author contributions

XW: methodology, software, writing—original draft. YS and WX: conceptualization, validation, investigation, data curation. SZ: validation, writing—review and editing. YZ: validation, visualization.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Baharlouei, Z., Hashemi, M., Narimani, H., and Mohsenian-Rad, H. (2013). Achieving optimality and fairness in autonomous demand response:Benchmarks and billing mechanisms. *IEEE Trans. Smart Grid* 4 (2), 968–975. doi:10.1109/tsg.2012.2228241

Bruninx, K., Pandžić, H., Le Cadre, H., and Delarue, E. (2020). On the interaction between aggregators, electricity markets and residential demand response providers. *IEEE Trans. Power Syst.* 35 (2), 840–853. doi:10.1109/tpwrs.2019.2943670

California Independent System Operator (2016). Addendum-draft final technical appendix-flexible ramping product. Report.

Chen, S., Chen, Q., and Xu, Y. (2018). Strategic bidding and compensation mechanism for a load aggregator with direct thermostat control capabilities. *IEEE Trans. Smart Grid* 9 (3), 2327–2336. doi:10.1109/TSG.2016.2611611

Du, Y., Lu, F., and Zandi, H. (2021). Approximating Nash equilibrium in day-ahead electricity market bidding with multi-agent deep reinforcement learning. *Mod. Power Syst. Clean Energy* 9 (3), 534–544.

Fang, X., Hu, Q., Li, F., Wang, B., and Li, Y. (2016). Coupon-based demand response considering wind power uncertainty: A strategic bidding model for load serving entities. *IEEE Trans. Power Syst.* 31 (2), 1025–1037. doi:10.1109/tpwrs.2015.2431271

Hu, J., Cao, J., Guerrero, J. M., Yong, T., and Yu, J. (2017). Improving frequency stability based on distributed control of multiple load aggregators[J]. *IEEE Trans. Smart Grid* 8 (4), 1553–1567. doi:10.1109/tsg.2015.2491340

Li, Z., Wang, S., Zheng, X., de Leon, F., and Hong, T. (2018). Dynamic demand response using customer coupons considering multiple load aggregators to simultaneously achieve efficiency and fairness. *IEEE Trans. Smart Grid* 9 (4), 3112–3121. doi:10.1109/tsg.2016.2627140

Liang, B., Yang, J., Hou, B., and He, Z. (2021). A pricing method for distribution system Aggregators considering differentiated load types and price uncertainty. *IEEE Trans. Power Syst.* 36 (3), 1973–1983. doi:10.1109/tpwrs.2020.3032593

Liang, B., Yang, J., Hou, B., and He, Z. (2018). A Pricing Method for Distribution System Aggregators Considering Differentiated Load Types and Price Uncertainty. *IEEE Trans. Power Syst* 36 (3), 1973–1983.

Pandžić, H., Kuzle, I., and Capuder, T. (2013). Virtual power plant mid-term dispatch optimization, *Appl. Energy* 101, 134–141. doi:10.1016/j.apenergy.2012.05.039

Shao, S., Pipattanasomporn, M., and Rahman, S. (2013). Development of physical-based demand response-enabled residential load models. *IEEE Trans. Power Syst.* 28 (2), 607–614. doi:10.1109/tpwrs.2012.2208232

Sumaiti, A., Konda, S. R., Panwar, L., Gupta, V., Kumar, R., and Panigrahi, B. K. (2020). Aggregated demand response scheduling in competitive market considering load behavior through fuzzy intelligence. *IEEE Trans. Ind. Appl.* 56 (4), 1–4247. doi:10.1109/tia.2020.2988853

Vivekananthan, C., Mishra, Y., Ledwich, G., and Li, F. (2014). Demand response for residential appliances via customer reward scheme. *IEEE Trans. Smart Grid* 5 (2), 809–820. doi:10.1109/tsg.2014.2298514

Yang, H., Zhang, S., Qiu, D., Zhao, J., Lai, M., Dong, Z. Y., et al. (2018). Distributionally robust optimal bidding of controllable load aggregators in the electricity market. *IEEE Trans. Power Syst.* 33 (1), 1089–1091. doi:10.1109/tpwrs.2017.2701650

Youbo, L., Junbo, Z., Lixiong, X., Liu, T., Qiu, G., and Liu, J. (2019). Online TTC Estimation using nonparametric analytics considering wind power integration. *IEEE Trans. Power Syst.* 34 (1), 494–505. doi:10.1109/tpwrs.2018.2867953

Youbo, L., Kunyu, Z., Xueqin Amy, L., Liu, J., and Kennedy, J. M. (2018). Dynamic pricing for decentralized energy trading in Micro-Grids. *Appl. Energy* 228, 689–699. doi:10.1016/j.apenergy.2018.06.124

10.3389/fenrg.2022.1030076

Appendix

TABLE A1 Parameter settings of load aggregators.

Market participants	LA1	LA2	LA3
$d_{n,w}$	30%	40%	25%
$\lambda_{n,t,c}^{bid}$	[30, 50, 70, 90]	[25, 50, 75, 100]	[35, 50, 65, 80]
Bus	7	12	24

TABLE A2 Parameter settings of thermal units.

Market participants	G1	G2	G3
P_{max}	120	120	75
Bus	1	2	22

TABLE A3 Parameter settings of wind turbine.

Market participants	WT1	WT2	WT3
P_{max} (MW)	120	80	80
Bus	10	15	28

frontiers | Frontiers in Energy Research

TYPE Original Research
PUBLISHED 12 January 2023
DOI 10.3389/fenrg.2022.1003315

OPEN ACCESS

EDITED BY
Youbo Liu,
Sichuan University, China

REVIEWED BY
Wenqian Yin,
The University of Hong Kong, Hong Kong SAR, China
Youjun Deng,
Tianjin University, China

*CORRESPONDENCE
Zhaobin Du,
epduzb@scut.edu.cn

SPECIALTY SECTION
This article was submitted to Smart Grids,
a section of the journal
Frontiers in Energy Research

RECEIVED 26 July 2022
ACCEPTED 24 October 2022
PUBLISHED 12 January 2023

CITATION
Fu K, Du Z, Li F, Li Z and Xia C (2023), Optimal allocation of phase shifting transformer with uncertain wind power based on dynamic programming.
Front. Energy Res. 10:1003315.
doi: 10.3389/fenrg.2022.1003315

Optimal allocation of phase shifting transformer with uncertain wind power based on dynamic programming

Kang Fu[1], Zhaobin Du[1,2]*, Feng Li[3], Zuohong Li[3] and Chengjun Xia[1,2]

[1]South China University of Technology, Guangzhou, China, [2]Guangdong Provincial Key Laboratory of Intelligent Operation and Control for New Energy Power System, Guangzhou, China, [3]The Grid Planning and Research Center of Guangdong Power Grid Corporation, Guangzhou, China

Phase Shifting Transformer (PST) can help improve the power flow distribution of the transmission section, which can increase the wind power consumption of the grid. In order to adapt the PST allocation to the grid evolution, this paper presents a dynamic programming method to allocate PST in each planning stage of the grid optimally. The optimal allocation model of PST under a single grid seeks to maximize the wind power consumption and the Total Transfer Capacity (TTC) between areas. A calculation method for TTC of grids containing PST and wind power is proposed. The Non-Dominated Sorting Genetic Algorithm II (NSGA2) is used to solve the Pareto sets under each planning stage of the grid. Then, the optimal planning path of PST is derived based on dynamic programming. The superiority of the proposed method is demonstrated by comparing the IEEE-118 system results of dynamic and static programming.

KEYWORDS

phase shifting transformer, dynamic programming, optimal allocation, wind power uncertainty, non-dominated sorting genetic algorithm II

1 Introduction

The application of new energy sources in the grid has been rapidly developed, with the continuous process of energy transition of the grid. In 2021, over 134 GW of renewable power capacity was added in China, making up 76.1% of the newly installed power generation capacity (NEA, 2022). The annual growth rate of installed wind power generation is over 25%, and wind power has become the most rapidly developing renewable energy. However, the construction cycles of the grid and wind farms are not in sync while the amount of wind power grid-connected is increasing rapidly, resulting in the limitations on wind power delivery capacity in areas where wind power is concentrated (Zhang et al., 2020).

The traditional solutions of grid strengthening have problems such as high investment costs, low utilization, and increased environmental pressure. Installing Flexible AC Transmission System (FACTS) devices is another solution (IEA, 2017). FACTS

devices can improve the wind power delivery capacity by re-dispatching the power flow distribution of the transmission section. It plays a good transition role during the planning cycle of the grid. The primary FACTS devices used currently are (Ghahremani et al., 2013): Static Var Compensator (SVC), static synchronous compensator, Thyristor Controlled Series Compensator (TCSC), PST, and Unified Power Flow Controller (UPFC). Among them, PST and UPFC have a more significant impact on wind power integration (Zhang et al., 2018). Still, the installation and operation costs of PST are much lower than those of UPFC, giving a substantial economic advantage (Brilinskii et al., 2020). Installing PST in the grid can effectively improve the power flow distribution of the transmission section, thus improving the transmission capacity and wind power consumption of the grid.

Currently, many studies have proposed optimization models for the allocation of PST. The models can be divided into two categories according to their optimization objectives. The first category mainly focuses on optimizing traditional power system indicators such as active power loss, power flow balance, grid transmission capacity, and voltage profile (Preedavichit et al., 1998). In (Verboomen et al., 2008), the phase shifter distribution factor based on the DC load flow has been derived, and a method is proposed to calculate the TTC for grids containing PST. In (Sebaa et al., 2014), a multi-objective optimization model is constructed to solve the optimal allocations of PST and SVC, considering active power loss, power flow balance, and voltage stability as optimization indicators (Gerbex et al., 2001). gives the installation numbers of TCSC, Thyristor-Controlled Phase Shifting Transformer (TCPST), Thyristor-Controlled Voltage Regulator (TCVR), and SVC. The optimization objective is to maximize the network transmission power, and the optimal allocations of each device are found by a genetic algorithm. In (Kazemi and Sharifi., 2006), the optimal location of PST is found with congestion management in normal and emergency conditions, resulting in reduced production cost and increased load capacity of the power market. In (Wu et al., 2008), a group search optimizer with multiple producers is presented to optimize the positions of TCSC, TCPST, and TCVR, and their control parameters to minimize the active power loss and improve the voltage profile. In (Lima et al., 2003), an optimal model for allocating TCPST is presented. It uses mixed integer linear programming to maximize the system load capacity. However, the results of the DC optimal power flow model are subject to errors. The second category mainly focuses on the optimization of wind power consumption. In (Miranda and Alves., 2014; Zhang et al., 2021), the optimal location of PST is found to maximize wind power consumption, but the investment cost of PST is not considered, and the number of PST is limited to one (Zhang et al., 2017). proposes a bilevel optimization model to solve the optimal locations of PST in the transmission network. The upper level problem seeks to minimize the investment costs on series FACTS, the cost of wind power curtailment, and possible load shedding. The lower level problems capture the market clearing under different operating scenarios. In (Zhang et al., 2018), a bilevel optimization model for the optimal locations of TCSC and PST in the transmission network is proposed, and the uncertainty of wind power is considered. The proposed optimal models for allocations of PST and other FACTS devices only consider the case of a single grid throughout the existing studies. These studies can only compose the planning path of PST by solving for the optimal allocation scheme of each stage, which may not be the optimal path if multi-stage grid planning is considered. This paper proposes a dynamic planning method to find the optimal allocation path of PST to solve this problem, considering the grid evolution and wind power uncertainty.

In terms of PST single-stage allocation, the existing optimization models (the second category of optimization models mentioned above) only consider wind power consumption and investment cost of PST without considering the traditional optimization indicators of the power system. This paper finds that installing PST in the grid can increase the TTC of the grid while increasing the wind power consumption. However, wind power's uncertainty will greatly affect TTC's calculation (Wang et al., 2021), so the traditional TTC calculation method (Verboomen et al., 2008) is no longer applicable. This paper proposes a method to calculate the TTC of grids with PST and wind power. Therefore, a multi-objective optimization model is proposed to find the Pareto sets of PST allocation schemes under a single grid, considering wind power consumption, investment and operation costs of PST, and TTC as optimization indicators.

In solving the optimal allocation model of PST, existing methods are mainly divided into two categories. The first category uses mixed integer programming to solve the DC optimal power flow model (Lima et al., 2003; Zhang et al., 2018). The time required for this solution is short, but a secondary verification is necessary because of the errors. The second category uses intelligent algorithms to solve the optimal model [GA (Gerbex et al., 2001), PSO (Zhang et al., 2021), *etc.*], which yields more accurate results but takes more time. Since solving the PST optimal allocation model PST does not require high computational speed, NSGA2 is applied in this paper to ensure the accuracy of the results.

To sum up, this paper proposes a dynamic programming method to find the optimal path planning of PST, considering the grid evolution. First, we build a single-stage multi-objective optimization model, with wind power consumption, investment and operation costs of PST, and TTC as optimization indicators. Second, NSGA2 is used to find the Pareto sets of each grid planning stage. Then a Technique for Order Preference by Similarity to an Ideal Solution (TOPSIS) is used to calculate the adaptation values for each allocation scheme in the Pareto sets. Finally, the optimal allocation path of PST is obtained

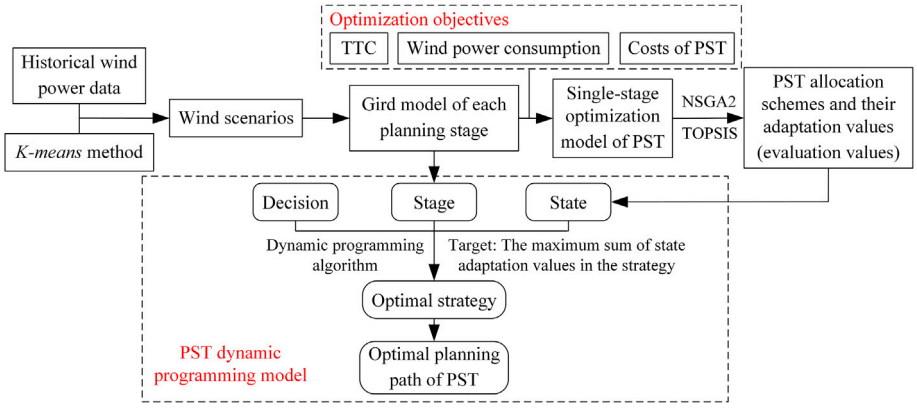


FIGURE 1
The framework of the PST allocation method.

based on the dynamic programming model. The main contributions of this paper are as follows:

1) A dynamic programming method is proposed to solve the optimal allocation path of PST in multi-stage grid planning, considering wind power uncertainty.
2) A method for calculating the TTC of grids containing PST and wind power is proposed. TTC is considered as an optimization indicator for the PST single-stage allocation model.

The rest of this paper is organized as follows. Section 2 shows the main framework of the proposed method. Section 3 introduces the basic principle and steady-state model of PST. Section 4 presents the PST single-stage allocation model. The solution based on NSGA2 is demonstrated in Section 5. Section 6 describes the dynamic programming model of PST. Section 7 verifies the effectiveness of the proposed method by comparing the IEEE 118-bus system results of dynamic and static programming. Finally, the main findings of this study are summarized with some prospects for future studies in the conclusion section.

2 Main framework of the proposed method

The proposed method is developed on the main framework of dynamic programming, and the main idea is to allocate PST dynamically in the multi-stage planning of the grid. The method has two main steps, and the framework of the proposed method is given in Figure 1.

First, the PST single-stage planning models are built for each planning stage of the grid. We consider TTC, wind power consumption, and PST costs as the optimization objectives to find the PST allocation schemes of each planning stage. TOPSIS is used to evaluate each allocation scheme by calculating its evaluation values.

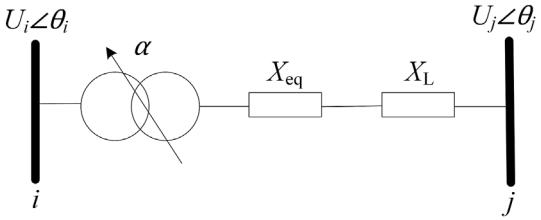


FIGURE 2
The equivalent circuit of the branch with PST.

Second, the dynamic programming model is constructed and solved for the optimal planning path of the PST. The dynamic planning model is built by considering each planning stage of the grid as a stage, the PST allocation scheme of each stage as a state, and the scheme's evaluation value as the state's adaptation value. The target is to maximize the sum of the adaptation values of the states in the strategy. Then the dynamic programming model is solved to obtain the optimal multi-stage decision, and the optimal PST planning path is obtained.

3 Basic principle and steady-state model of PST

PST changes the power flow by injecting a voltage vector into the line, resulting in the variation of the phase and voltage amplitude at both ends of the line (Ding et al., 2017).

Figure 2 depicts the equivalent circuit of the branch with PST (Yeo et al., 2019), and the resistances of the line and PST are ignored here. In Figure 2, X_L represents the equivalent reactance of the line, X_{eq} represents the equivalent reactance of the PST, α is the phase-shift angle, and U_i, θ_i, U_j, and θ_j represent the voltage amplitude and the phase of nodes i and j, respectively. P_{ij} is the active power of line i-j.

$$P_{ij} = \frac{U_i U_j}{X_L} \sin\left(\theta_i - \theta_j\right), \tag{1}$$

$$P_{ij} = \frac{U_i U_j}{X_L + X_{eq}} \sin\left(\theta_i + \alpha - \theta_j\right), \tag{2}$$

Eqs 1, 2 represent the active power of the line before and after the installation of the PST, respectively. It can be seen that the phase angle difference has changed after the PST installation, and the active power of the line has been regulated.

4 Multi-objective optimal model of power systems with PST

4.1 Objective functions

The large-scale grid-connected wind power will increase the power flow of transmission lines, resulting in problems such as transmission line overloading and inter-regional transmission capacity limitations. The main focus of this study is to improve the wind power consumption while considering the investment cost of PST, and the TTC between grid areas is also one of the concerns.

4.1.1 Wind power consumption

The first objective function (Eq. 3) seeks to maximize the wind power consumption.

$$\max f_1 = \frac{\sum_{k=1}^{S_n} c_k W_k}{\sum_{k=1}^{S_n} c_k W_{k\max}}, \tag{3}$$

where c_k is the probability of scenario k. W_k and $W_{k\max}$ are the wind power consumption and total wind power output under scenario k, respectively. S_n is the number of scenarios.

4.1.2 Total transfer capacity

The second objective function (Eq. 4) is to maximize the TTC between grid areas.

$$\max f_2 = TTC. \tag{4}$$

We use the scenario method and linear programming to solve the TTC between the source and receiving areas. The steps are as follows:

Step 1. Use *K-means* clustering to reduce the number of scenarios to S_n, $i = 1$.

Step 2. Select a typical condition of the grid and the wind power output of scenario i is substituted into the typical condition of the grid as the base grid. The AC power flow of the base grid is calculated. Whether the wind power in the source area can be fully consumed, if so, then go to Step3; if not, then TTC is calculated as the power flow of the transmission section between the source and the receiver area.

Step 3. Calculate the sensitivity factor for line l (s_l) when the power output of all generators in the source area vary except for the balancer's.

$$s_l = \frac{dP_l}{d\Delta E}, \tag{5}$$

where ΔE denotes the variation of the power output of all generators in the source area except the balancer.

Step 4. The expression of the active power of line l is obtained based on the sensitivity factor of PST to line l.

$$P_l = P_{l0} + \sum_{j=1}^{N_p} \alpha_j \xi_l^j, \tag{6}$$

where ξ_l^j denotes the sensitivity factor of the PST j to the active power of line l (Li et al., 2022a). P_{l0} is the original active power of line l in the base grid. N_p is the number of PSTs installed. α_j is the phase-shift angle of PST j.

Step 5. The final active power expression for line l is obtained from Eqs 5, 6.

$$P_l = P_{l0} + s_l \cdot \Delta E + \sum_{j=1}^{Np} \alpha_j \xi_l^j. \tag{7}$$

Step 6. A linear programming model is built to maximize the TTC, with the phase shift angle and ΔE as the decision variables.

$$\max TTC_i = \sum_{i \in \Omega_t} P_i, \tag{8}$$

$$\text{s.t.} \begin{cases} -P_{l\max} \le P_l \le P_{l\max}, \\ \alpha_{\min} \le \alpha \le \alpha_{\max}, \end{cases} \tag{9}$$

where Ω_t denotes the regional interlink lines set, and $P_{l\max}$ is the maximum permissible powers limit of line l.

Step 7. Solve the linear programming model to get the value of the phase shift angle and ΔE. Then consider these values as the initial point of the Repeated Power Flow method (RPF) and obtain the value of TTC, $i = i+1$.

Step 8. If $i = S_n$, go to Step 9; If $i < S_n$, back to Step 2.

Step 9. Calculate the expected value with Eq. 10 considering the probability of scenarios, and the obtained expectation value is recorded as the *ETTC*.

$$\mathbf{ETTC} = \sum_{i=1}^{S_n} \boldsymbol{TTC}_i \cdot \boldsymbol{c}_i. \tag{10}$$

The RPF in Step7 is based on the conventional AC power flow, and its basic idea is as follows. From a specific base state, the output of the generation area is gradually increased while the load of the receiving area is increased correspondingly. The AC power flow in the power increasing process is repeatedly calculated. A series of power flow solution points are obtained, and various constraints are checked on these solution points until the maximum adjustment of the generator that fulfils all constraints is found. At this point, the active power of the transmission section between the transmitting and receiving areas is considered as the TTC.

4.1.3 Investment and operation costs of PST

The third objective function (Eq. 11) means to minimize the investment and operation costs of PST (Ippolito and Siano., 2004).

$$\mathbf{min} \boldsymbol{f}_3 = \sum_{i=1}^{N_p} \boldsymbol{C}_{\boldsymbol{Pi}} + \boldsymbol{\beta} \cdot \boldsymbol{T}_i \cdot \boldsymbol{C}_{\boldsymbol{Pi}}, \tag{11}$$

$$\boldsymbol{C}_{\boldsymbol{Pi}} = \boldsymbol{\gamma} \cdot \boldsymbol{S}_{\boldsymbol{Pi}}, \tag{12}$$

where C_{Pi} is the investment cost of the PST i. β denotes the annual operating cost factor. T_i is the usage time of the PST i. γ is the cost factor of PST. S_{Pi} is the capacity of the PST i.

It is worth noting that the decision variables selected in the optimal model are the location of the PST, the number of PSTs, the phase shift angles, and the wind power output.

4.2 Power system constraints

4.2.1 Equation constraints

The active and reactive power constraint equations are as follows:

$$\boldsymbol{P}_{\mathrm{G}i} - \boldsymbol{P}_{\mathrm{L}i} - \boldsymbol{U}_i \sum_{j=1}^{n_b} \boldsymbol{U}_j \left(\boldsymbol{G}_{ij} \cos \boldsymbol{\theta}_{ij} + \boldsymbol{B}_{ij} \sin \boldsymbol{\theta}_{ij} \right) = \mathbf{0} \tag{13}$$

$$\boldsymbol{Q}_{\mathrm{G}i} - \boldsymbol{Q}_{\mathrm{L}i} - \boldsymbol{U}_i \sum_{j=1}^{n_b} \boldsymbol{U}_j \left(\boldsymbol{G}_{ij} \sin \boldsymbol{\theta}_{ij} - \boldsymbol{B}_{ij} \cos \boldsymbol{\theta}_{ij} \right) = \mathbf{0} \tag{14}$$

where P_{Gi} and Q_{Gi} are the active and reactive power output of the i-bus generator, respectively. P_{Li} and Q_{Li} are the load active and reactive power of the bus i. n_b is the number of lines connected to node i. G_{ij} and B_{ij} are the conductance and the susceptance of the line i-j, respectively.

4.2.2 Inequality constraints

Inequality constraints for the active and reactive power outputs of generators, voltage amplitudes of buses, and active power of branches are as follow:

$$\begin{cases} \boldsymbol{U}_{i\,\mathbf{min}} \le \boldsymbol{U}_i \le \boldsymbol{U}_{i\,\mathbf{max}}, \\ \boldsymbol{P}_{\boldsymbol{Gi}\,\mathbf{min}} \le \boldsymbol{P}_{\boldsymbol{Gi}} \le \boldsymbol{P}_{\boldsymbol{Gi}\,\mathbf{max}}, \\ \boldsymbol{Q}_{\boldsymbol{Gi}\,\mathbf{min}} \le \boldsymbol{Q}_{\boldsymbol{Gi}} \le \boldsymbol{Q}_{\boldsymbol{Gi}\,\mathbf{max}}, \\ -\boldsymbol{P}_{\boldsymbol{ij}\,\mathbf{max}} \le \boldsymbol{P}_{\boldsymbol{ij}} \le \boldsymbol{P}_{\boldsymbol{ij}\,\mathbf{max}}, \end{cases} \tag{15}$$

where $U_{i\max}$ and $U_{i\min}$ are the maximum and minimum voltage of node i, respectively. $P_{Gi\max}$ and $P_{Gi\min}$ are the maximum and minimum active power output of generator i, respectively. $Q_{Gi\max}$ and $Q_{Gi\min}$ are the maximum and minimum reactive power output limits of generator i, respectively. $P_{ij\max}$ is the maximum permissible powers limit of line i-j.

Since this paper focuses on the PST allocation in multi-stage grid planning, the time scale between the stages is in years, so the model of the generator is moderately simplified. It is assumed that the responsiveness of conventional power sources is strong enough to consume wind power. So some temporal constraints are not considered, such as ramping constraints of generators.

4.3 Constraints of decision variables

It is known that the PST is mainly used to regulate the power flow of the transmission section where PST is installed (Hadzimuratovic and Fickert., 2018). We have imposed some limits on the range of locations and the number of PSTs to avoid solving unreasonable PST allocation schemes, which are shown as follows.

1) The locations of PST are limited to the transmission sections, which contain the wind power transmission lines.
2) The maximum number of PSTs installed on a line is 1, and the maximum number of PSTs installed in a transmission section is 2.

4.3.1 Locations of PST

$$\boldsymbol{l}_p \in \boldsymbol{D}. \tag{16}$$

l_p is the line with PST installed. D is the set of transmission sections that deliver wind power.

4.3.2 Shift-phase angle

$$\boldsymbol{\alpha}_{\mathbf{min}} \le \boldsymbol{\alpha} \le \boldsymbol{\alpha}_{\mathbf{max}}. \tag{17}$$

$\alpha_{\max}$ and $\alpha_{\min}$ are the maximum and minimum phase angles of PST, respectively.

4.3.3 Number of PSTs

Installing PST in the transmission section can improve transmission capacity by lessening the power flow of heavily loaded lines. Since the number of heavy load lines on a specific transmission section is limited, the transmission capacity

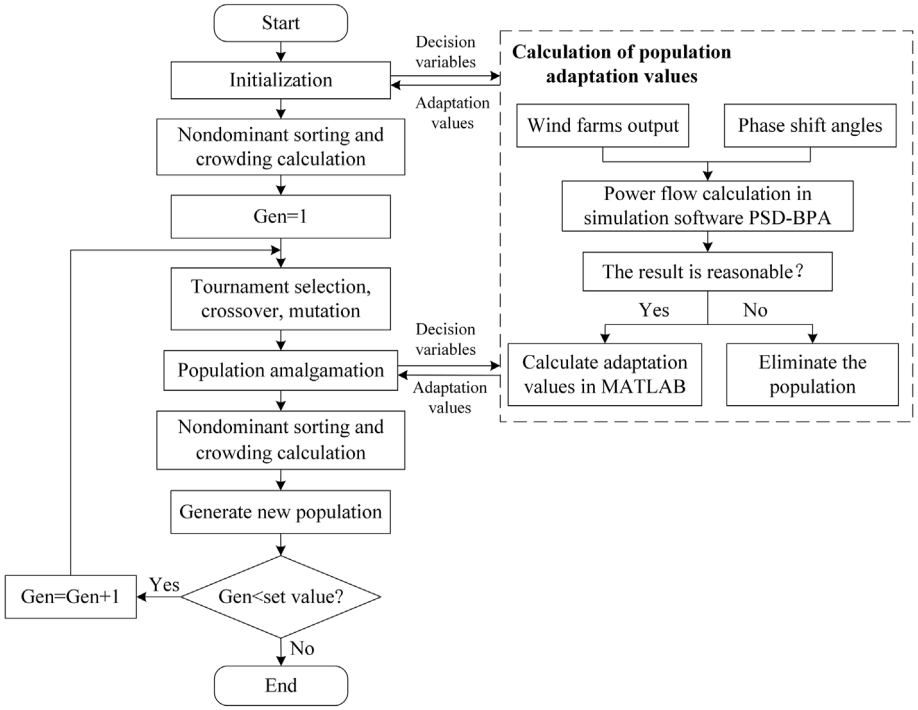


FIGURE 3
The flow chart of the multi-objective optimization model solving.

enhancement will naturally decrease as more PSTs are installed (numerical simulations are performed in Section 7.5 to verify this view).

The number of lines in the transmission section is limited in the real grid. Considering the economics of PST investing, this study limits the PSTs number of the transmission section to less than 2.

$$\boldsymbol{N_i \leq 2}. \tag{18}$$

N_i is the PSTs number of the transmission section-i.

4.3.4 Wind power output

$$\boldsymbol{0 \leq P_{wi} \leq P_{wi\,max}}. \tag{19}$$

P_{wi} and $P_{wi\max}$ are the actual and maximum output of wind power at node i, respectively.

5 Solution approach

The NSGA2 algorithm is used to solve the Pareto sets of the PST optimal allocation model. The phase shift angle and wind power output in the initial population are randomly generated, which may result in unsolvability or unreasonable power flow results [power flow reverse (Li et al., 2022)]. The following improvements are added to NSGA2 to solve this problem, and the flow chart of NSGA2 is shown in Figure 3.

1) An additional judgment is added to eliminate the unreasonable allocation schemes.
2) According to the power flow regulation characteristics of PST, the unreasonable range of phase shift angle is eliminated to accelerate the iteration speed of NSGA2

$$\begin{cases} \boldsymbol{0 < \alpha_i \leq \alpha_{\max} \ldots i \in \Omega_p}, \\ \boldsymbol{\alpha_{\min} \leq \alpha_i < 0 \ldots i \in \Omega_q}. \end{cases} \tag{20}$$

In Eq. 20, Ω_p denotes the set of lines with the lowest load factor of each wind power transmission section, so set $\alpha > 0$ (Over-regulation, increasing the active power of the line). Ω_q denotes the set of lines with the highest load factor of each wind power transmission section, so set $\alpha < 0$ (Hysteresis-regulation, reducing the active power of the line).

After obtaining the Pareto set by NSGA2, the scores of allocation schemes in the Pareto set are found by the TOPSIS method considering the weight coefficients.

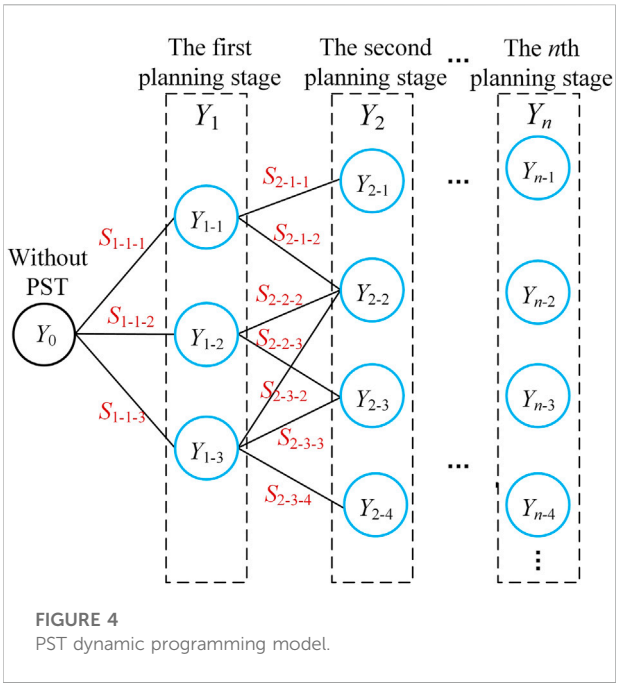


FIGURE 4
PST dynamic programming model.

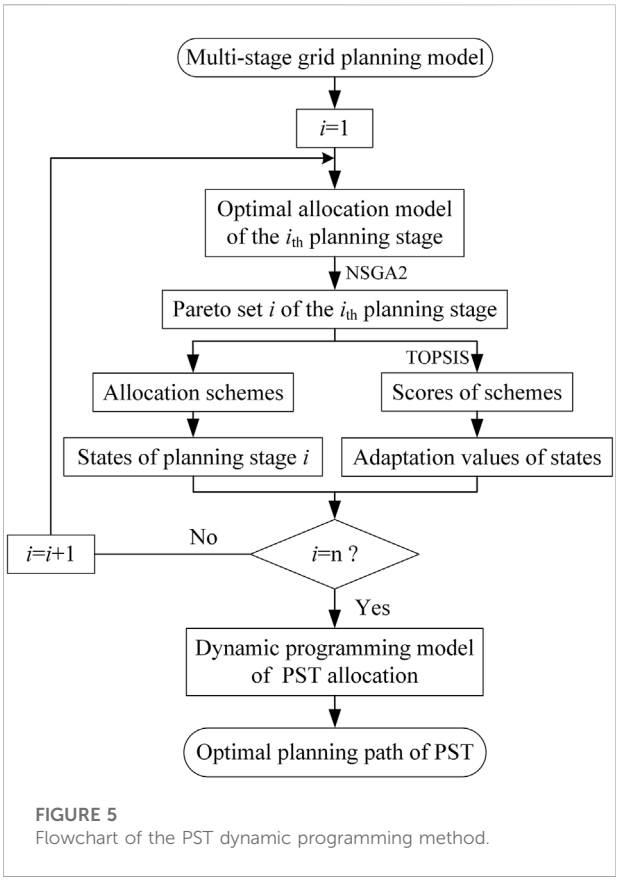


FIGURE 5
Flowchart of the PST dynamic programming method.

6 PST allocation based on dynamic programming

A dynamic programming model of PST is proposed to find the optimal planning path for PST. It ensures that the allocation of PST can be better adapted to the grid evolution. Figure 4 gives the PST dynamic programming model, and the details are as follows.

Stage: Define a stage as the transformation between two adjacent planning stages of the grid, and the number of planning stages is n.

State: Define states as the PST allocation schemes of each planning stage of the grid. In Figure 4, $Y_{i\text{-}j}$ indicates the jth PST allocation scheme under the ith planning stage of the grid.

Decision: The decisions represent the transition choices of PST allocation schemes between two adjacent planning stages of the grid, which are described in Figure 4 as connections between neighboring states. In addition, the set of lines with PST installed in the allocation scheme after the decision must include that of the allocation scheme before the decision (without considering the case of decommissioning or replacement of PST). In Figure 4, J_i is the number of decisions for stage i.

Strategy: Define each complete planning path through all planning stages as a strategy, represented in Figure 4 as a complete concatenation from Y_0 to Y_n.

Target: The target is to select the optimal planning path with the maximum sum of adaptation values. In Figure 4, $S_{i\text{-}p\text{-}q}$ is the adaptation value of decision $Y_{(i\text{-}1)\text{-}p}$-$Y_{i\text{-}q}$ at the current stage i, obtained by the TOPSIS. The specific calculation is as follows.

Step 1: Define all decisions in the current stage as evaluation objects and take the three objective functions in Section 4.1 as evaluation criteria.

$$\min f_3 = \sum_{i=1}^{N_p} (C_{P2} - C_{P1}) + \beta \cdot m \cdot C_{P2}. \tag{21}$$

Equation Eq. 21 shows the improved expression of the cost objective function, where m is the time of a stage. C_{P1} and C_{P2} denote the PST investment cost of the PST allocation scheme before and after the decision, respectively.

Step 2: Use the TOPSIS to obtain the unnormalized scores of all decisions in the current stage.

Step 3: Scores of all decisions at the current stage are processed as in Equation Eq. 22, ensuring that the decision scores of each stage are in the same order of magnitude.

$$S_{i\text{-}j} = h_i \cdot J_i \cdot \frac{S_j}{\sum_{n=1}^{J_i} S_n}, \tag{22}$$

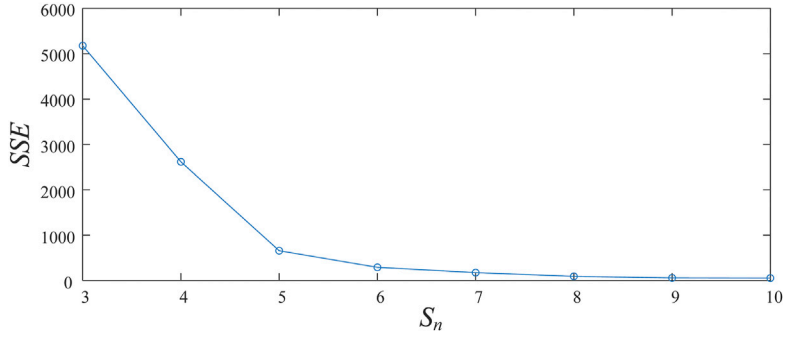


FIGURE 6
Distribution chart of *SSE* and S_n.

TABLE 1 Wind scenarios.

#	Wind intensities/p.u	Probabilities
1	0.78	0.1426
2	0.06	0.2570
3	0.61	0.1657
4	0.42	0.1826
5	0.22	0.2521

$$\boldsymbol{h_1 > h_2 > h_3 \cdots > h_n}, \tag{23}$$

where S_i is the unnormalized score of decision i. $S_{i\text{-}j}$ is the normalized score of decision j in stage i. h_i is the weighting factor of stage i in all stages, reflecting the importance of each grid planning stage. Since the PST plays a transitional role in the grid planning process, the demand for PST becomes smaller as the grid evolves. So we set continuous decreasing weighting factors for successive planning stages, as shown in Equation Eq. 23. The weighting factors values of grids are derived from experience in this study.

In conclusion, the flowchart of the proposed PST dynamic programming method is shown in Figure 5. The steps are as follows.

Step 1: Build a multi-stage planning model of the grid.

Step 2: Build the optimal allocation model of PST for each planning stage (Section 4).

Step 3: Use NSGA2 to solve the optimal allocation model based on the interaction between PSD-BPA (a simulation software that is used to calculate the power flow in this paper) and a calculation procedure in MATLAB (Tao et al., 2013). The scores of the allocation schemes are obtained by the TOPSIS method considering the weight coefficients (Section 5).

Step 4: Each planning stage's allocation schemes are considered states in the dynamic programming model. Each allocation scheme's score is recorded as the adaptation value of each state in the PST dynamic programming model. The optimal planning path of PST is obtained by the dynamic programming algorithm (Section 6).

7 Case study

7.1 Case parameters and wind scenarios

The proposed dynamic programming model and solution approach are tested on the IEEE 118-bus system. The system data is derived from the IEEE standard system. The thermal limits for the transmission lines refer to the values in (Blumsack., 2006). Three wind farms with a maximum capacity of 1600 MW each are assumed to be located at bus 5, 26, and 91 (Ziaee and Choobineh., 2017). The phase shift angle range is set to be (−50°, 50°). The usage time of PST is set to be 15 years (the time of a complete planning path in this study). The cost factor of PST is selected to be 10$/kVA, and the annual operating cost factor is chosen to be 5%.

We built a multi-stage planning grid model in this case. Three planning stages are considered base on the IEEE 118-bus system, and the cycle between two adjacent stages is 5 years. The specific grid models of each stage are as follows.

The first planning stage: Based on the IEEE 118-bus system, the wind farm with the maximum capacity of 1600 MW is located in bus 5, and the load has increased by 20%, $h_1 = 1$.

The second planning stage: Based on the grid of the first planning stage, the wind farm with the maximum capacity of 1600 MW is located in bus 26, and the load has increased by 10%, $h_2 = 0.9$.

The third planning stage: Based on the grid of the second planning stage, the wind farm with the maximum capacity of 1600 MW is located in bus 91. The load in area C (Zhang and

TABLE 2 IEEE 118-bus system results for different planning paths of PST.

Planning stage	Planning path	Wind power consumption (%)			PST locations	Investment cost on PST (M $)	ETTC (MW)	Wind power consumption of all wind farms (%)
		Bus 5	Bus 26	Bus 91				
1st	Without PST	55.84	—	—	—	—	235.5	55.84
	$Road_1$	72.34	—	—	8-30	2.2	333.5	72.34
	$Road_2$	89.84	—	—	8-30, 30-38, 5-3	8.8	535.5	89.84
2nd	Without PST	61.36	70.71	—	—	—	225.1	66.46
	$Road_1$	72.02	98.35	—	8-30, 30-38	4.4	460.6	85.19
	$Road_2$	85.84	99.51	—	8-30, 30-38, 5-3	8.8	586.8	92.67
3rd	Without PST	61.36	70.71	75.80	—	—	225.1	69.29
	$Road_1$	72.02	98.35	78.68	8-30, 30-38, 89-92 (2)	6.6	460.6	83.01
	$Road_2$	85.84	99.51	78.68	8-30, 30-38, 5-3, 89-92 (2)	11	586.8	88.01

Grijalva., 2013) has increased by 10%, and a new transmission line has been added between bus 91 and bus 92, with line parameters consistent with the original lines of bus 91 and bus 92, $h_3 = 0.8$.

To obtain the wind scenarios, the wind power intensities of each 5 minutes in 7 years provided by DR POWER (National Energy Renewable Laboratory, 2019) are used to represent the wind generation profile. We then use the *K-means* method (Baringo and Conejo., 2013) to conduct the scenario reduction, where the optimal number of clusters is obtained by the Elbow Method (Bandara et al., 2019). The SSE distribution plots for different S_n values are shown in Figure 6.

The optimal number of clusters was found to be 5, based on the results in Figure 6 and the Elbow Method. The probabilities and wind intensities for the final five scenarios are provided in Table 1.

7.2 Optimal planning path of PST

In terms of NSGA2 parameters, the initial population is set to be 300, the number of iterations is set to 200, the crossover rate is set to be 0.9, and the variation rate is set to be 0.1. The normalized weight of three evaluation criteria are set to be 0.6, 0.2, and 0.2, respectively.

The Pareto sets of PST allocation for three planning stages are derived from the optimization calculation in Section 5 (the results are shown in Supplementary Tables S1–S3 in the Supplementary Material).

The proposed PST dynamic programming method is a multi-stage decision-making problem, and the connection between each stage is considered. The existing PST allocation methods are mainly used for single-stage optimization. When they are used to solve multi-stage problems, the optimal PST allocation scheme of each stage needs to be found separately and combined into a planning path by order. We define this method as static programming and the result as the static planning path.

Table 2 provides the planning results based on the static planning and the dynamic planning for $Road_1$ (the static planning path) and $Road_2$ (the dynamic planning path), respectively. Columns 2–4 indicate the wind power consumption of each wind farm. The fifth column shows the locations of the PST. The sixth column represents the investment cost of PST. The seventh column denotes the TTC between areas A and B (Zhang and Grijalva., 2013). The last column gives the wind power consumption rate of all wind farms. Doublelines 89-92 (1) and 89-92 (2) denote the lines with smaller and larger impedances between bus 89 and bus 92, respectively. Figure 7 shows the wind power curtailment rate for the three wind farms in each scenario.

As observed in Table 2 and Figure 7, the consumption rate of wind power under $Road_1$ and $Road_2$ has been significantly improved, and the TTC has also been improved. The wind power consumption rate and TTC of each planning stage in $Road_2$ are higher than those in $Road_1$. However, the costs of the three planning stages are higher in $Road_2$, resulting in a better allocation scheme in $Road_1$ than in $Road_2$ when the grid of each planning stage is considered separately.

7.3 Comparison of dynamic programming and static programming

In this section, two comparison cases are proposed to verify the effectiveness of the dynamic programming method proposed in this paper. Both cases compare the results of multi-stage allocation of PST based on static and dynamic programming.

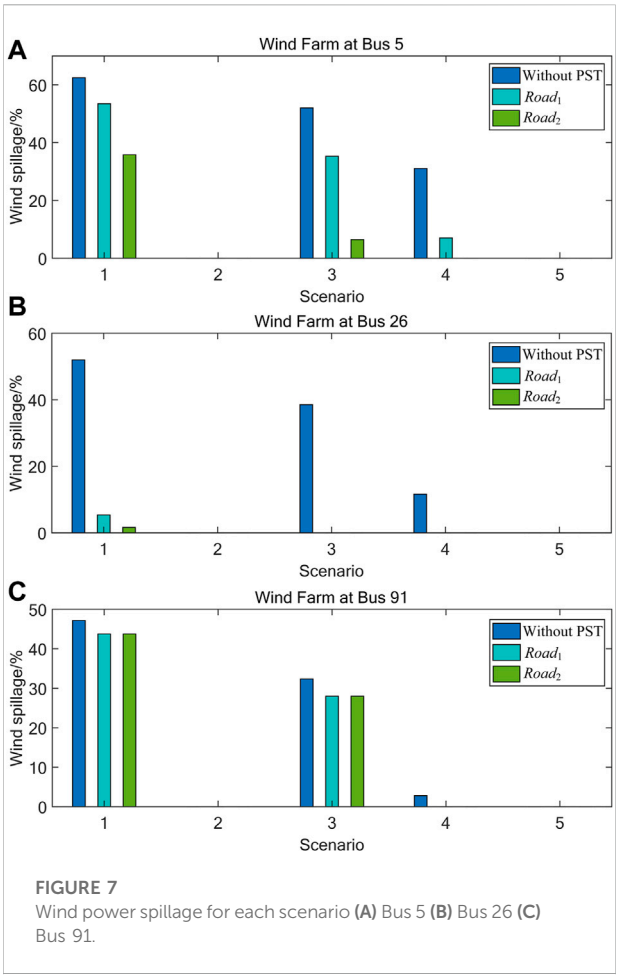


FIGURE 7
Wind power spillage for each scenario **(A)** Bus 5 **(B)** Bus 26 **(C)** Bus 91.

7.3.1 Comparison case 1

The first comparison case uses the single-stage allocation model in Section 4, and the results are presented in Table 3. It compares the IEEE 118-bus system results of $Road_1$ and $Road_2$. The second column indicates the wind power consumption rate for 15 years, namely the time of a complete path in dynamic programming. The third column gives the annual average TTC value. The fourth column shows the total costs of PST in 15 years. The last column represents the sum of adaptation values.

As observed in Table 3, the improvement of wind power consumption under $Road_2$ is 11.94% higher than that under $Road_1$. The increase in TTC of $Road_2$ is 47.36% higher than that of $Road_1$. The sum of adaptation values under $Road_2$ is higher than that under $Road_1$, although the total costs under $Road_2$ are higher than that under $Road_1$, which means that the planning path obtained from the dynamic programming scheme is better than that of the static programming.

7.3.2 Comparison case 2

To further demonstrate the superiority of the proposed dynamic approach, the single-stage optimization model in Section 4 is replaced with the single-stage PST allocation model in (Zhang et al., 2018) while the dynamic programming framework remains unchanged in this case. Then the results of the dynamic and static programming are compared. The optimal allocation model in (Zhang et al., 2018) converts wind curtailment into the corresponding cost by using the cost coefficient of wind curtailment. The multi-objective optimization is converted to a single-objective optimization by considering the sum of the annual wind curtailment cost and the annualized investment cost in PST as the total optimization objective. Therefore, compared to the case1 in Section 7.3.1, there is no need to evaluate the allocation schemes of each stage, and the total costs of the planning path can be obtained by directly summing the costs of all stages according to their weights. Using this result to compare the effects of various planning paths is more intuitive.

The results of the second case are shown in Table 4. The objective value in Table 4 represents the annualized investment cost in PST plus the annual wind curtailment cost. In Table 4, $Road_3$ represents the static planning path obtained based on the static programming method, and $Road_4$ represents the dynamic planning path obtained based on the proposed dynamic programming method.

As observed in Table 4, the wind power consumption rate is higher under $Road_4$ than $Road_3$. The objective value of $Road_4$ is \$8.141M, which is lower than the objective value of $Road_3$ at \$9.139M, which means the planning path under the dynamic programming is better since the sum of the costs for annualized investment in PST and annual wind curtailment lower. So the proposed dynamic planning method is still better after replacing the optimization model in a single stage, indicating the universality of the proposed dynamic planning method.

TABLE 3 IEEE 118-bus system results of $Road_1$ and $Road_2$.

Planning path of PST	Wind power consumption of 15 years (%)	Annual average of TTC (MW)	Costs of PST in 15 years (M $)	Sum of adaptation values
Without PST	66.25	247.7	—	—
$Road_1$	81.96	418.2	9.9	3.19
$Road_2$	89.87	535.5	18.15	3.38

TABLE 4 IEEE 118-bus system results of $Road_3$ and $Road_4$.

Planning path of PST	Wind power consumption of 15 years (%)	Costs of PST in 15 years (M $)	Objective value (M $)
Without PST	66.25	—	—
$Road_3$	89.6	22.0	9.139
$Road_4$	93.1	32.6	8.141

TABLE 5 IEEE 118-bus system results of installing PST and transmission network expansion.

	Wind power consumption of planning stage (%)			Wind power consumption in total (%)
	1	2	3	
Base case	55.84	66.46	69.29	66.11
Adding new lines	78.81	79.06	82.62	80.80
Installing PST	72.34	85.19	83.01	81.96

7.4 Comparison of installing PST and transmission network expansion

To better verify the effectiveness of installing PST, we compare the results of installing PST with those of installing new transmission lines. Grids of three planning stages in Section 7.2 are used to find the respective effects of adding PST and new transmission lines on wind power consumption. The details of each case are as follows.

The base case: No PST or new transmission lines are installed in the grids. The case of installing PST: The PST allocation schemes of each planning stage refer to the results of $Road_1$ in Table 2.

The case of transmission network expansion: New transmission lines are added to the heavy load lines of the wind power transmission section, which is line 8-30 in the first planning stage, lines 8-30, 30-38 in the second planning stage and lines 8-30, 30-38, 91-92 in the third planning stage. The new transmission line parameters are consistent with the original line parameters for simplicity.

It is worth mentioning that the number of PSTs added and new transmission lines in each planning year are equal.

Table 5 presents the results of the three cases. Columns 2–4 indicate the wind power consumption of each planning stage. The fifth column gives the total wind power consumption in three planning stages. The wind power consumption with the installation of PST is lower in the first planning stage compared to the transmission network expansion, but higher in the second and third planning stages. In terms of total wind power consumption rate, the effect of installing PST is also slightly better than adding new lines. In addition, transmission grid expansion usually requires higher investment costs, longer construction time and more stringent environmental approvals than installing PST in the real grid.

In conclusion, when the wind power transmission section of the grid has overload problems and the section has exploitable transmission potential, installing PST is a better transition option during the long construction cycle of the grid.

7.5 Validation of the limitations on the number of PST

In order to verify the point in Section 4.3.3, the 118-bus system in Section 7.1 is used for simulation. The wind farm with the maximum capacity of 1600 MW is located in bus 5. The wind power transmission section consists of lines 5-3, 5-4, 5-11, 5-6, and 8-30, where line 8-30 is the heaviest loaded line of the transmission section.

Table 6 shows the results for various numbers of PSTs. It can be seen that the bigger the number of PSTs, the higher the wind power consumption and the larger the TTC of the transmission section. However, the increment of the wind power consumption and the TTC is not proportional to the number of PSTs. In conclusion, excessive PSTs could inevitably lose the whole economic efficiency.

TABLE 6 IEEE 118-bus system results for various numbers of PSTs.

Number of PSTs	PST locations	Wind power consumption/% (increase rate)	TTC/MW (increase rate)
None	—	39.63 (/)	907.6 (/)
1	8-30	50.84 (28.28%)	1,047.7 (15.44%)
2	8-30, 5-3	56.44 (42.42%)	1,117.6 (23.14%)
3	8-30, 5-3, 4-11	58.45 (47.47%)	1,142.6 (25.89%)

TABLE 7 Computational comparison for different computational models.

	TTC (MW)	Computational error (%)	Computation time (s)
Model1	342.43	1.28	0.042
Model2	338.0	0.03	77.976
Model3	338.1	—	3833.218

7.6 Computational issues

All the simulations are conducted on a test computer with an Intel(R) Xeon(R) W-2255 CPU @ 3.70 GHz and 128.00 GB of RAM.

To demonstrate the computation accuracy of the proposed method for solving TTC in Section 4.1.2, we selected a typical scenario to compare the computational results in three computational models.

Model 1. Use the sensitivity method in 3.1.2 without considering the correction by RPF.

Model 2. Use the result of the model1 as the initial point for RPF, and a subsequent correction is applied;

Model 3. Use RPF to calculate TTC directly.

We consider the results obtained by RPF (model3) as the standard value and calculate the error of the results in model1 and model2. As observed in Table 7, the error of the results obtained by the sensitivity method is 1.28% and reduces to 0.03% after the correction by RPF.

In terms of computational speed, the computation time of the sensitivity method is extremely short. The calculation time of RPF is very long for two main reasons. First, not only the generator output but also the phase shift angle need to be adjusted, so the computation work became greater compared to the conventional RPF. Second, small adjustment steps for each generator and phase shift angle have been set to get accurate results. The proposed method (model2) obtained more accurate results after the correction by RPF, and the computation time was greatly reduced compared to model3. Since the initial point obtained based on the sensitivity method is closer to the final point.

8 Conclusion

This paper presents a dynamic programming method for optimal PST allocation, considering the evolution and wind power uncertainty of the grid. The proposed approach seeks to identify the optimal planning path of PST in multi-stage grid planning. To demonstrate the effectiveness of the proposed method, we compare the results under dynamic and static programming. The method is also applied to the model of a published manuscript, and the results prove the superiority of the proposed method. In addition, we present a calculation method for TTC of grids containing PST and wind power, for TTC is used as one of the optimization objectives.

The current research is aimed to optimize the planning path of PST by considering the evolution of the grid. However, the addition of other FACTS devices is not considered. Therefore, the main topics in future works are how the PST cooperates with other FATCS devices and the dynamic programming of multi-type FACTS.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

Author contributions

Investigation, ZL; methodology, KF and ZD; project administration, FL; supervision, CX; writing—original draft, KF; writing—review and editing, FL and ZD.

Funding

This research was funded by the Southern Power Grid Corporation's Science and Technology Project (Project No. 037700KK52190015 (GDKJXM20198313)) and Key-Area Research and Development Program of Guangdong Province (2019B111109001).

Conflict of interest

FL and ZL was employed by The Grid Planning and Research Center of Guangdong Power Grid Corporation.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The authors declare that this study received funding from the Southern Power Grid Corporation's Science and Technology Project. The funder had the following involvement in the study: Investigation, project administration, writing—review.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fenrg.2022.1003315/full#supplementary-material

References

Bandara, H. M. M. T., Samarasinghe, D. P., Manchanayake, S. M. A. M., Perera, L. P. J., Kumaradasa, K. C., Pemadasa, N., and Samarasinghe, A.P. (2019). "Analyzing payment behaviors and introducing an optimal credit limit," in Proceedings of the 2019 International Conference on Advancements in Computing (ICAC), Malabe, Sri Lanka, 05-07 December 2019 (IEEE), 68–72. doi:10.1109/ICAC49085.2019.9103404

Baringo, L., and Conejo, A. J. (2013). Correlated wind-power production and electric load scenarios for investment decisions. *Appl. Energy* 101, 475–482. doi:10.1016/j.apenergy.2012.06.002

Blumsack, S. (2006). Network topologies and transmission investment under electric-industry restructuring. Pittsburgh, PA, USA: Carnegie Mellon Univ. Thesis.

Brilinskii, A. S., Badura, M. A., Evdokunin, G. A., Chudny, V. S., and Mingazov, R. I. (2020). "Phase-Shifting transformer application for dynamic stability enhancement of electric power stations generators," in Proceedings of the 2020 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus), St. Petersburg and Moscow, Russia, 27-30 January 2020 (IEEE), 1176–1178. doi:10.1109/EIConRus49466.2020.9039410

Ding, T., Bo, R., Bie, Z., and Wang, X. (2017). Optimal selection of phase shifting transformer adjustment in optimal power flow. *IEEE Trans. Power Syst.* 32 (3), 2464–2465. doi:10.1109/TPWRS.2016.2600098

Gerbex, S., Stephane, R., and Cherkaoui, A. J. (2001). Optimal location of Multi-Type FACTS devices in a power system by means of genetic algorithms. *IEEE Trans. Power Syst.* 16 (3), 537–544. doi:10.1109/59.932292

Ghahremani, E., and Kamwa, I. (2013). Optimal placement of multiple-type FACTS devices to maximize power system loadability using a generic graphical user interface. *IEEE Trans. Power Syst.* 28 (2), 764–778. doi:10.1109/TPWRS.2012.2210253

Hadzimuratovic, S., and Fickert, L. (2018). "Determination of critical factors for optimal positioning of Phase-Shift Transformers in interconnected systems," in Proceedings of the 2018 19th International Scientific Conference on Electric Power Engineering (EPE), Brno, Czech Republic, 16-18 May 2018 (IEEE), 1–6. doi:10.1109/EPE.2018.8396033

IEA (2017). Getting wind and sun onto the grid. AvaliableAt: https://www.iea.org/reports/getting-wind-and-solar-onto-the-grid.

Ippolito, L., and Siano, P. (2004). Selection of optimal number and location of thyristor-controlled phase shifters using genetic based algorithms. *IEE Proc. Gener. Transm. Distrib.* 151 (5), 630–637. doi:10.1049/ip-gtd:20040800

Kazemi, A., and Sharifi, R. (2006). "Optimal location of thyristor controlled phase shifter in restructured power systems by congestion management," in Proceedings of the 2006 IEEE International Conference on Industrial Technology, Mumbai, India, 15-17 December 2006 (IEEE), 294–298. doi:10.1109/ICIT.2006.372211

Li, F., Li, Z., Yu, M., Liu, R., Fu, K., and Du, Z. (2022a). Research on optimal location method of phase shifter based on improved PageRank algorithm. *Guangdong Electr. Power* 35 (5), 42–52. doi:10.3969/j.issn.1007-290X.2022.005.006

Li, Z., Li, F., Liu, R., Yu, M., Chen, Z., Xie, Z., and Du, Z. (2022b). A Data-Driven genetic algorithm for power flow optimization in the power system with phase shifting transformer. *Front. Energy Res.* 9, 793686. doi:10.3389/fenrg.2021.793686

Lima, F. G. M., Galiana, F. D., Kockar, I., and Munoz, J. (2003). Phase shifter placement in large-scale systems via mixed integer linear programming. *IEEE Trans. Power Syst.* 18 (3), 1029–1034. doi:10.1109/TPWRS.2003.814858

Miranda, V., and Alves, R. (2014). "PAR/PST location and sizing in power grids with wind power uncertainty," in Proceedings of the 2014 International Conference on Probabilistic Methods Applied to Power Systems (PMAPS), Durham, UK, 07-10 July 2014 (IEEE), 1–6. doi:10.1109/PMAPS.2014.6960679

NEA (2022). NEA's first quarter of 2022 online press conferences. AvaliableAt: http://www.nea.gov.cn/2022-01/28/c_1310445390.htm.

National Energy Renewable Laboratory (2019). Wind data-site 126685. AvaliableAt: https://egriddata.org/dataset/wind-data-site-126685.

Preedavichit, P., and Srivastava, S. C. (1998). Optimal reactive power dispatch considering FACTS devices. *Electr. Power Syst. Res.* 46 (3), 251–257. doi:10.1016/S0378-7796(98)00075-3

Sebaa, K., Bouhedda, M., Tlemçani, A., and Henini, N. (2014). Location and tuning of TCPSTs and SVCs based on optimal power flow and an improved cross-entropy approach. *Int. J. Electr. Power & Energy Syst.* 54, 536–545. doi:10.1016/j.ijepes.2013.08.002

Tao, H., Xu, J., and Zou, W. (2013). Model conversion from BPA to PSCAD. *Electr. Power Autom. Equip.* 33 (8), 152–156. doi:10.3969/j.issn.1006-6047.2013.08.026

Verboomen, J., Van Hertem, D., Schavemaker, P. H., Kling, W. L., and Belmans, R. (2008). Analytical approach to grid operation with phase shifting transformers. *IEEE Trans. Power Syst.* 23 (1), 41–46. doi:10.1109/TPWRS.2007.913197

Wang, X., Wang, X., Sheng, H., and Lin, X. (2021). A Data-Driven sparse polynomial chaos expansion method to assess probabilistic total transfer capability for power systems with renewables. *IEEE Trans. Power Syst.* 36 (3), 2573–2583. doi:10.1109/TPWRS.2020.3034520

Wu, Q. H., Lu, Z., Li, M. S., and Ji, T. Y. (2008). "Optimal placement of FACTS devices by a group search optimizer with multiple producer," in Proceedings of the 2008 IEEE Congress on Evolutionary Computation (IEEE World Congress on Computational Intelligence), Hong Kong, China, 01-06 June 2008 (IEEE), 1033–1039. doi:10.1109/CEC.2008.4630923

Yeo, J. H., Dehghanian, P., and Overbye, T. (2019). "Power flow consideration of impedance correction for phase shifting transformers," in Proceedings of the 2019 IEEE Texas Power and Energy Conference (TPEC), College Station, TX, USA, 07-08 February 2019 (IEEE), 1–6. doi:10.1109/TPEC.2019.8662150

Zhang, H., Yang, J., Ren, X., Wu, Q., Zhou, D., and Elahi, E. (2020). How to accommodate curtailed wind power: A comparative analysis between the us, Germany, India and China. *Energy Strategy Rev.* 32, 100538. doi:10.1016/j.esr.2020.100538

Zhang, N., Zhu, X., and Liu, J. (2021). "Improving the consumption capacity of wind power in distributed network using a thyristor controlled phase shifting transformer," in Proceedings of the 2021 3rd Asia Energy and Electrical Engineering Symposium (AEEES), Chengdu, China, 26-29 March 2021 (IEEE), 80–84. doi:10.1109/AEEES51875.2021.9403143

Zhang, X., and Grijalva, S. (2013). "Multi-area ATC evaluation based on Kron reduction," in Proceedings of the 2013 IEEE International Conference on Smart Energy Grid Engineering (SEGE), Oshawa, Canada, 28-30 August 2013 (IEEE), 1–6. doi:10.1109/SEGE.2013.6707921

Zhang, X., Shi, D., Wang, Z., Yu, Z., Wang, X., Bian, D., and Tomsovic, K. (2017). "Bilevel optimization based transmission expansion planning considering phase shifting transformer)," in Proceedings of the 2017 North American Power Symposium (NAPS), Morgantown, WV, USA, 17-19 September 2017 (IEEE), 1–6. doi:10.1109/NAPS.2017.8107289

Zhang, X., Shi, D., Wang, Z., Zeng, B., Wang, X., Tomsovic, K., and Jin, Y. (2018). Optimal allocation of series FACTS devices under high penetration of wind power within a market environment. *IEEE Trans. Power Syst.* 33 (6), 6206–6217. doi:10.1109/TPWRS.2018.2834502

Ziaee, O., and Choobineh, F. (2017). Optimal Location-Allocation of TCSCs and transmission switch placement under high penetration of wind power. *IEEE Trans. Power Syst.* 32 (4), 3006–3014. doi:10.1109/TPWRS.2016.2628053

frontiers | Frontiers in Energy Research

TYPE Original Research
PUBLISHED 18 January 2023
DOI 10.3389/fenrg.2022.1088096

OPEN ACCESS

EDITED BY
Youbo Liu,
Sichuan University, China

REVIEWED BY
Wei Yang,
Southwest Petroleum University, China
Xia Lei,
Xihua University, China
Zhifang Yang,
Chongqing University, China

*CORRESPONDENCE
Zhaobin Wei,
xiaobiner6@163.com

SPECIALTY SECTION
This article was submitted to Smart Grids,
a section of the journal
Frontiers in Energy Research

RECEIVED 03 November 2022
ACCEPTED 15 November 2022
PUBLISHED 18 January 2023

CITATION
Zhang D, Wang Y, Xi L, Deng F, Deng Z, Liu J and Wei Z (2023), Low-Carbon and economic flexibility scheduling of power system with multiple generation resources penetration.
Front. Energy Res. 10:1088096.
doi: 10.3389/fenrg.2022.1088096

Low-Carbon and economic flexibility scheduling of power system with multiple generation resources penetration

Dawei Zhang[1], Yanfeng Wang[1], Litang Xi[1], Fuyao Deng[1], Zhisen Deng[1], Jichun Liu[2] and Zhaobin Wei[2]*

[1]State Grid Sichuan Electric Power Company, Chengdu, China, [2]College of Electrical Engineering, Sichuan University, Chengdu, China

The operation flexibility of the power system suffers great challenges due to the vigorously developing of renewable energy resources under the promotion of the carbon neutralization goal. To this end, this paper proposes an economical and flexible energy scheduling method for power system integrated with multiple generation resources while considering the operation of low-carbon. Specifically, flexibility evaluation indexes are constructed to describe the characteristics of the flexible generation units. Then they are connected with the flexibility of the power system in an economic and low-carbon flexible energy scheduling model. To coordinate the operation economy, flexibility, and carbon emission reduction, the model incorporates demand response, operational characteristics, and flexibility requirements. Further, the model is fully validated through the simulation on the modified IEEE 30-bus system. Results demonstrate that: the proposed method can reduce the system's carbon emission and total operating costs and promote photovoltaic consumption.

KEYWORDS

carbon emission, economic flexible energy scheduling, multiple generation resources, demand response, photovoltaic consumption

1 Introduction

In response to the carbon emission reduction goal proposed worldwide, the hydro-thermal-solar-gas multi-source system (HTSGS) has been significantly promoted for its advantage of renewable energy substitution and high electricity density (Buhan et al., 2020; Zhang et al., 2022). However, the operational flexibility of HTSGS is limited and challenged due to the randomness and intermittency of renewable energy (Du et al., 2019; Li et al., 2022). To handle this issue, it is significant to enhance the flexibility of HTSGS while guaranteeing its low-carbon and economic operation under the penetration of renewable energy.

The flexibility of the power system is supposed to be greatly enhanced to cope with the strong randomness and volatility of renewable energy resources. However, there are

presently no unified flexibility definitions. To this end, several opinions on flexibility have been put forward. Heggarty et al. (Heggarty et al., 2020) defined flexibility as the power system's ability to cope with variability and uncertainty. Yamujala et al. (Heggarty et al., 2021) proposed that flexibility is the energy, power, and ramp capability of a system to modify generation and demand in response to load variations at minimum cost. Emmanuel et al. (Emmanuel et al., 2020) defined flexibility as the ability of the power system to respond adequately to dynamic grid conditions at various timescales while operating at minimal cost within institutional frameworks and market designs. Ma et al. (Ma et al., 2013) described flexibility as the ability of a power system to cope with variability and uncertainty in both generation and demand. However, the aforementioned flexibility definitions were not combined with the operating characteristics of the generation units with fast response capability, i.e., the operation flexibility cannot be fully mobilized.

As the necessary description of power system flexibility, corresponding indexes are essential to guide the inflexibility-oriented operation. Until now, some efforts on flexibility indexes have been carried out. In (Lu et al., 2018), loss of flexibility probability, loss of flexibility duration, loss of flexibility expectation, and flexibility demand shortage were used to describe flexibility with renewable power curtailment. Flexibility with high penetration of renewables was characterized by four indexes: ramping limit, power capacity, energy capacity, and response time in (Mohandes et al., 2019). In (Brahma and Senroy, 2020), the flexibility index focusing on small-signal stability was developed. Nevertheless, this work suffers great limitations in the case where a large variety of renewable energy occurs in the power system, due to the failure of the linearization model. In low-carbon power systems, an index based on operating range and ramping was utilized to quantify operational flexibility (Yamujala et al., 2021). Considering the transmission capacity and energy conversion constraints, the flexibility margin index was proposed to evaluate the flexibility from the aspect of the acceptable wind power fluctuations range (Zhao et al., 2021). According to the different research objects concerned by scholars, the indexes designed to evaluate power system flexibility are also various. However, most of these flexibility indexes are restrictive and only applicable to a specific scenario; they fail to describe the fast-ramping capacity, such as that of cascade hydropower unit (CHU) and gas unit (GU) integrated into HTSGS, nor cover its flexibility adjustment advantage *via* the multi-source complementation.

Multiple operation scheduling strategies have been developed to improve the power system flexibility according to the guidance of specific flexibility indexes. Specifically, in photovoltaic (PV) embedded microgrid, Zhao and Xu (Zhao and Xu, 2017) proposed a two-stage ramp-limited optimal scheduling strategy considering flexibility check. Next, Considering large-scale wind power integration, Li et al. (Li et al., 2018) proposed a big-M method-based co-optimized scheduling model including the flexible capacity of automatic generation control units. However, the flexible capacity is only provided by conventional thermal power unit (TPU), which is not environmentally friendly nowadays. Shi et al. (Shi et al., 2021) formulated the real-time economic scheduling problem as a multi-stage robust program to leverage flexible resources in a broader timescale. Wang et al. (Wang et al., 2021) presented a model that considers three states of flexible resource-pumped storage hydro in look-ahead scheduling. The framework proposed in (Fan et al., 2022) used robust optimization to measure the system flexibility and considered the interaction between economic scheduling and automatic generation control. Despite the progress of the above works, most of them focus on the scheduling at the generation side to boost the system flexibility, while ignoring the fact that the source-load interaction can performs better in the flexibility enhancement. Additionally, these works above did not involve the initially vital objectives, i.e., carbon emission reduction, and operation economy, which may make them not feasible in actual application.

As mentioned above, carbon emission reduction is another important objective in the power system scheduling besides flexibility since the system also undertakes important responsibility in the decarbonization trend to cope with global warming. For the sake of this, some works have been carried out in the low-carbon scheduling of power system. To be more specific, a scheduling strategy based on carbon capture and fuel cost was proposed, and the modeling and analysis of a carbon capture technology was discussed to reduce carbon emission and generation cost (Reddy et al., 2017). Moreover, the economic-emission scheduling of combined renewable and coal power plants equipped with carbon capture systems was addressed (Akbari-Dibavar et al., 2021). In (Ma et al., 2015), they found out that DR could help to accommodate renewable energy, and economic and low-carbon day-ahead scheduling was addressed. A low-carbon optimal scheduling model with demand response carbon intensity control was also proposed in (Wang et al., 2022), which effectively reduced carbon emission. However, these scheduling strategies above simply aimed to gain more economic benefit and reduce carbon emission, where the flexibility is not considered. Note that carbon reduction and flexibility can interact through the resource allocation integrated into power system. To deal with this, more efforts should be put into practice.

Given the limitation, this paper focuses on the flexibility of HTSGS and carbon emission reduction, where the coordination of DR and the operating characteristics of the generation units are fully developed. The main contributions of this paper can be summarized as follows:

1) A low-carbon and economic flexibility scheduling strategy for HTSGS is proposed while considering the system flexibility cost, carbon emission cost, and operational cost. The

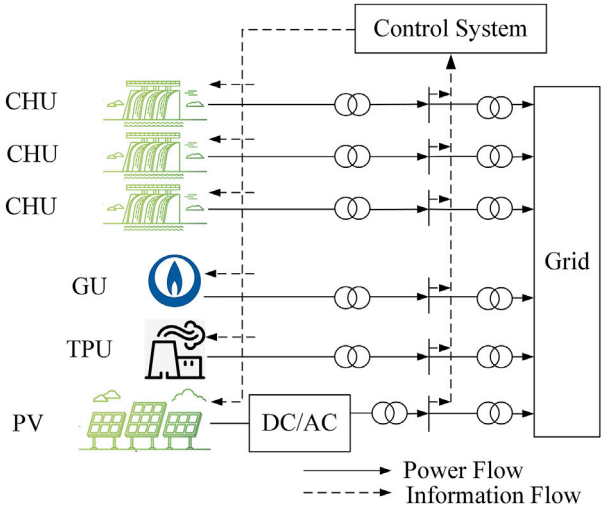


FIGURE 1
HTSGS structure diagram.

scheduling strategy can improve the adaptability of HTSGS in the low-carbon environment and satisfy the flexibility challenge brought by renewable energy.

2) This paper particularly constructs flexibility indexes for the HTSGS at the system level, including upward insufficient ramping resource probability (UIRRP), downward insufficient ramping resource probability (DIRRP), upward sufficient ramping resource expectation (USRRE), and downward sufficient ramping resource expectation (DSRRE). Besides, these flexibility indexes are also combined with the operating characteristics of generation units (i.e, the upward and downward flexibility supply of CHU and GU). This kind of integrated flexibility manner enables to provide an accurate operation direction for HTSGS and fully incentives the internal flexible resources.
3) To reveal the impact on system flexibility and economy, multiple comparison cases, including the power supply composition and carbon prices, are simulated, which can provide data support for the actual operation of HTSGS.

The remainder of this paper is organized as follows. **Section 2** introduces the flexible resource evaluation index of HTSGS. **Section 3** proposes a multi-objective optimal scheduling strategy considering the power system with HTSGS. **Section 4** studies the model and method proposed in this paper through multiple scenarios, then compares and analyzes the scheduling results under diverse system flexibility costs, carbon transaction cost, DR, and carbon emission price. The impact of different power source structures in HTSGS on system flexibility is also discussed. **Section 5** concludes the paper.

2 Modeling for power system with HTSGS

A variety of controllable power generation units are integrated into the HTSGS. The overall system flexibility can be improved through reasonable resource allocation, which is a necessary way to achieve a low-carbon power system. The HTSGS is mainly composed of CHU, PV, TPU, and GU. The system structure is shown in Figure 1.

As shown in Figure 1, the CHUs consist of multiple cascade hydropower stations, which are connected to the power grid through transformers together with TPU and GU; In the PV power station, the PV array is connected to the power grid through the DC/AC inverter; The power generation control system collects and analyzes source data in real-time, and regulates the multiple sources. The power generation control system makes sure HTSGS can provide sufficient power side flexibility resources.

2.1 Model of flexible resources

2.1.1 Flexibility model of CHU

Compared with conventional TPU, CHU has the advantages of wide-range regulation and low cost. CHU is a good flexible resource. The maximum upward flexibility supply *FSRup CHU*(σ,t) and downward flexibility supply *FSRdown CHU*(σ,t) are provided by CHU as follows:

$$FSR_{CHU}^{up}(\sigma,t) = \min\left(RU\sigma, P_{CHU}^{max}(t) - P_{CHU}(t)\right) \quad (1)$$

$$FSR_{CHU}^{down}(\sigma,t) = \min\left(RU\sigma, P_{CHU}(t) - P_{CHU}^{min}(t)\right) \quad (2)$$

where σ is the time scale unit; RU and RD are the maximum and the minimum upward climbing rate of CHU; Pmax *CHU* and Pmin *CHU* are the maximum and minimum output of CHU at time t; $P_{CHU}(t)$ is the actual output of CHU at time t.

2.1.2 Flexibility model of GU

The GU has the advantages of high-power generation efficiency and low pollution generation rate when operating under a high load rate. GU can also be used as a flexible resource to adjust the system operation performance.

The flexibility that GU can provide is defined as the maximum upward flexibility supply *FSRup GU* (σ,t) and downward supply *FSRdown GU* (σ,t):

$$FSR_{GU}^{up}(\sigma,t) = \min\left(RU_g\sigma, P_{GU}^{max}(t) - P_{GU}(t)\right) \quad (3)$$

$$FSR_{GU}^{down}(\sigma,t) = \min\left(RD_g\sigma, P_{GU}(t) - P_{GU}^{\min}(t)\right) \quad (4)$$

where σ is the time scale unit; RU_g and RD_g are the maximum and the minimum upward climbing rate of GU; Pmax *GU* and Pmin *GU* are the maximum and minimum output of GU at time t; $P_{GU}(t)$ is the actual output of GU at time t.

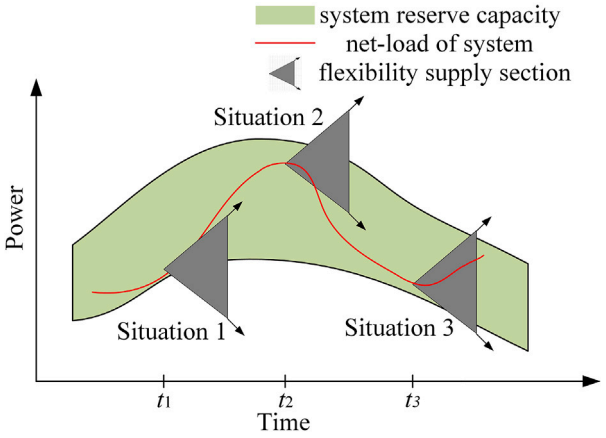


FIGURE 2
Schematic diagram of HTSGS flexibility evaluation.

2.2 Model of flexible indexes

As shown in Figure 2, the green interval is the system reserve capacity range, the red curve is the system load curve, and the gray triangle with arrows represents the flexible scheduling output of the power system with HTSGS integration. Based on this, situation one shows that its gray triangle representing the system's flexibility resources can only meet the lower limit of the system reserve capacity but cannot meet the upper limit, which means that the system is unable to balance the upward fluctuating load and the system has insufficient upward flexibility resources. Similarly, Situation two represents insufficient resources for the downward flexibility of the system. Situation three indicates that the system has sufficient upward and downward flexibility resources to balance the upward or downward fluctuation load. In this paper, the corresponding index model is developed to analyze the upward and downward flexibility resources.

2.2.1 Upward insufficient ramping resource probability

Upward insufficient ramping resource probability (UIRRP) $\eta_{UIRRP}(t)$ is the probability that the flexibility provided by CHU and GU cannot meet the upward flexibility resource demand at time t when the power system with HTSGS integration operating, as shown in Eq. 5.

$$\begin{cases} \eta_{UIRRP}(t) = \Pr FSR^{up}(t) - ((P_{NL}(t+1) - (P_{NL}(t)) < 0\} \\ FSR^{up}(t) = FSR^{up}_{CHU}(t) + FSR^{up}_{PSU}(t) \end{cases} \quad (5)$$

where $FSR^{up}(t)$ is the sum of the maximum upward flexibility supply provided by the system; $P_{NL}(t)$ is the net-load of the system.

2.2.2 Downward insufficient ramping resource probability

Downward insufficient ramping resource probability (DIRRP) $\eta_{DIRRP}(t)$ is the probability that the flexibility provided by CHU and GU cannot meet the downward flexibility resource demand at time t when the power system with HTSGS integration operating, as shown in Eq. 6.

$$\begin{cases} \eta_{DIRRP}(t) = \Pr\{FSR^{down}(t) - ((P_{NL}(t) - (P_{NL}(t+1)) < 0\} \\ FSR^{down}(t) = FSR^{down}_{CHU}(t) + FSR^{down}_{PSU}(t) \end{cases} \quad (6)$$

where $FSR^{down}(t)$ is the sum of the maximum downward flexibility supply provided by the system.

2.2.3 Upward sufficient ramping resource expectation

Upward sufficient ramping resource expectation (USRRE) $E_{USRRE}(t)$ is the expectation of the excess $FSRM^{up}(t)$ that the upward flexibility resources provided by the HTSGS exceed the upward flexibility resources required by the system at time t during the power system operation, as shown in Eq. 7.

$$\begin{cases} E_{USRRE}(t) = FSRM^{up}(t) \cdot \Pr\{FSR^{up}(t) \\ \qquad -(P_{NL}(t+1) - P_{NL}(t)) > 0\} \\ FSRM^{up}(t) = FSR^{up}(t) - (P_{NL}(t+1) - P_{NL}(t)) \end{cases} \quad (7)$$

2.2.4 Downward sufficient ramping resource expectation

Downward sufficient ramping resource expectation (DSRRE) $E_{DSRRE}(t)$ is the expectation of the excess $FSRM^{down}(t)$ that the downward flexibility resources provided by the HTSGS exceed the downward flexibility resources required by the system at time t during the power system operation, as shown in Eq. 8.

$$\begin{cases} E_{DSRRE}(t) = FSRM^{down}(t) \cdot \Pr\{FSR^{down}(t) \\ \qquad -(P_{NL}(t) - P_{NL}(t-1)) > 0\} \\ FSRM^{down}(t) = FSR^{down}(t) - (P_{NL}(t) - P_{NL}(t+1)) \end{cases} \quad (8)$$

2.3 Modeling of DR

DR means that users can change their electricity consumption behavior through the change in electricity price, which can improve the power system operational flexibility. The elasticity coefficient of demand price indicates the relationship between the electricity load change and the electricity price change. The elasticity coefficient can be calculated as follows:

$$\mu_{nm} = \frac{\Delta L_n / \lambda_L L_{n0}}{\Delta Q_m / Q_{m0}} \begin{cases} u_{nm} \le 0, & n = m \\ u_{nm} > 0, & n \ne m \end{cases} \quad (9)$$

where ΔL_{n0} and ΔQ_m are the electric load change and electricity price after DR; L_{n0} and Q_m are load and electricity price before DR; λ_L is the adjustable load proportion in the electric load.

The change of electric load after DR within t periods can be expressed as:

$$\begin{bmatrix} \frac{\Delta L_1}{L_1} \\ \frac{\Delta L_2}{L_2} \\ \vdots \\ \frac{\Delta L_t}{L_t} \end{bmatrix} = \mu \begin{bmatrix} \frac{\Delta Q_1}{Q_1} \\ \frac{\Delta Q_2}{Q_2} \\ \vdots \\ \frac{\Delta Q_t}{Q_t} \end{bmatrix} \tag{10}$$

$$\mu = \begin{bmatrix} \mu_{11} & \mu_{12} & \cdots & \mu_{1t} \\ \mu_{21} & \mu_{22} & \cdots & \mu_{2t} \\ \vdots & \vdots & \vdots & \vdots \\ \mu_{t1} & \mu_{t2} & \cdots & \mu_{tt} \end{bmatrix} \tag{11}$$

where μ is the price elasticity coefficient matrix of DR.

The DR model based on electricity price and load is established as follows:

$$L_n = L_{n0}\left(\sum_{m=1}^{t}\mu\frac{\Delta Q_m}{Q_{m0}} + \mu\frac{\Delta Q_n}{Q_{n0}} + 1\right) \tag{12}$$

where L_n is the electric load after DR.

3 Economic scheduling model of power system with HTSGS

The objective function of the traditional economic scheduling model is to minimize the system operation cost. On this basis, this paper comprehensively considers the carbon emission reduction and the system flexibility guarantee and introduces the carbon emission cost and the risk cost of insufficient flexibility. A system economic scheduling model is built, and the multi-dimensional optimal coordinated scheduling of overall system operation economy, flexibility, and carbon emission reduction is realized.

3.1 System operation cost

The annual system operation cost consists of TPU operation cost C_{TPU}, CHU operation cost C_{CHU} and GU operation cost C_{GU}, as shown in Eq. 13.

$$C_r = C_{TPU} + C_{CHU} + C_{GU} \tag{13}$$

The TPU operation cost includes the startup cost and unit cost of power generation, as shown in Eq. 14; CHU operation cost includes startup cost, shutdown cost, and flexibility cost, as shown in Eq. 15; GU operation cost includes the startup cost, shutdown cost, and gas cost, as shown in Eq. 16.

$$C_{TPU} = \sum_{t=1}^{T}\sum_{x=1}^{\theta_{TPU}}\left[c_x^{su}SU_x(t) + \left(a_x + b_xP_x(t) + c_xP_x(t)^2\right)\right] \tag{14}$$

$$C_{CHU} = \sum_{t=1}^{T}\sum_{y=1}^{\theta_{CHU}}\left[c_y^{su}SU_y(t) + c_y^{sd}SD_y(t) + c_y^{op}P_y(t)\right] \tag{15}$$

$$C_{GU} = \sum_{t=1}^{T}\sum_{z=1}^{\theta_{GU}}\left[\left[\begin{array}{l} c_z^{su}SU_z(t) + c_z^{sd}SD_z(t) + \\ \left(a_zP_z(t)^2 + b_zP_z(t) + c_z\right) \end{array}\right]\right] \tag{16}$$

Where T is the total number of scheduling periods; θ_{TPU}, θ_{CHU} and θ_{GU} are the set of TPU, CHU, GU; *csu x* is the startup cost of the *x*th TPU; a_x, b_x and c_x are the power generation cost coefficients of TPU; $SU_x(t)$ is the 0–1 state variable to indicate *x*th TPU startup. When it is 0, the *x*th TPU is not started, and when it is 1, the TPU *x* is in the startup state; $P_x(t)$ is the generating power of the TPU *x* at time *t*; *csu y* and *csd y*are the startup cost and shutdown cost of the *y*th CHU; $SU_y(t)$ and $SD_y(t)$ are 0–1 state variables for the startup and shutdown of the *y*th CHU; $P_y(t)$ is the generating power of the *y*th CHU at time *t*; and *cop y* is the power generation cost coefficient of the *y*th CHU; *csu z*, *csd z* are the startup and shutdown costs of the *z*th GU; $P_z(t)$ is the generating power of the *z*th GU at time *t*; a_z, b_z and c_z are the operating consumption coefficients of *z*th GU.

3.2 Carbon emission cost

The system purchases carbon emission rights in the carbon trading market according to carbon emission quotas and actual carbon emissions. The carbon emission quota allocation based on power generation is adopted in this paper. The carbon emission quota allocated to the system is approximately proportional to the total power generation of the system. The system carbon emission quota is as follows:

$$E_D = \sum_{t=1}^{T}\gamma P_a(t) \tag{17}$$

where E_D is the allocated carbon emission quota; γ is the carbon emission quota coefficient; $P_a(t)$ is the total power generation of all generator units at time *t*.

In this paper, the carbon emissions of TPU and GU are considered. The actual carbon emission is as follows:

$$E_C = \sum_{t=1}^{T}\sum_{n=1}^{\theta_{TU}}\eta_n^{TU}P_n(t) + \sum_{t=1}^{T}\sum_{m=1}^{\theta_{GU}}\eta_m^{GU}P_m(t) \tag{18}$$

where E_C is the actual carbon emission of TPU and GU; $P_n(t)$ and $P_m(t)$ are the output power of the *n*th TPU and the *m*th GU at time *t*; $\eta TU\ n$ and $\eta GU\ m$ are the carbon emission coefficients of the *n*th TPU and the *m*th GU, respectively.

The carbon emission cost of this paper adopts the divided-interval ladder carbon trading model. The higher the carbon emission, the higher the carbon trading price required, as shown in Eq. 19.

$$C_{co} = \begin{cases} \sigma(E_C - E_D) \\ E_C < E_D \\ \sigma(1+\mu)(E_C - E_D) \\ E_D \le E_D < E_D + \omega \\ \sigma(1+\mu)\omega + \sigma(1+\xi\mu)(E_C - E_D) \\ E_D + \omega \le E_C \end{cases} \tag{19}$$

where σ Is the unit price of carbon trading; μ and ξ are the growth coefficients of carbon trading price, where $\xi \geq 2$; ω is the corresponding carbon emission range.

3.3 System flexibility cost

The risk cost of insufficient flexibility includes the risk cost caused by insufficient upward flexibility and the risk cost caused by insufficient downward flexibility. It is determined by the multiplication of the upward and downward flexibility supply and demand difference and the risk cost coefficient of the corresponding lack of flexibility.

The upward flexibility resources shortage $FSRS^{up}(t)$ is the difference between the upward flexibility demand and the actual upward flexibility supply capacity at time t.

$$FSRS^{up}(t) = (P_{NL}(t+1) - P_{NL}(t)) - FSR^{up}(t) \tag{20}$$

Similarly, the downward flexibility resources shortage $FSRS^{down}(t)$ is the difference between the downward flexibility demand and the actual downward flexibility supply capacity at time t.

$$FSRS^{down}(t) = (P_{NL}(t) - P_{NL}(t+1)) - FSR^{down}(t) \tag{21}$$

Thus, the risk cost of insufficient flexibility C_{FSRS} is:

$$C_{FSRS} = \sum_{t\in T} \left(\lambda_{up} FSRS^{up}(t) + \lambda_{down} FSRS^{down}(t)\right) \tag{22}$$

where λ_{up}, λ_{down} are the risk cost coefficients that lack upward and downward flexibility.

3.4 Objective function

The minimum total cost C_a is taken as the objective function, and is composed of system operation cost C_r, carbon emission cost C_{CO} and risk cost of insufficient flexibility C_{FSRS}, as shown in Eq. 23.

$$\min C_a = C_r + C_{co} + C_{FSRS} \tag{23}$$

3.5 Constraint conditions

3.5.1 Power flow constraint

The sum of the output of CHU, GU, and PV at any time during the whole scheduling period in the HTSGS should balance the system load. For any node b, the flexibility requirements for PV and load are as follows:

$$\sum_{x\in b} P_x(t) + \sum_{y\in b} P_y(t) + \sum_{z\in b} P_z(t) + \sum_{l\in L_b} P_l(t) + \sum_{m\in b} P^{UL}_{s,m}(t) \geq \sum_{n\in b} P^{LL}_{d,n}(t) \tag{24}$$

$$\sum_{x\in b} P_x(t) + \sum_{y\in b} P_y(t) + \sum_{z\in b} P_z(t) + \sum_{l\in L_b} P_l(t) + \sum_{m\in b} P^{LL}_{s,m}(t) \leq \sum_{n\in b} P^{UL}_{d,n}(t) \tag{25}$$

where $P_y(t)$ is the power generated by the CHU y at time t; $P_z(t)$ is the power generated by the GU z at time t; l, L_b are the transmission line and the transmission line set of the corresponding node b; $PUL\ s,m(t)$ and $PLL\ s,m(t)$ are the upper and lower limits of the generated power of the PV m; $PUL\ d,n(t)$ and $PLL\ d,n(t)$ are the upper and lower limits of the load n.

Calculations and constraints for the DC power flow of the transmission line are as follows:

$$P_l(t) = \frac{(\theta_b(t) - \theta_{b'}(t))}{X_l} \tag{26}$$

$$P_{l,min} \leq P_l(t) \leq P_{l,max} \tag{27}$$

where $P_l(t)$ is the value of the transmission line power flow; $P_{l,max}$ and $P_{l,min}$ are the maximum and minimum power flow of the transmission line; X_l is the reactance of the transmission line l; θ_b is the phase angle of node b.

3.5.2 Operation constraints of TPU

The output constraint of TPU is as follows:

$$P_{x,\min} \leq P_x(t) \leq P_{x,\max} \tag{28}$$

where $P_{x,\min}$ and $P_{x,\max}$ are the minimum and maximum output of the TPU x.

Ramping rate constraint of TPU is as follows:

$$-RU_x \Delta t \leq P_x(t) - P_x(t-1) \leq RD_x \Delta t \tag{29}$$

where RU_x and RD_x are the upper and lower limits of the ramping rate of TPU x.

3.5.3 Operation constraints of CHU

The operation constraints of CHU include the upper and lower limits of storage capacity, power generation flow and discharge flow, water balance constraints, upper and lower output limits, and ramping constraints.

The upper and lower limits of storage capacity, power generation flow and discharge flow of CHU y are as follows:

$$V_{y,\min} \leq V_y(t) \leq V_{y,\max} \tag{30}$$

$$F^{dc}_{y,\min} \leq F^{dc}_{y}(t) \leq F^{dc}_{y,\max} \tag{31}$$

$$F^{out}_{y,\min} \leq F^{out}_{y}(t) \leq F^{out}_{y,\max} \tag{32}$$

$$F_y^{out}(t) = F_y^{fw}(t) + F_y^{dc}(t) \tag{33}$$

where $V_y(t)$, $V_{y,\min}$ and $V_{y,\max}$ are the real-time, minimum and maximum storage capacity of the CHU y at time t; *Fdc y*(*t*), *Fdc y,* min and *Fdc y,* max are the real-time, minimum and maximum power generation flow of the CHU y at time t; *Ffw y*(*t*) and *Fout y*(*t*) are the waste and discharge water flow at time t; *Fout y,* min and *Fout y,* max are the minimum and maximum discharge water flow.

The water balance constraints of CHU are the change balance of the storage capacity of a single CHU over a certain time period and the change balance of the storage capacity of multiple CHUs in a certain space, as shown in Eq. 34 and Eq. 35.

$$V_y(t+1) = V_y(t) + \Delta\tau\Delta s\left(F_y^{in}(t) - F_y^{dc}(t) - F_y^{fw}(t)\right) \tag{34}$$

$$F_y^{in}(t) = F_y^{n}(t) + \sum_{ur\in\Omega_{URCHU}}\left(F_{ur}^{dc}(t) + F_{ur}^{fw}(t)\right) \tag{35}$$

where *Fin y*(*t*) is the water inflow of the CHU y at time t, Δs is the number of seconds per hour; *Fn y*(*t*) is the natural flow of the CHU y at time t; ur and Ω_{URCHU} are the number and number set of upstream CHUs.

The output of CHU shall be between the upper and lower limits of its output:

$$P_{y,\min} \le P_y(t) \le P_{y,\max} \tag{36}$$

where $P_{y,\min}$ and $P_{y,\max}$ are the minimum and maximum output of CHU y.

The output of CHU is determined by the storage capacity and power generation flow:

$$\begin{aligned} P_y(t) = {} & a_{1,y}\left(V_y(t)\right)^2 + a_{2,y}\left(F_y^{dc}(t)\right)^2 \\ & + a_{3,y}V_y(t)F_y^{dc}(t) + a_{4,y}V_y(t) \\ & + a_{5,y}F_y^{dc}(t) + a_{6,y} \end{aligned} \tag{37}$$

where $a_{1,y}$, $a_{2,y}$, $a_{3,y}$, $a_{4,y}$, $a_{5,y}$, $a_{6,y}$ are the calculation coefficients of CHU output.

The difference between CHU output at time t and t+1 shall not exceed the maximum ramping output:

$$\begin{cases} P_y(t+1) - P_y(t) \le \left(1 - SU_y(t)\right)RU_y \\ P_y(t) - P_y(t+1) \le \left(1 - SD_y(t)\right)RD_y \end{cases} \tag{38}$$

where RU_y and RD_y are the maximum upward ramping output and maximum downward climbing output.

3.5.4 Operation constraints of GU

Operation constraints of GU include startup and shutdown constraints, output constraints, gas control constraints, and ramping rate constraints.

The increased times of GU startup and shutdown will shorten the GU's service life and increase the startup and shutdown cost, so the upper limit of startup and shutdown times is taken as the operation constraint of GU.

$$\sum_{t=1}^{T}\left[SU_z(t)\cdot SD_z(t-1)\right] \le N_z \tag{39}$$

where N_z is the upper limit of the startup times of the GU z.

The output constraint of GU is as follows:

$$P_{z,\min} \le P_z(t) \le P_{z,\max} \tag{40}$$

where $P_{z,\min}$ and $P_{z,\max}$ are the minimum and maximum output of the GU z.

In reality, the natural gas supply is still insufficient, so the daily power generation of GU is set as a constraint in scheduling as follows:

$$\sum_{t=1}^{T}P_z(t)\Delta t = E_z \tag{41}$$

where E_z is the daily power generation of GU z.

Ramping rate constraint of TPU is as follows:

$$-RU_z\Delta t \le P_z(t) - P_z(t-1) \le RD_z\Delta t \tag{42}$$

where RU_z and RD_z are the upper and lower limits of the ramping rate of GU z.

3.5.5 Constraints of PV

The amount of radiation is the critical factor to determine the PV output power. The prediction error of radiation ΔR can be expressed as a normally distributed random variable with a mean of 0 and a standard deviation of σ_R. If the predicted value of radiation is R_P, the probability density function of the actual radiation $R = R_P+\Delta R$ is:

$$f_R(R) = \frac{1}{\sqrt{2\pi}\sigma_R}e^{-(R-R_P)^2/2\sigma_R^2} \tag{43}$$

where the variance of the prediction error of the radiation is set to m% of the predicted value R_P.

The PV output power and radiation meet the relationships as follows:

$$P_{PV} = \lambda_{PV}RS_{PV} \tag{44}$$

where λ_{PV} is the conversion efficiency of PV power generation; S_{PV} is the total area of the PV array.

The PV output shall not exceed its upper limit.

$$0 < P_{PV} \le P_{PV,\max} \tag{45}$$

where $P_{PV,\max}$ is the maximum PV output.

3.5.6 Operation constraints of DR

Users participating in DR determine their own power load use according to the electricity price information. The total load used in the whole scheduling period remains unchanged and only changes in the power consumption time, that is, the change in the total load is 0, as shown in Eq. 46.

$$\sum_{t=1}^{T}\Delta L_m = 0 \tag{46}$$

After DR, the power consumption cost of users shall not exceed the power consumption cost before DR, as shown in Eq. 47.

$$\sum_{t=1}^{T}Q_n L_m \le \sum_{t=1}^{T}Q_{n0}L_{m0} \tag{47}$$

where Q_{n0} and Q_n are the electricity price before and after DR.

3.5.7 Operational flexibility constraints

The upward and downward flexibility supply required by the power system with HTSGS integration is provided by the TPU, CHU, and GU, as shown in Eq. 48.

$$\begin{cases} FSR^{up}(t) = \sum\limits_{x\epsilon\theta_{TPU}} FSR_x^{up}(t) + \sum\limits_{y\epsilon\theta_{CHU}} FSR_y^{up} \\ \qquad + \sum\limits_{z\epsilon\theta_{GU}} FSR_z^{up}(t) \\ FSR^{down}(t) = \sum\limits_{x\epsilon\theta_{TH}} FSR_x^{down}(t) + \sum\limits_{y\epsilon\theta_{CHU}} FSR_y^{down} \\ \qquad + \sum\limits_{z\epsilon\theta_{GU}} FSR_z^{down}(t) \end{cases} \tag{48}$$

The flexibility supply provided by the TPU x at time t is constrained by the ramping rate and the upper and lower limits of output, as shown in Eq. 49.

$$\begin{cases} FSR_x^{up}(t) \le \min\left(RU_x, P_{x,\max} - P_x(t)\right) \\ FSR_x^{down}(t) \le \min\left(RD_x, P_x(t) - P_{x,\min}\right) \end{cases} \tag{49}$$

The flexibility supply provided by the CHU y at time t is also constrained by the ramping rate and the upper and lower limits of output, as shown in Eq. 50.

$$\begin{cases} FSR_y^{up}(t) \le \min\left(RU_y, P_{y,\max} - P_y(t)\right) \\ FSR_y^{down}(t) \le \min\left(RD_y, P_y(t) - P_{y,\min}\right) \end{cases} \tag{50}$$

The flexibility supply provided by the GU z at time t is also constrained by the ramping rate and the upper and lower limits of output, as shown in Eq. 51.

$$\begin{cases} FSR_z^{up}(t) \le \min\left(RU_z, P_{z,\max} - P_z(t)\right) \\ FSR_z^{down}(t) \le \min\left(RD_z, P_z(t) - P_{z,\min}\right) \end{cases} \tag{51}$$

4 Case studies

4.1 Parameters setting

The proposed model comprehensively considers the optimal utility of system flexibility index, carbon trading, and DR to the power system with HTSGS integration. The dispatcher can adjust the source and load on both sides according to the day-ahead load forecast results to minimize the total operation cost of the system. Then, a modified IEEE 30-bus system is used for case studies. The system includes 1 PV, 2 TPUs, 1 GU, and 3 CHUs. The location and technical parameters of each generator unit are shown in Table 1 and Figure 3. The model in this paper is a mixed integer linear programming model, which is globally optimized by calling the Yalmip/Gurobi commercial solver on Windows 7 computer (3.2 Ghz, 8GB, 4-core) under the MATLAB 2016 platform. The solution time is 24.73s, meeting the scheduling time requirement.

To illustrate the advantages of the model proposed in this paper, four comparison scenarios are set as follows:

Scenario 1: system flexibility cost, carbon trading cost, and DR are considered at the same time;

Scenario 2: only the system carbon trading cost and DR are considered;

Scenario 3: only the system flexibility cost and carbon trading cost are considered;

Scenario 4: only the system flexibility cost and DR are considered.

4.2 Simulation results analysis

4.2.1 Comparison results based on flexibility indicators

This section will compare the scheduling results considering the system flexibility cost with those not considering the flexibility cost. The scheduling output sequence results are shown in Figures 4,5. It can be seen that in the scheduling scenario, without considering the flexibility cost, TPUs and CHUs mainly supply power. After considering the system flexibility, the GU keeps shut down at the time when the PV dominates the power supply at the source side, i.e., 3:00 pm-5:00 p.m. In contrast, the GU starts up at the time when the PV output is small, which significantly reduces the TPU output level. The power supply pressure of the CHUs is also alleviated. The above process considering flexibility cost reflects the flexible supply of the source side. In addition, the system's PV abandonment rate is also reduced. This is because scenario one considers the flexibility limit of the system itself. By introducing the risk cost of insufficient flexibility into the objective function, the balance between the optimal economic cost and the reduction of the clean energy abandonment is achieved better.

Scenario one is further refined to obtain the calculation results of the system flexibility index at all times over a day, as shown in Figures 6,7. Figure 6 corresponds to two upward flexibility indexes of UIRRP and USRRE; Figure 7 corresponds to the two downward flexibility indexes of DIRRP and DSRRE. It can be seen from Figures 6,7 that when it is between 2:00 p.m. and 6:00 p.m., the probability of insufficient system flexibility is 0. This is because the PV output is large during this period. The system can improve flexibility by switching the PV power station,

TABLE 1 Operating parameters of each unit.

Operating parameters	T1	T2	C1	C2	C3	G1
Unit capacity (MW)	150	60	50	150	250	40
Startup cost (¥)	1800	2000	140	160	200	230
Shutdown cost (¥)	330	310	140	160	200	120
Upward flexibility supply cost (¥/MW)	160	160	120	120	120	100
Downward flexibility supply cost (¥/MW)	130	130	100	100	100	70
Maximum ramping rate (MW/min)	0.25	0.12	0.41	1.24	2.05	1.73

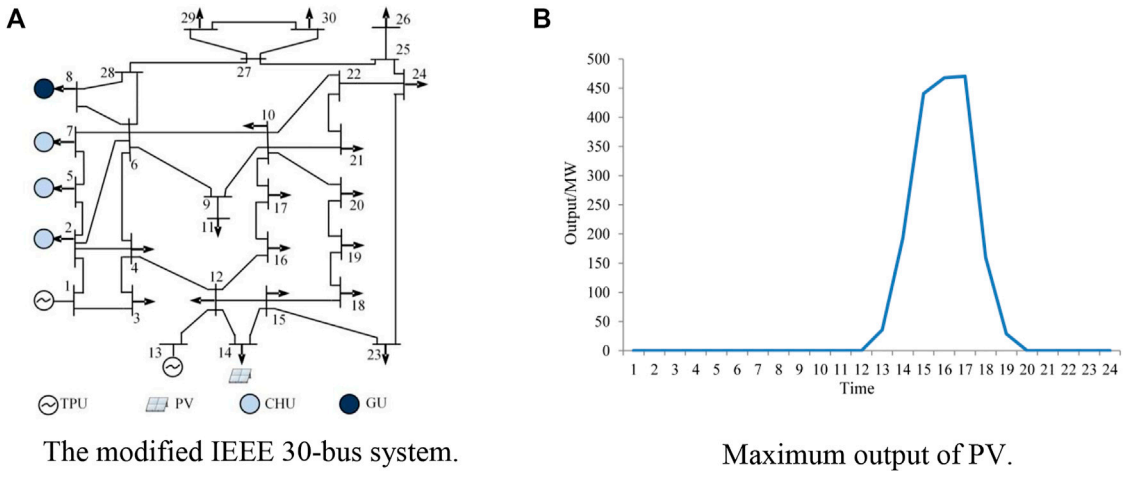


FIGURE 3
The location of each generator unit and maximum output of PV. **(A)** The modified IEEE 30-bus system. **(B)** Maximum output of PV.

and the sufficient system flexibility expectation is large. The maximum value of the USRRE index appears at 3:00 p.m., which is 15.73 MW, and the maximum value of the DSRRE index appears at 5:00 p.m., which is 9.02 MW. In contrast to the sufficient flexibility expectation index, the maximum values of insufficient flexibility probability and insufficient flexibility expectation appear at night, and the maximum values of UIRRP and UIRRE appear at 2:00 a.m., which are 8.48% and 19.20 MW, respectively; The maximum values of DIRRP and DIRRE indexes also appeared at 2:00 a.m., which are 5.51% and 5.23 MW, respectively. It can be seen that the system may still have insufficient flexibility at night. This provides data support for the system to plan further, invest and build high-quality flexible resources, and avoid PV abandonment caused by the imbalance between supply and demand of flexibility.

4.2.2 Comparison results of DR

The settings of scenario three and scenario one are mainly used to analyze the impact of DR implementation on system scheduling results. The scheduling output sequence results of scenario one have been given in the previous section. Here, the total load trend of the system with and without DR and the scheduling output sequence results of the units without DR are given, as shown in Figures 8, 9.

It can be seen that the overall trend of system scheduling output time sequence before and after DR implementation are similar, but there is a significant difference in output level. Specifically, since the cost of PV output is 0, the users participating in DR are stimulated to power consumption at 14:00–18:00, that is, the time when the PV output is large to reduce the system operation cost. The rest of the time, the power consumption of the system decreases, resulting in a decrease in TPU generation with the lowest output priority. However, due to system constraints, the total daily load remains unchanged. In general, the abandoned PV power is reduced after DR, which further verifies the technical advantages of DR implementation in promoting PV consumption.

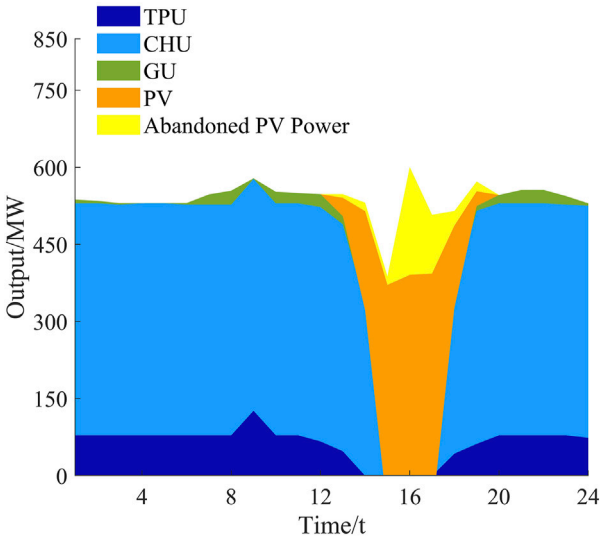


FIGURE 4
HTSGS scheduling results in Scenario 1.

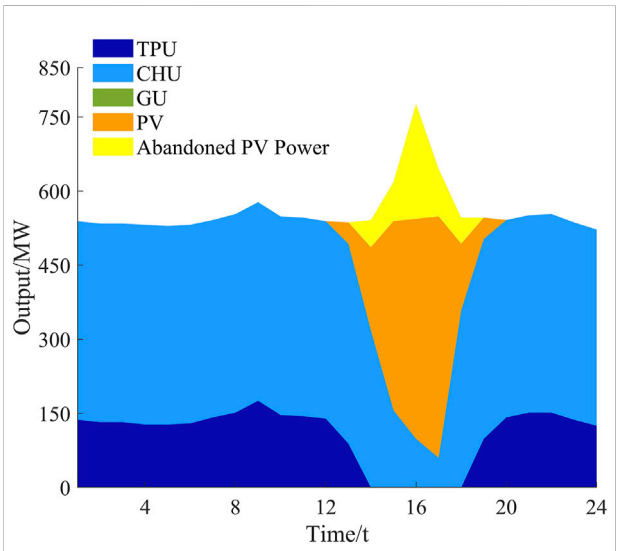


FIGURE 5
HTSGS scheduling results of in Scenario 2.

4.2.3 Comparison results of carbon trading cost

The inclusion of carbon trading costs will lead to changes in the output of TPUs in the system, which will have a chain effect on the scheduling output of other units. Therefore, this paper sets scenarios one and four to make a comparative analysis on whether carbon trading cost is included. The scheduling output sequence results of scenario 4, which is not considering the carbon trading cost, are shown in Figure 10.

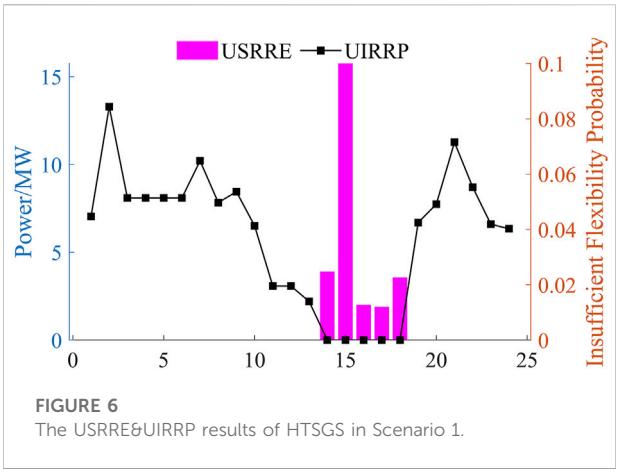


FIGURE 6
The USRRE&UIRRP results of HTSGS in Scenario 1.

It can be observed that the time series curve trend of system scheduling output is the same as that of scenario one considering carbon trading cost, but the TPU output level in scenario 4 has increased by about 50%. In addition, the output level of TPU in the early stage is relatively high, basically maintained at about 150 MW. When PV output is large, the abandoned PV power increases slightly due to an insufficient downward flexible supply of TPU itself.

4.2.4 Economic and technical comparison of all scenarios

Table 2 illustrates the results of operation cost and abandoned PV power under different scenarios. Taking scenario 1 as a comparison benchmark, it can be seen that scenario two is not guided by the flexibility target, so its flexibility cost is significantly increased. Meanwhile, as its output is mainly generated by CHUs and TPUs, the power generation cost and carbon emission cost are also increased. Similarly, DR is not implemented in scenario 3, and the power generation cost rises by 12.8%, which fully reflects the economic deployment effectiveness deficiency in load following source side output. In scenario 4, because the carbon emission cost limit is ignored, the carbon emission cost increases by 31.37%, corresponding to the increase in the TPU output level. At the same time, the power generation cost in this scenario four also increases. Among the four scenarios, the total amount of abandoned PV power in scenario one is the lowest, which fully verifies the promotion effect on PV consumption of the scheduling model proposed in this paper.

4.2.5 Flexibility analysis under different power supply structures

To reveal the impact of different power constructures on system flexibility, the following four power constructive scenarios are set in this section:

Scenario 5: the system includes CHUs and GU;

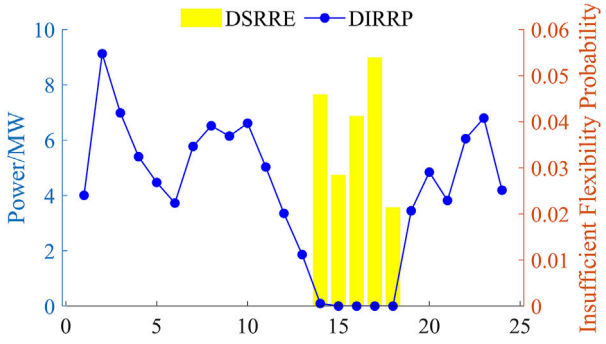


FIGURE 7
The DSRRE&DIRRP results of HTSGS in Scenario 1.

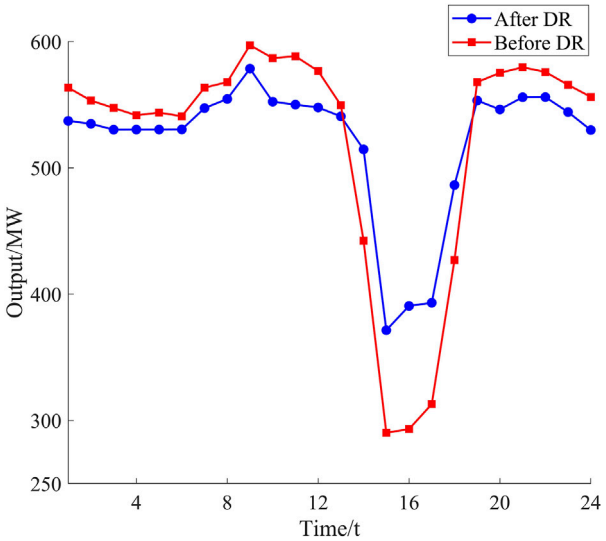


FIGURE 8
The total load comparison before and after DR.

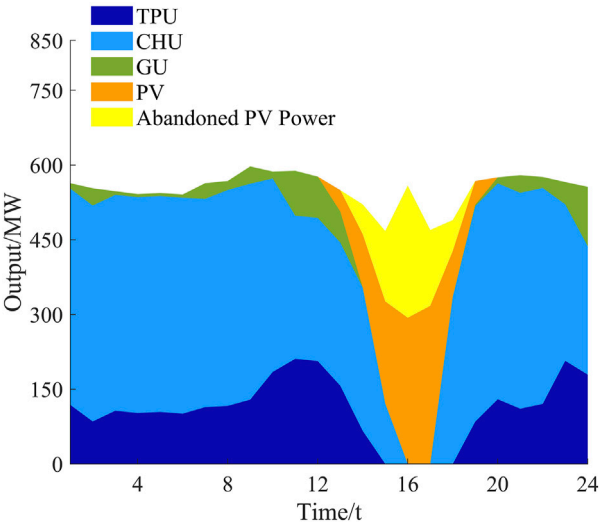


FIGURE 9
HTSGS Scheduling results in Scenario 3.

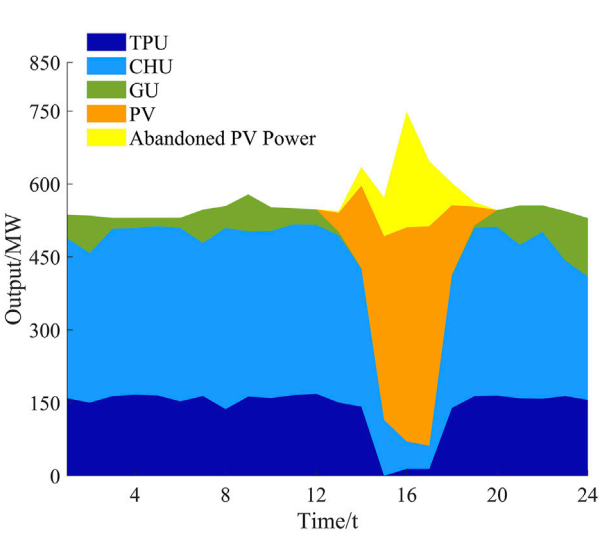


FIGURE 10
HTSGS Scheduling results in Scenario 4.

Scenario 6: the system only includes CHUs, and other parameter settings are consistent;

Scenario 7: the system only includes GU, and the capacity of the original TPUs increases proportionally;

Scenario 8: the system does not include CHU and GU, and the capacity of the original TPUs increases proportionally.

The flexibility indexes of the system under different scenarios are shown in Table 3.

It can be seen from Table 3 that the flexibility index UIRRP and DIRRP in scenario five is the smallest, which indicates that the power system under the HTSGS integration has the smallest probability of insufficient flexibility. In addition, the USRRE and DSRRE of scenario five reach 1.13 MW and 1.33MW, which indicates that the expectation of sufficient flexibility is high in scenario 5. Compared with scenario 5, the probabilities of insufficient flexibility are both increased in scenario six and scenario 7. The probability of insufficient flexibility in scenario eight reaches the maximum, where the UIRRP and DIRRP increase by 23.01% and 19.93% compared with scenario 5. In addition, the flexibility

TABLE 2 The system costs and abandoned PV power among different scenarios.

	Generation Cost/106 ¥	Flexibility Cost/106 ¥	Carbon emission Cost/106 ¥	Total Cost/106 ¥	Abandoned PV Power/MW
Scenario 1	90.86	13.87	8.24	112.97	412.99
Scenario 2	95.27	35.46	8.95	139.68	516.93
Scenario 3	102.49	11.34	10.28	124.11	681.72
Scenario 4	97.86	12.95	11.39	122.2	433.96

TABLE 3 The system flexibility results under different power source constructures.

	Scenario 5	Scenario 6	Scenario 7	Scenario 8
UIRRP	3.75%	8.69%	9.37%	26.76%
DIRRP	2.46%	7.51%	8.60%	22.39%
USRRE/MW	1.13	0.84	0.76	0.20
DSRRE/MW	1.33	0.50	0.45	0.14

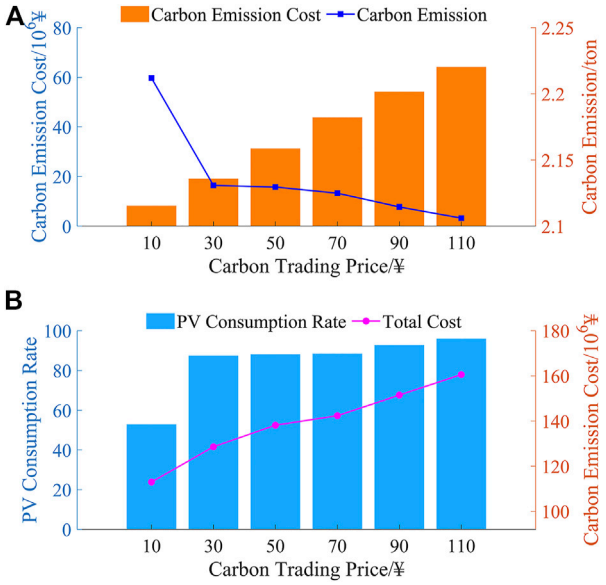


FIGURE 11
The economic and technical sensitivities at different carbon prices.

index USRRE and DSRRE are significantly reduced. This shows that when the HTSGS integration is not considered, the probability of insufficient power system flexibility is high, with 26.76% and 22.39% of probability of upward and downward insufficient flexibility. This is because the original TPUs of the system have the poor flexible regulation ability, and the flexible resource allocation of scenario eight is insufficient. From the above analysis, it can be seen that the HTSGS integration plays a positive role in improving the power system flexibility, which can effectively reduce the low flexibility probability and improve the sufficient flexibility probability of the system.

4.2.6 System economy analysis under different carbon trading prices

The carbon trading price will affect the system carbon cost, thus affecting the system scheduling results. This paper explores the operation cost and PV consumption rate under different carbon trading prices, and the results are shown in Figure 11.

It can be seen that with the rise of carbon trading price, the PV consumption rate gradually rises, and the carbon emission continues to decline. However, the system's carbon emission cost and total cost show an increasing trend. Among them, the carbon emission and PV consumption rate change obviously when the carbon emission price is among 10 and 30¥ and increases slowly after 30¥. At this time, the continued rise of carbon trading prices will bring little economic benefits to the system. In addition, due to the lower limit of TPU output, the PV consumption cannot be increased all the time, so the PV consumption trend slows down. In general, although the moderate increase in carbon emission price leads to an increase in system operation cost, the PV consumption rate and carbon emissions performances have been significantly improved, meeting the overall strategic requirements of energy conservation and emission reduction in the power industry.

5 Conclusion

In this paper, multi-dimensional economic scheduling is carried out for the power system with HTSGS integration considering its flexibility, carbon emission reduction, and DR. Firstly, the system flexibility resource is modeled, and the flexibility index is proposed. Then, considering the power flow constraints, source side operation constraints, system flexibility constraints, and DR constraints, a low-carbon flexible economic scheduling model of the system is constructed. The model minimizes the total cost of system operation, carbon emissions, and flexibility.

The case studies results show that the introduction of flexibility cost and carbon trading cost is beneficial to the system to adjust the scheduling strategy and reduce the PV power abandonment; DR can improve the source load timing balance and promote PV consumption. In general, the comprehensive consideration of system flexibility cost, carbon trading cost, and DR can effectively reduce system cost, improve system flexibility, and reduce carbon emission.

The internal multiple sources of HTSGS can complement each other and promote PV consumption. In addition, the sufficient flexibility of the system has been significantly improved after the HTSGS integration.

The system's carbon emissions will significantly decrease with the increase of the carbon trading price within a certain limit, and the PV consumption rate and the carbon emission price show the same trend.

In future work, we can study the effect of various renewable energy power generation scenarios, i.e., different weather conditions, on the proposed scheduling model. Besides, establishing a robust economic optimization scheduling model by fully considering the relationship between the robustness and economic constraints of the proposed scheduling strategy is worth studying. (Bouffard and Ortega-Vazquez, 2011), (Lannoye et al., 2012)., (Menemenlis et al., 2011).

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

All authors have made substantial contributions to the conception, design of the work; analysis; and drafting of the paper. All authors read and approved the final manuscript.

Funding

This research is supported by the Science and Technology Project of State Grid Corporation of China (Grant No. 5108-202226031A-1-1-ZN). The funder was not involved in the study design, collection, analysis, interpretation of data, the writing of this article, or the decision to submit it for publication.

Conflict of interest

Authors DZ, YW, LX, FD, and ZD was employed by the State Grid Sichuan Electric Power Company.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Akbari-Dibavar, A., Mohammadi-Ivatloo, B., Zare, K., Khalili, T., and Bidram, A. (2021). Economic-emission dispatch problem in power systems with carbon capture power plants. *IEEE Trans. Ind. Appl.* 57 (4), 3341–3351. doi:10.1109/tia.2021.3079329

Bouffard, F., and Ortega-Vazquez, M. (July 2011). "The value of operational flexibility in power systems with significant wind power generation", Proceedings of the 2011 IEEE Power and Energy Society General Meeting, Detroit, MI, USA, 1–5.

Brahma, D., and Senroy, N. (2020). Sensitivity-based approach for assessment of dynamic locational grid flexibility. *IEEE Trans. Power Syst.* 35 (5), 3470–3480. doi:10.1109/tpwrs.2020.2985784

Buhan, S., Kucuk, D., Cinar, M. S., Guvengir, U., Demirci, T., Yilmaz, Y., et al. (2020). A scalable river flow forecast and basin optimization system for hydropower plants. *IEEE Trans. Sustain. Energy* 11 (4), 2220–2229. doi:10.1109/tste.2019.2952450

Du, E., Zhang, N., Hodge, B.-M., Wang, Q., Lu, Z., Kang, C., et al. (2019). Operation of a high renewable penetrated power system with CSP plants: A look-ahead stochastic unit commitment model. *IEEE Trans. Power Syst.* 34 (1), 140–151. doi:10.1109/tpwrs.2018.2866486

Emmanuel, M., Doubleday, K., Cakir, B., Marković, M., and Hodge, B.-M. (2020). A review of power system planning and operational models for flexibility assessment in high solar energy penetration scenarios. *Sol. Energy* 210, 169–180. doi:10.1016/j.solener.2020.07.017

Fan, L., Zhao, C., Zhang, G., and Huang, Q. (2022). Flexibility management in economic dispatch with dynamic automatic generation control. *IEEE Trans. Power Syst.* 37 (2), 876–886. doi:10.1109/tpwrs.2021.3103128

Heggarty, T., Bourmaud, J.-Y., Girard, R., and Kariniotakis, G. (2020). Quantifying power system flexibility provision. *Appl. Energy* 279, 115852. doi:10.1016/j.apenergy.2020.115852

Lannoye, E., Flynn, D., and O'Malley, M. (2012). Evaluation of power system flexibility. *IEEE Trans. Power Syst.* 27 (2), 922–931. doi:10.1109/tpwrs.2011.2177280

Li, H., Zhang, N., Fan, Y., Dong, L., and Cai, P. (2022). Decomposed modeling of controllable and uncontrollable components in power systems with high penetration of renewable energies. *J. Mod. Power Syst. Clean Energy* 10 (5), 1164–1173. doi:10.35833/mpce.2020.000674

Li, P., Yu, D., Yang, M., and Wang, J. (2018). Flexible look-ahead dispatch realized by robust optimization considering CVaR of wind power. *IEEE Trans. Power Syst.* 33 (5), 5330–5340. doi:10.1109/tpwrs.2018.2809431

Lu, Z., Li, H., and Qiao, Y. (2018). Probabilistic flexibility evaluation for power system planning considering its association with renewable power curtailment. *IEEE Trans. Power Syst.* 33 (3), 3285–3295. doi:10.1109/tpwrs.2018.2810091

Ma, J., Silva, V., Belhomme, R., Kirschen, D. S., and Ochoa, L. F. (2013). Evaluating and planning flexibility in sustainable power systems. *IEEE Trans. Sustain. Energy* 4 (1), 200–209. doi:10.1109/tste.2012.2212471

Ma, R., Li, K., Li, X., and Qin, Z. (2015). An economic and low-carbon day-ahead Pareto-optimal scheduling for wind farm integrated power systems with demand response. *J. Mod. Power Syst. Clean. Energy* 3 (3), 393–401. doi:10.1007/s40565-014-0094-7

Menemenlis, N., Huneault, M., and Robitaille, A. (July 2011). "Thoughts on power system flexibility quantification for the short-term horizon", Proceedings of the 2011 IEEE Power and Energy Society General Meeting, Detroit, MI, USA, 1–8.

Mohandes, B., Moursi, M. S. E., Hatziargyriou, N., and Khatib, S. E. (2019). A review of power system flexibility with high penetration of renewables. *IEEE Trans. Power Syst.* 34 (4), 3140–3155. doi:10.1109/tpwrs.2019.2897727

Reddy, S. K., Panwar, L. K., Panigrahi, B. K., and Kumar, R. (2017). Modeling of carbon capture technology attributes for unit commitment in emission-constrained environment. *IEEE Trans. Power Syst.* 32 (1), 662–671. doi:10.1109/tpwrs.2016.2558679

Shi, Y., Dong, S., Guo, C., Chen, Z., and Wang, L. (2021). Enhancing the flexibility of storage integrated power system by multi-stage robust dispatch. *IEEE Trans. Power Syst.* 36 (3), 2314–2322. doi:10.1109/tpwrs.2020.3031324

Wang, S., Liu, J., Chen, H., Bo, R., and Chen, Y. (2021). Modeling state transition and head-dependent efficiency curve for pumped storage hydro in look-ahead dispatch. *IEEE Trans. Power Syst.* 36 (6), 5396–5407. doi:10.1109/tpwrs.2021.3084909

Wang, Y., Qiu, J., and Tao, Y. (2022). Optimal power scheduling using data-driven carbon emission flow modelling for carbon intensity control. *IEEE Trans. Power Syst.* 37 (4), 2894–2905. doi:10.1109/tpwrs.2021.3126701

Yamujala, S., Kushwaha, P., Jain, A., Bhakar, R., Wu, J., and Mathur, J. (2021). A stochastic multi-interval scheduling framework to quantify operational flexibility in low carbon power systems. *Appl. Energy* 304, 117763. doi:10.1016/j.apenergy.2021.117763

Zhang, S., Xiang, Y., Liu, J., Liu, J., Zhao, X., Jawad, S., et al. (2022). A regulating capacity determination method for pumped storage hydropower to restrain PV generation fluctuation. *CSEE J. Power Energy Syst.* 8 (1), 304–316. doi:10.17775/cseejpes.2020.01930

Zhao, J., and Xu, Z. (2017). Ramp-limited optimal dispatch strategy for PV-embedded microgrid. *IEEE Trans. Power Syst.* 32 (5), 4155–4157. doi:10.1109/tpwrs.2017.2670920

Zhao, Y., Wang, C., Zhang, Z., and Lv, H. (2021). Flexibility evaluation method of power system considering the impact of multi-energy coupling. *IEEE Trans. Ind. Appl.* 57 (6), 5687–5697. doi:10.1109/tia.2021.3110458

frontiers | Frontiers in Energy Research

TYPE Original Research
PUBLISHED 20 January 2023
DOI 10.3389/fenrg.2022.1088100

OPEN ACCESS

EDITED BY
Yaser Qudaih,
Higher Colleges of Technology, United Arab Emirates

REVIEWED BY
Weiye Zheng,
South China University of Technology, China
Wanjun Huang,
Beihang University, China
Ao Qiao,
Chengdu University of Information Technology, China

*CORRESPONDENCE
Zhiyuan Tang,
tangzhiyuan@scu.edu.cn

SPECIALTY SECTION
This article was submitted to Smart Grids,
a section of the journal
Frontiers in Energy Research

RECEIVED 03 November 2022
ACCEPTED 05 December 2022
PUBLISHED 20 January 2023

CITATION
Tang Y, Tang Z, Gao Y, Wei L, Zhou J, Han F and Liu J (2023), An adaptive forecasting method for the aggregated load with pattern matching.
Front. Energy Res. 10:1088100.
doi: 10.3389/fenrg.2022.1088100

An adaptive forecasting method for the aggregated load with pattern matching

Yikun Tang[1], Zhiyuan Tang[1]*, Yi Gao[2], Lianbin Wei[3], Jin Zhou[2], Fujia Han[4] and Junyong Liu[1]

[1]College of Electrical Engineering, Sichuan University, Chengdu, Sichuan, China, [2]State Grid Tianjin Power Economics and Technology Research Institute, Tianjin, China, [3]State Grid Tianjin Electric Power Company, Tianjin, China, [4]China Electric Power Research Institute, Beijing, China

Electrical load forecasting plays a vital role in the operation of power system. In this paper, a novel adaptive short-term load forecasting method for the aggregated load is built. The proposed method consists of two stages: load forecast model preparation stage and adaptive load forecast model selection stage. In the first stage, based on historical load data of all consumers, the typical monthly load patterns are firstly identified in an optimal fashion with the aid of the cosine similarity. Then, for each identified monthly load pattern, a stacking ensemble learning method is proposed to train the load forecasting model. In the second stage, according to the similarity between individual load data of the latest month and the identified monthly load pattern, all the consumers are firstly classified into different groups where each group corresponds to a particular load pattern. Then, for each group, the corresponding trained load forecasting model is employed for short-term load forecast and the final forecast of the aggregated load is calculated as a simple aggregation of the produced load forecast for each group of consumers. Case studies conducted on open dataset show that, compared with the single forecasting model, the proposed adaptive load forecasting method can effectively improve the load forecasting accuracy.

KEYWORDS

adaptive load forecasting, aggregated load, pattern matching, cosine similarity, stacking ensemble learning

1 Introduction

Load forecasting is an important part of the smart grid, which plays an important role in the security, stability and economic dispatch of the power system (Diamantoulakis et al., 2015). Rapid progress in the consumer-side advanced measurement system brings about a large amount of consumer-side power consumption data which has opened the possibility of designing various load forecasting methods (Stephen et al., 2017; Chaouch, 2014). The key challenge becomes how to effectively process these data and improve the consumer-side load prediction accuracy.

As mentioned above, load forecasting has always been a research hotspot due to its vital role in the operation of power systems. Many experts and scholars have done in-

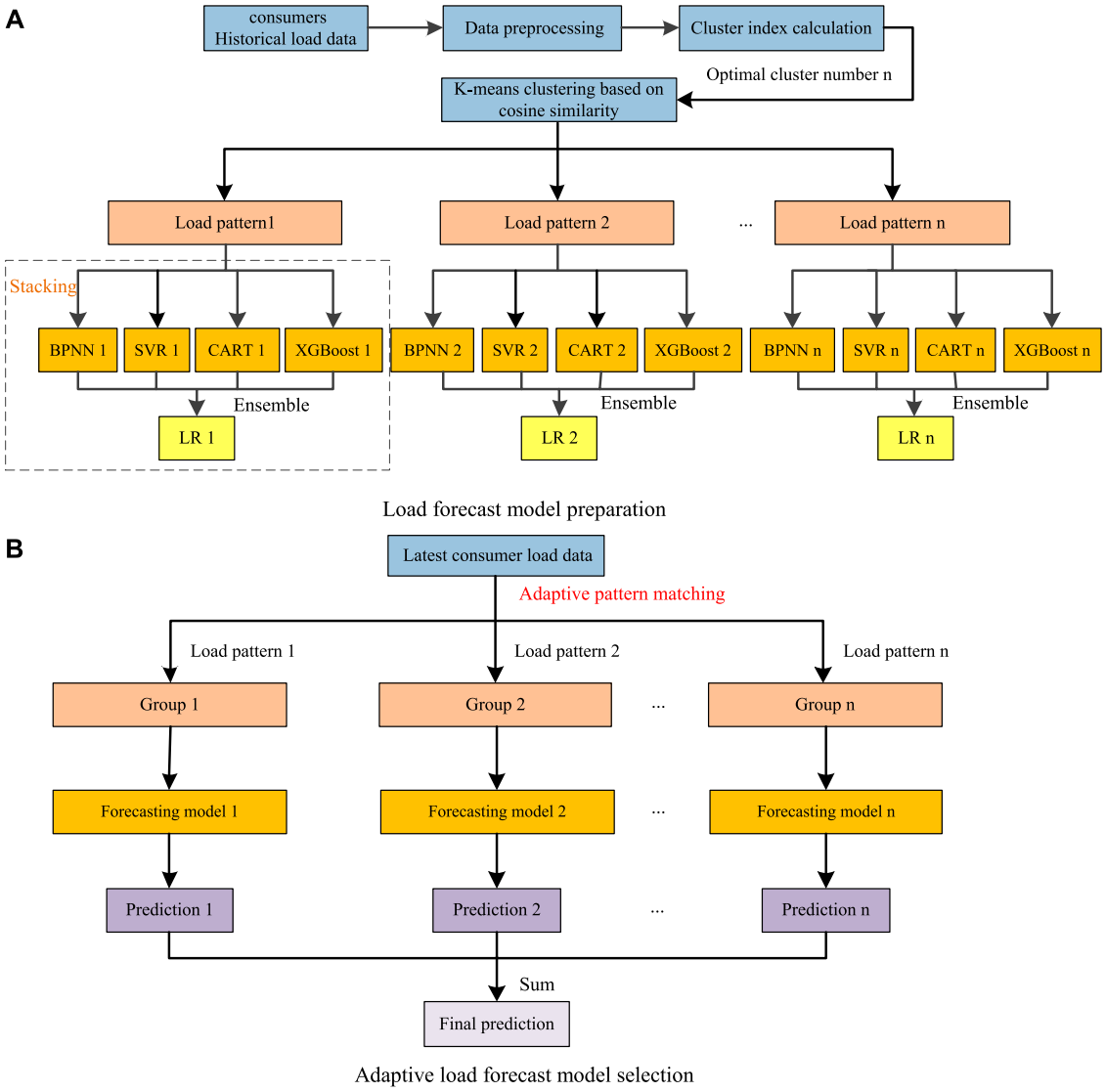


FIGURE 1
The full flowchart of the proposed method. **(A)** Load forecast model preparation. **(B)** Adaptive load forecast model selection.

depth research on it, and proposed many models and algorithms. As the typical time-series forecasting algorithms, auto regression (Mbamalu and El-Hawary, 1993) and auto regression moving average (Huang and Shih, 2003) have been proposed for load forecast. Although these methods require a small amount of data for calculation and also have a fast computation speed, they are often criticized for their limited success in handling complex non-linear relationship and relatively poor learning ability. Moreover, they have high requirements for data stability which however is not the intrinsic nature of load in distribution networks.

To overcome above mentioned issues, machine learning techniques have been proposed for load forecasting, such as artificial neural network (ANN) (Hippert et al., 2001; Quan et al., 2014), support vector regression (SVR) (Tan et al., 2020a), Classification And Regression Tree (CART) (Zhong and Tam, 2015), Random Forest (RF) (Xuan et al., 2021), extreme gradient boosting (XGBoost) (Yao et al., 2022), and so on. These algorithms have been successfully applied to tackle non-linear feature fitting in load forecasting. For example, in (Guan et al., 2013), separate neural networks (NN s) are applied to wavelet decomposed filtered load data. In (Haque, 2019), a SVR based model is presented to determine the impact of HVAC set point adjustments on building-level electrical load. In (Liao, 2019), a XGBoost based model is designed to analyze the common laws of meteorological and daily types on the load. Besides the machine

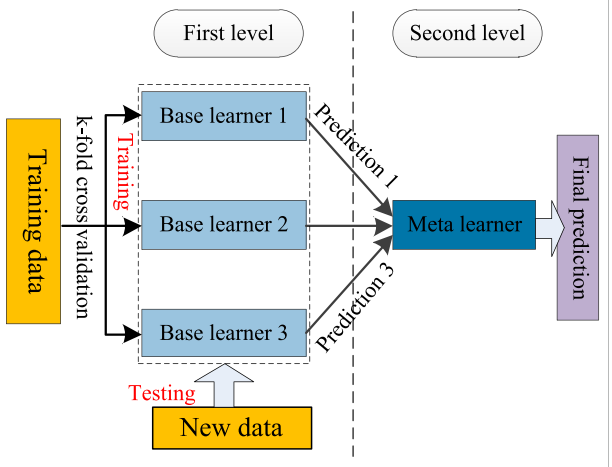


FIGURE 2
Ensemble learning method based on Stacking.

learning methods, deep learning is also used extensively for load forecasting due to its powerful learning ability. For example, Recurrent Neural Network (RNN) (Zhang et al., 2020) is the commonly used deep learning model for time series forecasting. In (Shi et al., 2018), a RNN based model is designed to forecast the household load forecasting, whose performance is obviously better than time-series analysis and machine learning approaches. Long Short-term Memory network (LSTM) (Tan et al., 2020b) and Gated Recurrent Unit network (GRU) (Sajjad et al., 2020) are two of variants of RNN, and can achieve better performance than RNN. In (Kumar et al., 2018), LSTM-RNN and GRU-RNN based model is applied to spark clusters, which brings better performance than RNN. In Kong et al. (2017) proved that LSTM based model is suitable, and claimed that the model can be further improved if more energy sequences are obtained. Although machine learning and deep learning approaches can process massive amounts of data and obtain better prediction results than traditional methods, they need to consume more memory and time. In addition, parameter adjustment is also one of the difficulties of machine learning and deep learning approaches.

Although the above-mentioned forecasting models have achieved great accuracy in load forecasting, there is still one critical challenge in the real-world load forecasting, i.e., the forecasting accuracy of a certain forecasting model cannot be guaranteed in all possible load patterns. This is because that the generalization ability of a certain forecasting model is usually limited so that they cannot adapt to all various load patterns with different fluctuation characteristics. Compared with the aggregated load in transmission networks, the loads in distribution networks usually have severer fluctuations and uncertainties due to its much smaller size and more random factors such as weather, holiday, working days shifting or some social activities (Fallah et al., 2018). Thus, it is unexpected that in the distribution network a single forecasting model can provide an accurate non-linear function that maps the input vector to the future value of the load. In view of this, a forecasting method of pattern adaptive matching is proposed in this paper. In the proposed method, different load patterns are firstly obtained by optimal clustering of massive consumer historical loads, which can well characterize the fluctuation characteristics of various loads. Then, for each load pattern, a load forecasting model based on stacking ensemble learning is built, which can establish a multi-model fusion load forecasting model to improve the performance of load forecasting. Next, according to the fluctuation characteristics of the latest consumer load, all the consumers are classified into different groups, where each group corresponds to a particular identified load pattern. Finally, forecast the load of each group separately and sum the forecasts of different groups to obtain the aggregated load forecast. The result based on open dataset demonstrates that compared with the overall forecast by a single method, in the proposed method, the impact of load fluctuations on the forecast accuracy is reduced, and the forecast accuracy increased by 0.94%.

The key contributions of this paper can be briefly summarized as follows:

1) A novel two-stage short-term load forecasting method for the aggregated load is proposed, where the adaptive pattern matching technique is designed to improve the adaptive capability of the proposed forecasting method.
2) Forecasting models based on stacking ensemble learning are built, where the advantages of integrated heterogeneous models are fully utilized.
3) We validate our proposed method with a public dataset and compare its performance with several single load forecasting methods, where its superior performance is demonstrated for the proposed method.

The remainder of this letter is organized as follows: Section 2 focuses on the implementation of the forecasting model. Section 3 conducts case studies and makes comparisons to verify the effectiveness of our proposed method. Finally, Section 4 draws conclusion and gives an outlook on future work.

2 Proposed methodology

The construction of the adaptive forecasting method with pattern matching proposed in this paper is shown in Figure 1.

The overall forecasting process of the proposed method can be divided into two stages: load forecast model preparation stage and adaptive load forecast model selection stage. The construction is shown in Figure 1. In the load forecast model preparation stage, load patterns are optimally obtained from historical load data. And for each load pattern, stacking ensemble learning based forecasting models is well prepared. In this stage, we first utilize the cosine

similarity to measure the morphological distance between consumers' historical load data, then the load patterns are clustered by using the k-means algorithm. The optimal number of clusters is determined by cluster index. Finally, the Stacking based load forecasting model is built for each load pattern. The stacking based model has two levels. In the first level, the load pattern data are used to train the Back Propagation Neural Network (BPNN) (Park et al., 1991), SVR, CART and XGBoost based models. Then, in the second level, the liner regression (LR) model is employed to determine the weights and fuses the multiple level-one models. While in the adaptive load forecast model selection stage, different groups of consumers are constructed based on their load patterns and forecasting is conducted on the grouped load individually. In this stage, according to the similarity between the latest consumer load data and the load pattern, all consumer load data are firstly classified into the previously identified load pattern. Afterwards, the corresponding forecasting model of each load pattern is employed for prediction. Finally, sum the predictions of different groups to obtain the final prediction.

2.1 Load forecast model preparation stage

It is clear that load consumption differs in both magnitude and time of use, and is dependent on lifestyle, weather, and many other relevant factors. In order to classify dissimilar loads into different clusters, this paper uses cosine similarity to measure the difference in consumer load curve shape, and uses k-means algorithm for clustering. Let $x_i=(x_{i1}, x_{i2}, \ldots, x_{in})$,$x_j=(x_{j1}, x_{j2}, \ldots, x_{jn})$ be two load curves, and the similarity between them is defined as:

$$S\left(x_i, x_j\right) = 1 - \frac{\sum_{q=1}^{n}\left(x_{iq} \times x_{jq}\right)}{\sqrt{\sum_{q=1}^{n}\left(x_{iq}\right)^2} \times \sqrt{\sum_{q=1}^{n}\left(x_{jq}\right)^2}} \tag{1}$$

It can be concluded from Eq. 1 that the more similar the two curves are in shape, the lower the cosine similarity is. When the two curves are completely identical in shape, the cosine similarity is 0.

Next, we construct a clustering indicator for *K*-means clustering method. As a classic clustering algorithm, *K*-means algorithm is widely used in power load clustering analysis. But the number of clusters directly affects the clustering effect. In order to determine the optimal number of clusters, we adopt the graph-based clustering effect evaluation method to establish the clustering evaluation index.

$$V(K) = \sum_{c=1}^{K}\sum_{i=1}^{n_K} S(x_i, v_c) \tag{2}$$

where K is the number of clusters, v_c denotes the cluster center of each cluster; n_K represents the number of samples contained in the current K-th cluster. V represents the sum of distances from various samples to their corresponding cluster centers. V decreases with the increase of the number of clusters. When the decreasing trend of V value tends to be gentle, the corresponding cluster number is the best cluster number.

By applying the above clustering method, we can optimally obtain the load pattern from the historical load data. Then, for each load pattern, a load forecasting model is to be trained. In order to combine the advantages of several heterogeneous models, a stacking based ensemble learning method is proposed to train the load forecasting model. Ensemble learning is a method that combines the prediction results of multiple models to obtain better and more robust prediction than any one model. Ensemble learning based on stacking is an integration technology that combines many different models to form a new model with two-level structure. Figure 2 illustrates its basic structure.

In the first level, multiple different types of base learners are trained to predict the data set. In order to prevent the occurrence of overfitting, the training data is learned in the base learner by means of k-fold cross-validation, and the training results of the base learner are used as the input of the meta learner for training. After the training, the final prediction result can be obtained by taking the test set as input.

In this paper, we select BPNN, SVR, CART, and XGBoost as the base learners. BPNN has self-learning, self-organization, self-adaptive, and strong non-linear function approximation ability, and has strong fault tolerance. The BPNN algorithm is mainly composed of two parts: one is the forward propagation of the signal; the other is the back-propagation of the error between the output value and the real value, and constantly corrects the parameters of each neuron. SVR is an important application branch of support vector machine (SVM) (Zhang et al., 2011). It maps the input vector to the high-dimensional feature space through non-linear mapping and obtains a regression hyperplane through calculation, so that the sum of distances from all sample points in the set to the hyperplane is minimum. The CART algorithm recursively performs binary segmentation on each feature, The segmentation method adopts the Gini index estimation function based on the minimum distance, and divides the current sample set into two sub-sample sets so that each non-leaf node generated has two branches. XGBoost is an engineering implementation of gradient boosting decision tree (GBDT) algorithm (Liu et al., 2018). Its basic idea is that multiple weak regression trees are linearly combined into strong regression trees. Its model is defined as:

$$F(x) = \sum_{m=1}^{M} f_m(x) \tag{3}$$

where $F(x)$ is the final output value and $f_m(x)$ is the output value of the m-th weak regression tree.

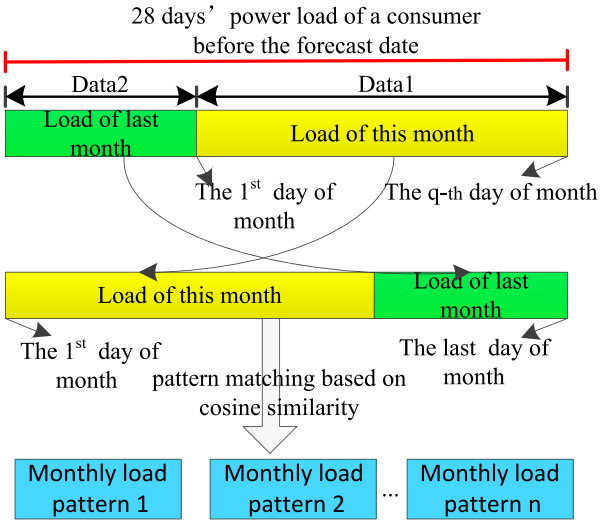


FIGURE 3
Pattern matching diagram.

In the second level, linear regression is chosen as the meta learner. That is, the predictions obtained by the base learners are optimally weighted by the least squares algorithm to obtain the final prediction. The model is defined as:

$$P = w_0 + w_1 y_1 + w_2 y_2 + \ldots + w_n y_n \tag{4}$$

where w represents the weight coefficient, y is the prediction of the base learners, P is the final prediction.

In the process of model training, the prediction $\hat{P}$ of the sample is known, and the weighting coefficient W can be obtained by the following formula.

$$W = \left(Y^T Y\right)^{-1} Y^T \hat{P} \tag{5}$$

where $Y = [\boldsymbol{1}, y_1, y_2, \ldots, y_n]$, $W = [w_0, w_1, w_2, \ldots, w_n]^T$, $\boldsymbol{1}$ is the column vector whose elements are all 1.

2.2 Adaptive load forecast model selection stage

When performing load forecasting in real time, the consumer's electricity load should be firstly matched with the load pattern that has been divided according to the morphological similarity between the consumers' latest electricity consumption data and the identified load pattern. Then the consumer's power loads are predicted by the prediction model corresponding to the same load pattern.

Figure 3 shows the diagram of pattern matching approach. In this approach, we set a month as 28 days (the same as below). When forecasting the load on the day q + 1 (q = 0, 1, ..., 27) of a month, the selected data includes the data1 of the first q days of this month and the load data2 for the (28-q) days of the previous month. In order to make the selected data time consistent with the monthly load pattern, move data1 to data2 and splice it into a consumer's completed monthly load data3 from the beginning of the month to the end of the month. Calculate the cosine similarity between each consumer data3 and each monthly load pattern, and classify each consumer into the most similar monthly load pattern with its lowest value. Thus, all the consumers are divided into different groups. Each group is forecasted separately by the corresponding forecasting model. Then, these predictions are summed to form the aggregated load prediction.

3 Case study

In order to assess the performance of our model, we chose the grid-side power consumption data of 300 households in Australia for analysis (https://www.ausgrid.com.au/Industry/Innovation-and-research/Data-to-share/Solar-home-electricity-data). The data length is from 1 July 2010, to 30 June 2011. The time resolution is 30 min.

To evaluate the overall forecasting accuracy of the proposed model, two commonly used evaluation metrics are adopted in this work. They are Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE), whose expressions are given as

$$MAPE = \sum_{i=1}^{N} \left| \frac{e_i - o_i}{e_i} \right| \times \frac{100}{N} \tag{6}$$

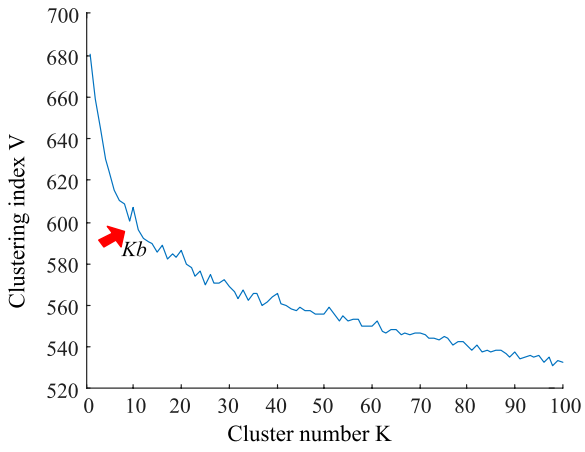


FIGURE 4
Clustering index.

TABLE 1 Pattern matching results on November 7.

Pattern	Pattern 1	Pattern 2	Pattern 3	Pattern 4	Pattern 5	Pattern 6	Pattern 7	Pattern 8	Pattern 9
Numbers of consumer	10	28	27	83	6	74	2	64	6
Matching index	0.036	0.018	0.016	0.009	0.078	0.006	0.123	0.018	0.139

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(e_i - o_i)^2} \tag{7}$$

where e is the actual value, o is the predicted value, and N is the predicted number.

3.1 Identification of load patterns

Firstly, select the data of 300 households in the first 10 months from 1 July 2010, to formulate a total of 3,000 monthly load data samples. The time resolution is 30 min. We increase the number of clusters to obtain the change graph of clustering index, as shown in Figure 4, using the K-means algorithm based on cosine similarity. When the K value is between 9 and 12, the downward trend of the evaluation index V tends to slow down. Select $K_b = 9$ as the best cluster number. Then the K-means clustering is carried out again to obtain nine kinds of monthly load patterns, as shown in Figure 5.

Let the load at the predicted time t P_t be the output, and select historical load, working day, weekend, and time of day as feature inputs, the input feature are:

$$x = \{P_{t-h}, P_{t-h-1}, P_{t-h-2}, P_{t-h+1}, P_{t-h+2}, P_{t-2h}, P_{t-2h-1}, P_{t-2h+1}, P_{t-3h}, P_{t-7h}, weekday, hour\} \tag{8}$$

where the first seven elements are historical load data, h is the time of the day when the load data is collected, weekday and hour respectively represent the week and hour. 1,344 samples can be

TABLE 2 Load forecasting results of different methods.

Method	MAPE (%)	RMSE
Pattern matching	4.04	5.23
BPNN	6.88	8.19
SVR	4.98	6.10
CART	6.10	8.18
XGBoost	5.25	5.90

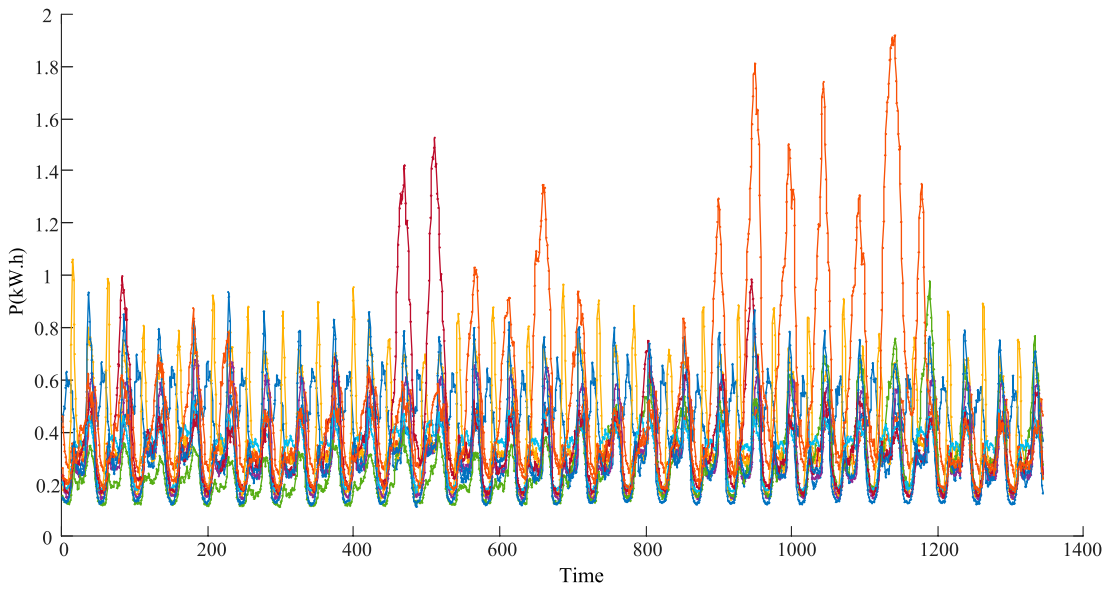


FIGURE 5
Monthly load pattern.

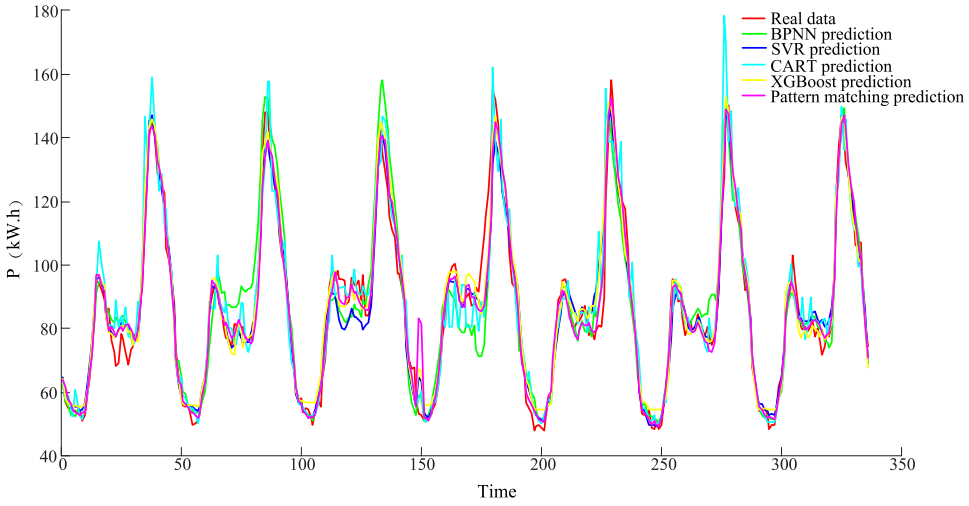


FIGURE 6
Prediction results of different methods.

constructed in each month's load pattern, 70% of the samples are randomly selected as training samples, and the remaining samples are used as validation samples. The best super parameters of each base learner are obtained by the grid search method.

3.2 Pattern matching and load forecasting

Select the first 7 days of November for load forecasting. Pattern matching is required for each day from the first day to the seventh day of November. Table 1 give the pattern matching results of the first day of November. The matching index in the table refers to the cosine similarity between the data obtained by summing the data of consumers that match the same load pattern with the corresponding monthly load pattern It can be seen from the table that the fewer consumers the model matches, the higher the matching index, that is, the lower the similarity between the consumer load data and the monthly load pattern, which will make the prediction effect worse to a certain extent, but the number of consumers is small, the impact on the overall load prediction is small.

After the pattern matching, all consumers are assigned to different groups, each of which corresponds to a previous load pattern. Forecast the load separately for each group, and sum the predictions from all the groups to get the load forecasting results for the first 7 days of November, as shown in Figure 6 and Table 2. In order to verify the effectiveness of the method proposed in this letter, the overall prediction by the single method of BPNN, SVR, CART, and XGBoost is introduced.

Comparing the above results, it is found that, compared with other methods using a single model, the adaptive pattern matching prediction method proposed in this letter can achieve the best prediction effect and can improve the prediction accuracy by at least about 0.94%.

4 Conclusion and future works

In this paper, we have proposed an adaptive forecasting method for the aggregated load with pattern matching. The main advantage is that it can select the proper load pattern corresponding to the load in time according to the change of load fluctuation, and construct the best prediction model for prediction. Compared with a single load forecasting model, the forecasting accuracy of the proposed model is greatly improved. We will extend the method from deterministic forecasting to probabilistic forecasting. The application of our proposed method to wind or solar forecasting will also be investigated.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

Author contributions

The first author, YT, and the corresponding author, ZT, contributed to the main conception, analysis frame of this study and the manuscript preparation. JL gave constructive discussion,

contributed to the analysis, and made conclusion. YG, LW, JZ, and FH gave some advices.

Funding

This work is supported by Science and Technology Project of State Grid Corporation of China (No. 5400-202112507A-0-5-ZN).

Conflict of interest

LW was employed by State Grid Tianjin Electric Power Company.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The authors declare that this study received funding from State Grid Corporation of China. The funder had the following involvement in the study: Data provision and the suggestions for writing.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Chaouch, M. (2014). Clustering-based improvement of nonparametric functional time series forecasting: Application to intra-day household-level load curves. *IEEE Trans. Smart Grid* 5 (1), 411–419. doi:10.1109/TSG.2013.2277171

Diamantoulakis, P. D., Kapinas, V. M., and Karagiannidis, G. K. (2015). Big data analytics for dynamic energy management in smart grids. *Big Data Res.* 2, 94–101. doi:10.1016/j.bdr.2015.03.003

Fallah, S. N., Deo, R. C., Shojafar, M., Conti, M., and Shamshirband, S. (2018). Computational intelligence approaches for energy load forecasting in smart energy management grids: State of the art, future challenges, and research directions. *Energies* 11 (3), 596–626. doi:10.3390/en11030596

Guan, C., Luh, P. B., Michel, L. D., Wang, Y., and Friedland, P. B. (2013). Very short-term load forecasting: Wavelet neural networks with data pre-filtering. *IEEE Trans. Power Syst.* 28 (1), 30–41. doi:10.1109/TPWRS.2012.2197639

Haque, A. IEEE.An SVR-based building-level load forecasting method considering impact of HVAC set pointsProceedings of the 2019 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)February 2019Washington, DC, USA.

Hippert, H. S., Pedreira, C. E., and Souza, R. C. (2001). Neural networks for short-term load forecasting: A review and evaluation. *IEEE Trans. Power Syst.* 16 (1), 44–55. doi:10.1109/59.910780

Huang, S. J., and Shih, K. R. (2003). Short-term load forecasting via ARMA model identification including non-Gaussian process considerations. *IEEE Trans. Power Syst.* 18, 673–679. doi:10.1109/tpwrs.2003.811010

Kong, W., Dong, Z. Y., Hill, D. J., Luo, F., and Xu, Y. (2017). Short-term residential load forecasting based on resident behaviour learning. *IEEE Trans. Power Syst.* 33, 1087–1088. doi:10.1109/tpwrs.2017.2688178

Kumar, S., Hussain, L., Banarjee, S., and Reza, M., "Energy load forecasting using deep learning approach-LSTM and GRU in spark cluster", Proceedings of the 2018 Fifth International Conference on Emerging Applications of Information Technology (EAIT), January 2018, Kolkata, India, pp. 1–4. doi:10.1109/EAIT.2018.8470406

Liao, X. "Research on short-term load forecasting using XGBoost based on similar days", Proceedings of the 2019 International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS), January 2019, Changsha, China 2019.doi:10.1109/ICITBS.2019.00167

Liu, S., Cui, Y., Ma, Y., and Liu, P., "Short-term load forecasting based on GBDT combinatorial optimization", Proceedings of the 2018 2nd IEEE Conference on Energy Internet and Energy System Integration (EI2), Beijing, China, October 2018, pp. 1–5. doi:10.1109/EI2.2018.8582108

Mbamalu, Gan, and El-Hawary, M. E. (1993). Load forecasting via suboptimal seasonal autoregressive models and iteratively reweighted least squares estimation. *IEEE Trans. Power Syst.* 8, 343–348. doi:10.1109/59.221222

Park, D. C., El-Sharkawi, M., Marks, R., Atlas, L., and Damborg, M. (1991). Electric load forecasting using an artificial neural network. *IEEE Trans. Power Syst.* 6, 442–449. doi:10.1109/59.76685

Quan, H., Srinivasan, D., and Khosravi, A. (2014). Short-Term Load and Wind Power Forecasting Using Neural Network-Based Prediction Intervals. [J]. *IEEE Trans. Neural Netw. Learn. Syst.* 25 (2), 303–315.

Sajjad, M., Khan, Z. A., Ullah, A., Hussain, T., Ullah, W., Lee, M. Y., et al. (2020). A novel CNN-GRU-Based hybrid approach for short-term residential load forecasting. *IEEE Access* 8, 143759–143768. doi:10.1109/ACCESS.2020.3009537

Shi, H., Xu, M., and Li, R. (2018). Deep learning for household load forecasting—a novel pooling deep RNN. *IEEE Trans. Smart Grid* 9 (5), 5271–5280. doi:10.1109/TSG.2017.2686012

Stephen, B., Tang, X., Harvey, P. R., Galloway, S., and Jennett, K. I. (2017). Incorporating practice theory in sub-profile models for short term aggregated residential load forecasting. *IEEE Trans. Smart Grid* 8 (4), 1591–1598. doi:10.1109/TSG.2015.2493205

Tan, M., Yuan, S., Li, S., Su, Y., Li, H., and He, F. H. (2020). Ultra-short-term industrial power demand forecasting using LSTM based hybrid ensemble learning. *IEEE Trans. Power Syst.* 35, 2937–2948. doi:10.1109/tpwrs.2019.2963109

Tan, Z., Zhang, J., He, Y., Zhang, Y., Xiong, G., and Liu, Y. (2020). Short-term load forecasting based on integration of SVR and stacking. *IEEE Access* 99, 227719–227728. doi:10.1109/access.2020.3041779

Xuan, Y., Si, W., Zhu, J., Sun, Z., Zhao, J., and Xu, M. (2021). Multi-model fusion short-term load forecasting based on random forest feature selection and hybrid neural network. *IEEE Access* 9, 69002–69009. doi:10.1109/ACCESS.2021.3051337

Yao, X., Fu, X., and Zong, C. (2022). Short-term load forecasting method based on feature preference strategy and LightGBM-XGboost. *IEEE Access* 10, 75257–75268. doi:10.1109/ACCESS.2022.3192011

Zhang, M., Yu, Z., and Xu, Z. (2020). Short-term load forecasting using recurrent neural networks with input attention mechanism and hidden connection mechanism. *IEEE Access* 8, 186514–186529. doi:10.1109/ACCESS.2020.3029224

Zhang, Q., Lai, K. K., and Niu, D., "Optimization combination forecast method of SVM and WNN for power load forecasting", Proceedings of the 2011 Fourth International Joint Conference on Computational Sciences and Optimization, April 2011, Kunming and Lijiang City, China, pp. 249–253. doi:10.1109/CSO.2011.193

Zhong, S., and Tam, K. (2015). Hierarchical classification of load profiles based on their characteristic attributes in frequency domain. *IEEE Trans. Power Syst.* 30 (5), 2434–2441. doi:10.1109/TPWRS.2014.2362492

frontiers | Frontiers in Energy Research

TYPE Original Research
PUBLISHED 23 January 2023
DOI 10.3389/fenrg.2022.1087487

OPEN ACCESS

EDITED BY
Xin Zhang,
Brunel University London,
United Kingdom

REVIEWED BY
Bi Liu,
Sichuan Agricultural University, China
Jianjun Chen,
University of Electronic Science and Technology of China, China
Bin Zhang,
Aalborg University, Denmark

*CORRESPONDENCE
Jichun Liu,
jichunliu@scu.edu.cn

SPECIALTY SECTION
This article was submitted to Smart Grids, a section of the journal Frontiers in Energy Research

RECEIVED 02 November 2022
ACCEPTED 25 November 2022
PUBLISHED 23 January 2023

CITATION
Fan R, Ming Z, Xu W, Li T, Han Y, Ma R, Liu J and Wu Y (2023), Optimization of photovoltaic panel deployment in centralized photovoltaic power plant under multiple factors.
Front. Energy Res. 10:1087487.
doi: 10.3389/fenrg.2022.1087487

Optimization of photovoltaic panel deployment in centralized photovoltaic power plant under multiple factors

Rongquan Fan[1,2], Ziqiang Ming[3], Weiting Xu[2], Ting Li[1], Yuqi Han[1], Ruiguang Ma[1], Jichun Liu[4]* and Yiyang Wu[4]

[1]State Grid Sichuan Economic Research Institute, Chengdu, China, [2]Sichuan New Electric Power System Research Institute, Chengdu, China, [3]State Grid Sichuan Electric Power Company, Chengdu, China, [4]College of Electrical Engineering, Sichuan University, Chengdu, China

Solar energy is one of the main renewable energy sources and has rapidly developed in many countries. However, the photovoltaic (PV) output power will be different under various meteorological and geographical conditions. Therefore, this paper presents an optimization method for the deployment of PV panels in a centralized PV power plant considering multiple factors. Firstly, the whole planning area is divided into a certain amount of sub-areas according to a given area, and fuzzy C-means algorithm is used for terrain clustering according to the geographical characteristics of the sub-areas. Secondly, the correlation analysis between each meteorological factor and PV output power is carried out separately to select the main factors affecting PV output power, and then the expected annual PV output power under the joint action of several main meteorological factors in each terrain is calculated by dual-stage attention mechanism based long short-term memory algorithm. Finally, according to the expected annual PV output of each terrain, considering the constraints including cost, area and so on, the deployment optimization of PV panels is obtained to maximize the annual PV output of the whole PV power plant and minimize the construction cost. The results of case studies show that the proposed methods effectively improve the expected PV output power of the PV power plant and reduce the construction cost.

KEYWORDS

centralized PV power plant, DA-LSTM, PV output power, PV panel deployment, solar energy, terrain clustering

1 Introduction

Rapid development of renewable energy technologies such as solar is required due to climate change mitigation strategies worldwide (Dong et al., 2021). Moreover, the development of photovoltaic (PV) power technologies plays an important role in achieving the goals of emission peak and carbon neutrality (Zhang et al., 2021), and poverty alleviation (Zhang et al., 2020). However, the performance of PV systems is generally affected by the meteorological conditions (Hachicha et al., 2019; Li et al., 2021),

and geographical location (Al-Rousan et al., 2018; Cotfas and Cotfas, 2019). Furthermore, the centralized PV power plant covers a large area, and its continuous expansion also causes the problem of insufficient land. Therefore, it is necessary to make full use of the meteorological and geographical conditions in different regions to develop PV.

Different meteorological factors effect on PV output power to different extent. Therefore, the first step of centralized PV plant planning is to start with the effect caused by meteorological factors. Mekhilef et al. (2012) studied the effect of dust deposition, wind speed and relative humidity on the efficiency of solar cells. Li et al. (2020) analyzed the impacts of wind speed, wind direction, ambient temperature and solar radiation on PV considering dynamic line parameters. The main factors in the PV power calculation are the solar insolation. Ambient temperature and wind speed have transitive relation with PV power through irradiance, and the humidity and the atmospheric pressure have a negative correlation with PV output (Ziane et al., 2021). Dust is a cause of PV output power reduction, and models of dust deposition on PV modules using local meteorological events were developed in (Sengupta et al., 2021). Gowid and Massoud (2020) developed a PV maximum power point identification tool considering temperature and solar insolation. Mayer (2021) revealed the effects of the meteorological data resolution on the of simulation and optimization reliability of PV power plants. Those studies points out what and how meteorological factors affect PV output power, but most of them did not consider enough kinds of meteorological factors.

After confirming what and how meteorological factors affect the PV output power, PV output power calculation or prediction based on multiple meteorological factors has become a hot topic in recent years. Correlation analysis and regression analysis are basic methods to calculate PV power. A model of PV output power was obtained through regression analysis by selecting main factors that affect PV power in (Kim et al., 2019). Agoua et al. (2018) proposed a statistical spatio-temporal model based on correlation analysis to improve short-term forecasting of PV production. However, the methods mentioned above have relatively low efficiency and accuracy. To overcome this problem, a widely-used kind of method for PV calculation or prediction is deep learning. Considering solar radiation, sunlight, wind speed, temperature, cloud cover, and humidity, a modified long short-term memory (LSTM) is used to predict PV power in medium and long term (Son and Jung, 2020). There are also many other deep learning based methods such as methods based on convolutional neural network (Yan et al., 2021), spatiotemporal feedforward neural network (Rodríguez et al., 2022) and so on. Besides, graph modeling method is a novel method for PV power prediction. The graph modeling method is used to describe the relationship between various meteorological factors and PV power and predict the PV power, but the graph modeling method is more complex (Cheng et al., 2021). There are already various practical PV calculation methods taking meteorological factors into account, but there are few studies considering geographical conditions which also affect PV output power. To sum up, the method for PV output power calculation still needs to develop.

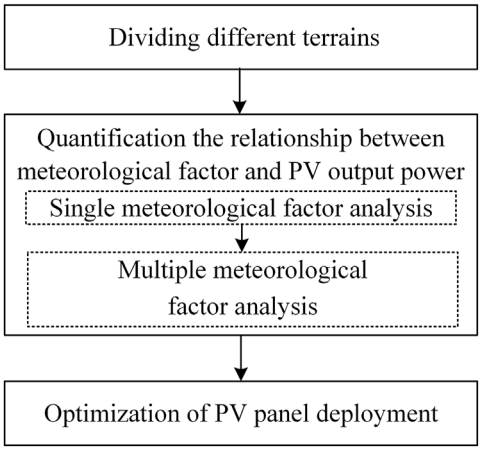


FIGURE 1
Three steps of the method for deployment of PV panels.

As mentioned above, the performance of PV is also affected by geographical conditions. For example, according to a study in a water pumping system with PV installation in Brazil, when the ratio between flow and radiation was taken into account, the monocrystalline PV system was more efficient (Nogueira et al., 2015). Polycrystalline solar module showed a better performance in semi-arid Region (Ettah et al., 2021). Ingenhoven et al. (2019) analyzed the performance loss rate of six different PV module types in five locations in Italy. Huld (2017) promoted a set of tool and data named PVMAPS, which could calculate PV performance in any region covered by the data. Due to the relationship between geographical conditions and PV performance, site selection became an important factor of PV power plant execution. To find the best location, many researches presented methods based on geographical information system (Hashemizadeh et al., 2019; Mensour et al., 2019). Although some literature dealt with the site selection of PV power plant, there were not many literature focus on the deployment of PV panels in centralized PV power plant.

In summary, the existing works lack an optimization method for the deployment of PV panels in a centralized PV power plant considering not only the geographical difference but the meteorological difference. To solve the above problems, this paper proposes an optimization method for the deployment of PV panels in a centralized PV power plant considering multiple factors. By optimizing the deployment position and quantity of PV panels, the method aims at higher PV output power and lower cost under certain capacity and approximate planning area

for a centralized PV power plant. It also provides possibility for more efficient application of PV.

The remainder of the paper is organized as follows. Section 2 introduces terrain clustering, quantification of the relationship between meteorological factors and PV output, and optimization model. Section 3 shows the result of case studies. Section 4 makes a conclusion for this paper.

2 Methodologies

As shown in Figure 1, the method for optimizing the deployment of PV panels in a centralized PV power plant under multiple factors is divided into three steps: dividing different terrains in the PV power plant according to geographical characteristics, modeling and quantifying the relationship between meteorological factors and PV output power in each terrain (including single meteorological factor and multiple meteorological factor analysis), optimizing the deployment of PV panels in the centralized PV power plant.

2.1 Terrain clustering

Before studying meteorological factors, we should first explore the influence of geographical factors. This is because the meteorological conditions in different locations are obviously different. Thus, to carefully study the relationship among PV output power, meteorological factors and geographical factors, it is necessary to combine the sub-areas with similar geographical locations to a terrain first. Moreover, the installation of PV panels is also affected by geographical factors. For example, according to the different slope direction and gradient, the installation area of unit PV panels varies, and the unit installation area will affect the subsequent optimal deployment of PV panels. Therefore, based on the above considerations, this section first divides the surrounding area of a planned centralized PV power station into several terrains according to the geographical characteristics.

Noted that the PV panels in a centralized PV power plant are often orderly concentrated in a certain area, the distance and direction between the terrain center and the gathering station (GS) are included in the geographical characteristics, so as to make each sub-area within the divided terrain roughly similar in location and more accord with the actual situation of the centralized PV power plant construction. Firstly, PV power plant are divided into several sub-areas from west to east and from north to south according to a rectangular area with the same area of A_N (determined according to the actual cases). Then the annual sunshine durations, average altitude, slope direction, slope, distance between the terrain center and the GS, the direction of the terrain center relative to the GS are used as the geographical features for dividing different terrains. Define the geographical feature of the ith sub-area as $X_i^{terrain}$ and establish the dataset $X_i^{terrain} = \{x_{i1}^{terrain}, x_{i2}^{terrain}, \cdots, x_{i6}^{terrain}\}$. Subscripts $i1$ to $i6$ represent the annual sunshine durations, average altitude, slope direction, slope, distance between the terrain center and the GS, the direction of the terrain center relative to the GS in the ith sub-area, respectively.

As one of the main unsupervised machine learning technologies, fuzzy clustering analysis is a method of analyzing and modeling important data using fuzzy theory, which establishes the uncertainty description of sample categories. The vector of fuzzy clustering algorithm can belong to multiple clusters at the same time, which can objectively reflect the real world. It has been effectively applied in many fields such as large-scale data analysis, data mining, vector quantization and so on, which is proved to have important theoretical and practical application value. With the further development of application, the research of fuzzy clustering algorithm is constantly enriched. In this section, we have chosen fuzzy c-means (FCM) algorithms due to its good performance (Benmouiza et al., 2016), and the FCM algorithm introduced by Dunn and improved by Bezdek (Nayak et al., 2015).

Using the FCM algorithm, a total of K sub-areas in the whole planning area of the PV power plant are classified into N terrains. The idea of FCM algorithm is to calculate the membership matrix $U = [u_{ij}]_{N\times K}$ and the cluster centers $V = \{v_1, v_2, \cdots, v_N\}$ from the terrain geographic feature dataset $X^{terrain} = \{X_1^{terrain}, X_2^{terrain}, \cdots, X_K^{terrain}\}$ through continuous iteration, and to minimize the function value in Eq. 1.

$$J(U,V) = \sum_{i=1}^{K}\sum_{j=1}^{N} u_{ij}^{m} d_{ij}^{2} \tag{1}$$

$$d_{ij} = \left\| X_i^{terrain} - v_j \right\| \tag{2}$$

In Eq. 1, u_{ij} is the membership degree of the ith sub-area belonging to the jth terrain; m is the membership factor; d_{ij} is the Euclidean distance from the ith sub-area to the jth cluster center, which is calculated by Eq. 2.

The calculation steps of the FCM algorithm are as follows:

Step 1. N cluster centers are randomly selected and the initial membership matrix $\boldsymbol{U}^{(0)}$ is calculated. Let $l = 1$ and start the first iteration.

Step 2. Calculate the cluster center $V^{(l)}$ and membership matrix $U^{(l)}$ of the lth iteration, and calculate the function value of $J^{(l)}$, as shown in follows.

$$v_j^{(l)} = \frac{\sum_{i=1}^{K}\left(u_{ij}^{(l-1)}\right)^{m} X_i^{terrain}}{\sum_{i=1}^{K}\left(u_{ij}^{(l-1)}\right)^{m}},\ j = 1, 2, \cdots, N \tag{3}$$

$$u_{ij}^{(l)} = \frac{1}{\sum_{i=1}^{K}\left(\frac{d_{ij}^{(l)}}{d_{ik}^{(l)}}\right)^{\frac{2}{m-1}}},\ i = 1, 2, \cdots K;\ j = 1, 2, \cdots, N \tag{4}$$

$$J^{(l)}\left(U^{(l)}, V^{(l)}\right) = \sum_{i=1}^{K}\sum_{j=1}^{N}\left((u)_{ij}^{(l)}\right)^{m}\left((d)_{ij}^{(l)}\right)^{2} \tag{5}$$

$$d_{ij}^{(l)} = \left\| X_i^{terrain} - v_j^{(l)} \right\| \tag{6}$$

Step 3. Set the termination value. Membership termination value $\varepsilon_u > 0$ or function termination value $\varepsilon_J > 0$ can be used. If $\max\left\{\left|u_{ij}^{(l)} - u_{ij}^{(l-1)}\right|\right\} < \varepsilon_u$ or $\left|J^{(l)} - J^{(l-1)}\right| < \varepsilon_J$, the iteration stops, otherwise, increase l and go to step 2.

When $u_{ij} = max_{1 \le j \le N}\left\{u_{ij}\right\}$, the ith sub-area belongs to the jth terrain and can be expressed as follows:

$$X_i^{terrain} \in D_j \tag{7}$$

2.2 Single meteorological factor analysis

Quantification the relationships between meteorological factors and PV output power in each terrain includes single meteorological factor and multiple meteorological factor analysis. There are many kinds of meteorological factors, and the impact of various meteorological factors on PV output power may be significantly different. Therefore, in order to accurately measure the relationship between meteorological factors and PV output power, and simplify the calculation complexity and time of subsequent multiple meteorological factors analysis without losing accuracy, a single meteorological factor analysis of PV output power is conducted first. The single meteorological factor analysis is to select the daily output data of an existing PV power plant near the planning area and various meteorological data collected by the corresponding meteorological station as historical data, preliminarily analyze the relationships between PV output power and various meteorological factors through Pearson correlation coefficient, and select several major meteorological factors that have a great impact on the PV output.

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \cdot \sum_{i=1}^{n}(y_i - \bar{y})^2}} \tag{8}$$

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{9}$$

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i \tag{10}$$

Where x_i is the ith time component of the meteorological factor x, y_i is the ith time component of PV output y, r is the correlation coefficient, $r \in [-1,1]$, and the closer the absolute value is to 1, the stronger the correlation between the meteorological factor x and PV output.

According to the value of correlation coefficient, k meteorological factors with the strongest correlation with PV output are selected for further mining the relationship between meteorological factors and PV output. Considering that the dimensions of various meteorological factors and PV output are not uniform, the MinMaxScaler method is used to normalize the data to improve the convergence speed of the subsequent deep learning model and reduce the error. MinMaxScaler method is shown in Eq. 11:

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{11}$$

where, x' is the normalized data, x is the original data, $x_{\min}$ and $x_{\max}$ are the minimum and maximum of the original data.

After normalizing the k main meteorological factors and PV output data, sequential feature $X^{me} = (x_1^{me}, x_2^{me}, \cdots, x_T^{me}) = (x^{me1}, x^{me2}, \cdots, x^{mek})^T$, containing k main meteorological factors, is constructed. It can be expressed by the following matrix:

$$X^{me} = \begin{bmatrix} x_1^{me1} & x_1^{me2} & \cdots & x_1^{mek} \\ x_2^{me1} & x_2^{me2} & \cdots & x_2^{mek} \\ \vdots & \vdots & \ddots & \vdots \\ x_T^{me1} & x_T^{me2} & \cdots & x_T^{mek} \end{bmatrix} \tag{12}$$

where, $x_t^{me} = (x_t^{me1}, x_t^{me2}, \cdots, x_t^{mek})$ is expressed as the above k meteorological feature sets at time t, and $x^{mep} = (x_1^{mep}, x_2^{mep}, \cdots, x_T^{mep})$ is expressed as the each value of the pth relevant meteorological variable for the whole period T.

2.3 Multiple meteorological factor analysis

The multiple meteorological factor analysis uses the dual-stage attention mechanism based LSTM (DA-LSTM). LSTM is an improved recurrent neural network. Each hidden layer is no longer a single neural network, but consists of four interconnected neural networks (forget gate, input gate, update gate and output gate). It compares the memory information with the current information and learns through self-evaluation and selectively forgetting mechanism, which can alleviate the problem of gradient vanishing and exploding in general recurrent neural network (Greff et al., 2017; Shewalkar, 2019).

2.3.1 Overview of long short-term memory

In the general LSTM networks, the long-term memory information at time t is defined as the cell state C_t. The LSTM cell receives the meteorological feature set x_t^{me} at time t and the short-term memory information h_{t-1} of cells at the previous time, and inputs the cell state C_{t-1} at the previous time into each gate as internal information. Through the forget gate f_t, the input gate i_t and the output gate o_t, the cell information is read and modified as follows:

$$f_t = \sigma\left(W_f \cdot \left[h_{t-1}, x_t^{me}\right] + b_f\right) \tag{13}$$

$$i_t = \sigma\left(W_i \cdot \left[h_{t-1}, x_t^{me}\right] + b_i\right) \tag{14}$$

$$C_t^{'} = tanh\left(W_C \cdot \left[h_{t-1}, x_t^{me}\right] + b_C\right) \tag{15}$$

$$o_t = \sigma\left(W_o \cdot \left[h_{t-1}, x_t^{me}\right] + b_o\right) \tag{16}$$

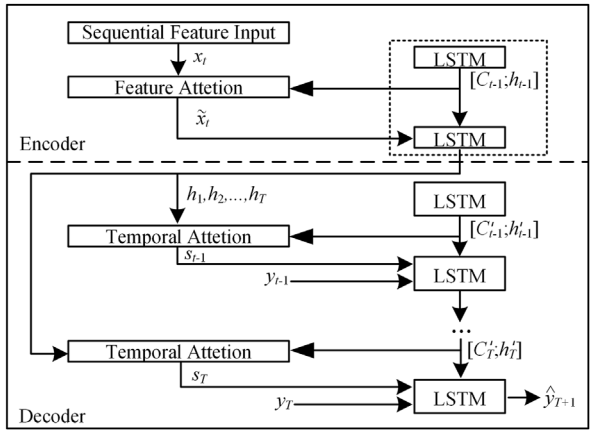


FIGURE 2
The structure of DA-LSTM.

where, W_f, W_i and W_o are the weight matrices of forget gate, input gate and output gate respectively, and b_f, b_i and b_o are the corresponding biases. W_C is the input cell state weight matrix, and b_C is the bias of the input cell state. σ is the sigmoid activation function, which transforms the output into [0,1] interval, tanh is a hyperbolic tangent activation function, which transforms the output into the [−1,1] interval. Eq. 13 represents a forget gate for judging whether the previous information is retained. Eq. 14 represents an input gate that determines whether the current information is written to the cell state at time t; Eq. 15 creates a new vector containing all possible values through the tanh activation function and adds it to the cell state; Eq. 16 represents an output gate that determines the information to be included in the output content. Input the hidden layer output h_t to obtain the PV output y_t, and the calculation is as shown in Eqs. 17–19.

$$C_t = f_t * C_{t-1} + i_t * C_t^{'} \tag{17}$$

$$h_t = o_t * tanh(C_t) \tag{18}$$

$$y_t = W_d h_t + b_d \tag{19}$$

where, * denotes the multiplication of matrix elements. W_d and B_d are the adjustable weight matrix and the bias of the output layer, respectively.

2.3.2 Dual-stage attention

The performance of only using ordinary LSTM under sudden weather and extreme weather conditions is often poor, so the attention mechanism is introduced. The attention mechanism imitates how the human brain processes information, which improves the performance of the neural network (Qu et al., 2021). In this paper, the feature attention mechanism is introduced at the encoder side, and the time attention mechanism is introduced at the decoder side to build a dual-stage attention mechanism, so as to obtain a more accurate relationship between PV output power and meteorological factors and historical information. The dual-stage attention structure can be shown in Figure 2.

Feature attention models the importance of each feature and assigns different attention to each dimension of the input (Zeng et al., 2022). In order to obtain the contribution rate of each meteorological feature to the PV output at the current time, the relevant meteorological features at time t are input into the feature attention mechanism to obtain the attention weight vector e_t:

$$e_t = V_e^T tanh(W_e[h_{t-1}, C_{t-1}] + U_e x^{mep} + b_e) \tag{20}$$

where, $e_t = (e_t^1, e_t^2, \cdots, e_t^k)$ is the combination of attention weight coefficients corresponding to each meteorological feature at time t; V_e, W_e and U_e are the weight matrix of attention mechanism, and b_e is the bias. Normalization is performed according to Eq. 21, using the Softmax function so that the sum of the feature attention weights is 1. The normalized feature attention weight is expressed as $\alpha_t = (\alpha_t^1, \alpha_t^2, \cdots, \alpha_t^p, \cdots, \alpha_t^k)$, where α_t^p is the feature attention weight value of the pth relevant meteorological feature at time t.

$$\alpha_t^p = \frac{exp(e_t^p)}{\sum_{i=1}^{k} exp(e_t^i)} \tag{21}$$

Multiplying the feature attention weight value α_t^p with the corresponding meteorological feature value x_t^{mep} to obtain the correlation feature $\tilde{x}_t^{mep}$ considering the contribution rate of different meteorological features:

$$\tilde{x}_t = \left(\alpha_t^1 x_t^{me1}, \alpha_t^2 x_t^{me2}, \cdots, \alpha_t^k x_t^{mek}\right) \tag{22}$$

By introducing the feature attention mechanism, the input to the LSTM network is no longer the original meteorological feature value, but the correlation feature weighted by the contribution rate. It adaptively strengthens the key factors affecting the PV output, weakens the less relevant meteorological factors, and improves the modeling accuracy.

Temporal attention mechanism is introduced for finding the characters of trend and cycle and determining the key node for the current PV output power adaptively (Zhu et al., 2022). In order to obtain the contribution rate of the sequential state in a period of time on the PV output at the current time, obtain the weight of the hidden state of the corresponding time sequence, and extract the historical key node, introduce the temporal attention mechanism at the decoding output side. The temporal attention weightof the hidden layer state at the current time depends on the hidden layer state $h_t = (h_t^1, h_t^2, \cdots, h_t^T)$ of the selected historical time sequence of LSTM, T is the time length of the input sequence. Taking it as an input, the temporal attention weight coefficient $l_t = (l_t^1, l_t^2, \cdots, l_t^\tau, \cdots, l_t^T)$ at the current time t is obtained, as shown in the Eq. 23:

$$l_t^\tau = V_d^T tanh\left(W_d\left[h_{t-1}^{'}, C_{t-1}^{'}\right] + U_d h_t^\tau\right) \tag{23}$$

where, V_d and W_d are the corresponding weights of the temporal attention, and U_d is the bias. The temporal attention weight is also obtained by normalizing with the Softmax function, and the comprehensive information s_t, related to the sequential state characteristics, at time t is obtained by considering the contribution rate of the information at each time in the input sequence.

$$\beta_t^\tau = \frac{exp\left(l_t^\tau\right)}{\sum_{j=1}^{T} exp\left(l_t^j\right)} \tag{24}$$

$$s_t = \sum_{\tau=1}^{T} \beta_t^\tau h_t^\tau \tag{25}$$

Combine comprehensive information s_t with original output y_t:

$$y_t^{'} = \tilde{W}\left[y_t, s_t\right] + \tilde{b} \tag{26}$$

where, $\tilde{W}$ and $\tilde{b}$ are the weights and bias input by the front-end fusion of the LSTM network. Considering the contribution rate of historical information, the hidden layer state at time t is obtained:

$$h_t^{'} = f_1\left(h_{t-1}^{'}, y_{t-1}^{'}\right) \tag{27}$$

where, f_1 is the LSTM network. The PV output at time T+1 can be expressed as:

$$\hat{y}_{T+1} = V_y^T\left(W_y\left[h_T^{'}, s_T\right] + b_w\right) + b_y \tag{28}$$

where, W_y and b_w are the weights and bias of the LSTM network. V_y and b_y are the weights and bias of the whole network before dimensional transformation.

2.4 Optimization of photovoltaic panel deployment in the power plant

2.4.1 Objective function

The optimization of PV panel deployment in the power plant takes the actual output power of the whole PV power plant and the lowest cost of PV panel deployment as the optimization objective, takes the whole year as the time scale, and makes full use of the meteorological advantages of each terrain to improve the efficiency and economy of the whole power plant.

$$F_1 = max\sum_{j=1}^{N} n_j y_{j,year} \tag{29}$$

$$y_{j,year} = \sum_{d}^{365} y_{j,d} \tag{30}$$

Eq. 29 is the objective function that maximizes total of the annual PV output power in each terrain, that is, the annual expected output of the whole planning PV power plant, where n_j represents the number of PV panels installed in the jth terrain. $y_{j,year}$ is the annual expected output power of a single PV panel in the jth terrain. The LSTM model is trained using historical data, and then the annual expected output of the unit PV panel in each terrain is calculated separately based on the actual meteorological conditions of each terrain and on the time scale of day. The annual output in the jth terrain can be expressed by Eq. 30, where $y_{j,d}$ represents the expected output of a single PV panel in the jth terrain on day d.

$$F_2 = min\sum_{j=1}^{N} C_j \tag{31}$$

$$C_j = n_j C_{sj} + C_{lj} \tag{32}$$

Eq. 31 is the objective function to minimize the deployment cost of all PV panels in the whole PV power plant, where C_j represents the total cost of installing PV panels in the jth terrain. Eq. 32 represents the installation cost of a single PV panel. Because the purchase cost of each PV panel is the same, the purchase cost can be ignored. C_{sj} represents the cost of occupying land for the installation of a single PV panel in the jth terrain, and C_{lj} represents the cost required for the line routing between the PV field in the jth terrain and the GS.

2.4.2 Constraint conditions

Since the total amount of PV panels is unchanged, only the cost required for the line routing from each terrain to the GS is considered when calculating the total line routing cost. It is assumed that the jth terrain contains J sub-areas, that is, $D_j = \left\{X_1^{terrain}, X_2^{terrain}, \cdots, X_J^{terrain}\right\}$.The line routing cost can be calculated as follows:

$$C_{lj} = C_{line,j} D_{j5} \tag{33}$$

$$C_{line,j} = \sum_{u=1}^{L} z_u C_{line,j,u} \tag{34}$$

$$\sum_{u=2}^{L} n_{u-1,limit} z_u \le n_j \le \sum_{u=1}^{L} n_{u,limit} z_u \tag{35}$$

$$\sum_{u=1}^{L} z_u \le 1 \tag{36}$$

$$D_{j5} = \frac{1}{J}\sum_{i=1}^{J} x_{i5}^{terrain} \tag{37}$$

$$C_{sj} = C_{space,j} A_j \tag{38}$$

Eq. 33 shows that the line routing cost C_{lj} is related to the distance from each terrain to the GS and the unit cost of the line, and $C_{line,j}$ is the unit cost of line routing per kilometer of the jth terrain. Eq. 34 is the calculation method of line routing cost per kilometer for the jth terrain. When the number of PV panels installed in a terrain exceeds a certain limit, the voltage level of the line needs to be increased to meet its maximum transmission power, and then the unit cost of the line will increase accordingly.

After the preliminary evaluation, it is assumed that there may be L voltage levels in the whole planning area. $n_{u,limit}$ represents the maximum number of PV panels that can be accessed with the uth voltage level. By introducing a set of 0–1 variables z_u, the range to which n_j belongs can be defined, as shown in Eqs. 35, 36. D_{j5} is the distance from the jth terrain to the GS, which is the average of the distance from the J sub-areas included in the jth terrain to the GS, as shown in Eq. 37. Eq. 38 calculates the cost of land consumption, where A_j is the actual area occupied by a unit of PV panel installed in the jth terrain.

The deployment of PV panels is also constrained by certain geographical conditions as follows:

$$n_j A_j \le A_{j,\,sum} \tag{39}$$

$$A_j = A_{PV} A_{slope} \tag{40}$$

$$A_{j,\,sum} = J A_N \tag{41}$$

$$D_{j4} \le D_{j4,limit} \tag{42}$$

Eq. 39 is the area constraint. The actual area of PV panel installation is related to the gradient and slope direction of the installation site, as shown in Eq. 40, where A_{PV} is the area of a PV panel itself, and A_{slope} is the area coefficient affected by the gradient and slope direction. $A_{j,sum}$ is the total area of the jth terrain, which is calculated by Eq. 41, and is the sum of the total J sub-areas included in the jth terrain. Also, when the gradient of a terrain is too steep, it is not proper to install a PV panel there, as shown in Eq. 42.

According to the planned capacity of PV power plant, the number of PV panels installed in the power plant can be obtained as follows:

$$n_{sum} = \sum_{j=1}^{N} n_j \tag{43}$$

3 Case studies

3.1 Parameter setting and dataset description

The case studied in this paper is the planning area of a large-scale centralized PV power plant in Southwest China, and the estimated installed capacity of the whole PV power plant is 210 MW. Obtain the PV output power data of an existing PV power plant near the planning area in 2021 and the meteorological data during the same period and at the same place. The temporal resolution of the data is 1 day. The meteorological data of each terrain in the planning area are obtained through interpolation due to the lack of actual weather station. The geographical data are obtained through SRTM data released by NASA (Jarvis et al., 2008). The parameters of the selected PV panel are: the power is 600 W and the size is 2172 mm × 1303 mm.

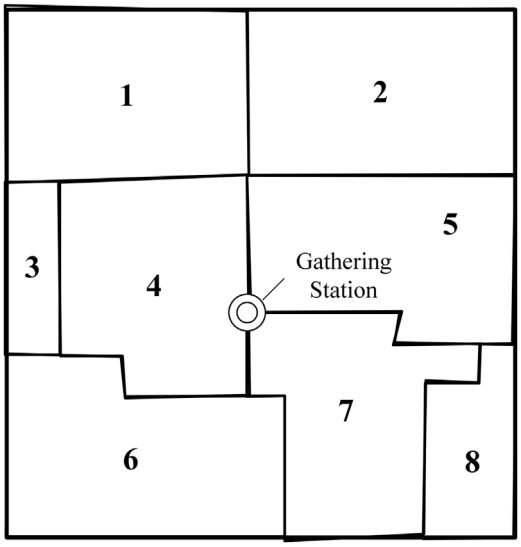


FIGURE 3
The results of terrain division.

3.2 Terrain division

The terrain in the planning area is divided by FCM algorithm. The annual sunshine durations, average altitude, slope direction, slope, the distance between the terrain center and the GS, and the direction of the terrain center relative to the GS in each sub-area are used to complete the terrain clustering. The result is shown in Figure 3.

After the terrain division, the geographic conditions of each terrain are collected and counted in Table 1. From the Table 1, it can be seen that there are great differences in geographical conditions among the terrains in the planning area. After the terrain is divided, the relationship between meteorological factors and PV output power can be analyzed more accurately, and then more accurate PV expected annual output power can be obtained.

3.3 Main meteorological factors selection

The first step of meteorological analysis is single factor analysis which is aimed to select main meteorological factors. Table 2 presents the results of single meteorological factor analyzed by Pearson correlation. As mentioned in Section 2.2, the Pearson correlation coefficient is a value between −1 and 1. The closer the absolute value of Pearson correlation coefficient is to 1, the stronger the correlation between this meteorological factor and PV output power is. If the value is positive, the meteorological factor is positively related to the PV output power, otherwise, it is negatively related.

TABLE 1 The geographic conditions of each terrain.

Terrain	N/°	E/°	Total Area/km²	Distance to GS/km	Gradient/°	Slope direction	Average altitude/m
1	28.851	99.394	4.66	2.487	10.3	North	4565
2	28.851	99.417	5.00	2.493	22.6	South	4349
3	28.834	99.394	1.06	2.071	35.1	North	4401
4	28.830	99.399	4.34	0.875	6.9	North	4584
5	28.835	99.417	4.50	1.320	7.4	North	4321
6	28.817	99.394	4.95	1.810	32.4	South	4213
7	28.817	99.417	4.04	1.391	7.8	South	4477
8	28.817	99.434	1.75	2.578	37.8	South	4248

TABLE 2 The Pearson correlation coefficient of each meteorological factor.

Meteorological factors	Pearson correlation coefficient
Total radiation intensity	0.821
Average temperature	0.311
Relative humidity	−0.691
Wind speed	0.410
Total cloud cover	−0.526
Sea level pressure	−0.108
Precipitation	−0.165
Snowfall	−0.215
Surface pressure	−0.087

Five main meteorological factors are selected depending on the Pearson correlation coefficient. In this case, the total radiation intensity, average temperature, relative humidity, wind speed and total cloud cover are selected as the main meteorological factors. Among them, the total sunshine intensity, average temperature and wind speed are positively correlated with PV output power, and the relative humidity and total cloud cover are negatively correlated with PV output power. Moreover, the total sunshine intensity is the most important factors.

3.4 Photovoltaic output power calculation

The PV output power and five main meteorological factors data are analyzed by using the DA-LSTM model, and then compared with the ordinary LSTM model. The accuracy of the model is verified by using the root mean square error (RMSE) e_{RMSE} and the mean absolute error (MAE) e_{MAE}. In order to analyze the performance of the DA-LSTM model, this paper firstly calculates the output power of the existing PV power plant. This paper compares the calculation results of the DA-LSTM model for PV output power 1 day in advance with the calculation results of the LSTM model without attention mechanism, as shown in Figure 4. It is obvious from Figure 4 that the calculation results of the DA-LSTM model are closer to the real value. Table 3 presents the RMSE and the MAE e_{MAE} of two different models. The RMSE and MAE produced by DA-LSTM model is smaller than ordinary LSTM, which means the DA-LSTM model performs better in PV output power calculation considering multiple meteorological factors.

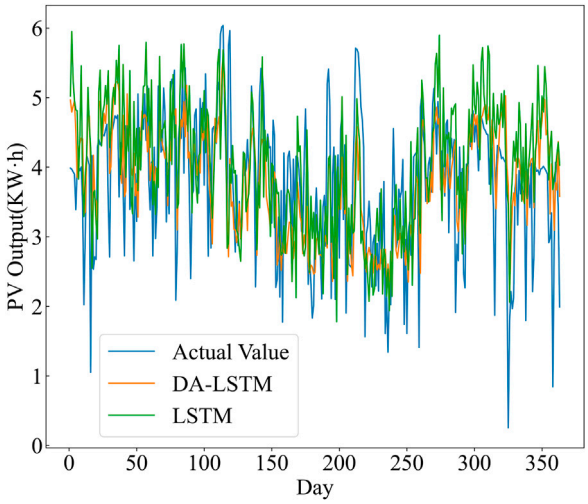


FIGURE 4
The actual value and the results of different models.

TABLE 3 The performances of different models.

Model	e_{RMSE}	e_{MAE}
DA-LSTM	0.896	0.697
LSTM	1.316	1.025

After verifying the accuracy of the DA-LSTM model, the annual expected output power of a single PV panel in each

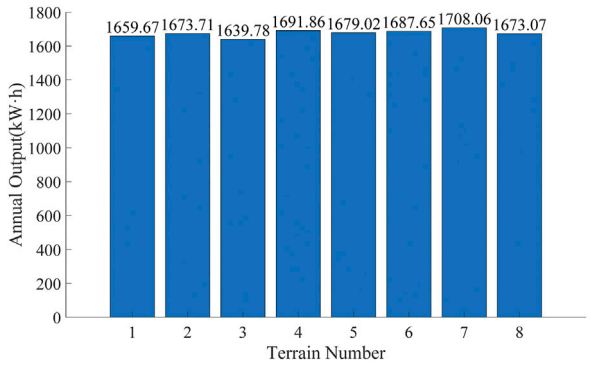


FIGURE 5
The annual expected output power of a single PV panel in each terrain.

terrain in the planning area is calculated, as shown in Figure 5. It can be seen that there is a certain difference in the output power of PV panels in different terrain. Considering the annual total sunshine intensity and annual sunshine durations of each terrain, it can be found that places with good solar resources are not necessarily the regions with the largest PV output power. That is because other meteorological factors do have a visible impact on PV output. In addition, terrain 7 has the highest PV output power. Moreover, since there are a large number of PV panels in a PV power plant, the priority of the installation site has a great impact on the output power of the final PV power plant.

3.5 Results of photovoltaic panel deployment

The deployment method proposed in this paper also considers conventional capacity constraints, area constraints and so on. However, its characteristic is that its objective function considers the maximum actual PV output power. To analyze the benefits of the method proposed in this paper, scheme B and C are set as comparison, and the proposed method is scheme A. The difference between the comparison methods is mainly in the objective function, which considers the maximum total sunshine intensity and the maximum total sunshine durations respectively. The PV resource parameters required by the three different methods are shown in Table 4. The parameters required by the two methods for comparison are easy to obtain.

Figure 6 shows the deployment results of three different methods. Obviously, the installation location and quantity of PV panels are different under three methods. Also, it is not difficult to understand that, considering the minimum cost, all deployment method will prefer to select the terrain nearest to the collection station. Terrain 4, 5, 7 are the nearest, therefore, these terrains are given high priority to install PV panels. But the actual deployment of each terrain depends on constraints and different objective functions. It is worth noting that all three methods choose to install a certain number of PV panels on terrain 4, which also proves that terrain four is the best comprehensively.

Figure 7 shows the deployment performance of three different methods. Obviously, the installation location and quantity results of three different methods are different. When the estimated PV installed capacity is given, it can be seen that the PV expected output of the purposed method is higher than that of the methods for comparison. The proposed method increases the expected output power by 0.71% and 1.71% compared with scheme B and scheme C respectively. Scheme A also lower the cost, which reduces the cost by 0.89% and 2.32% compared with scheme B and scheme C, respectively. Comparing the effects of the three deployment schemes, we can see that the method proposed in this paper improves the annual PV output power and reduces the total construction cost. In other words, the proposed method effectively improves the efficiency of solar energy

TABLE 4 The PV resource parameters of each terrain.

Terrain	Annual output/kW·h	Annual solar irradiation/MJ·m^{-2}	Annual sunshine duration/h
1	1659.67	6594	2655
2	1673.71	6586	2646
3	1639.78	6632	2622
4	1691.86	6635	2642
5	1679.02	6635	2672
6	1687.65	6628	2635
7	1708.06	6622	2605
8	1673.07	6632	2600

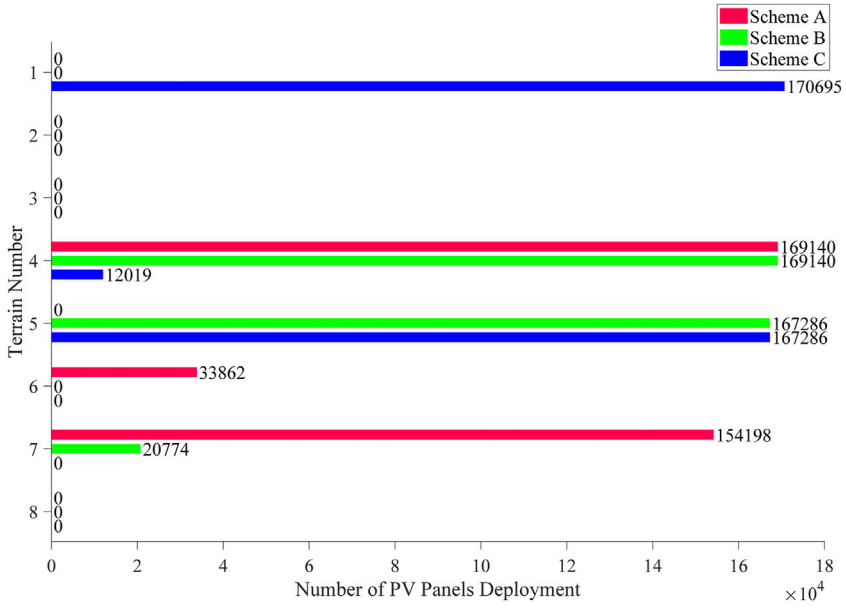


FIGURE 6
The deployment results of three different methods.

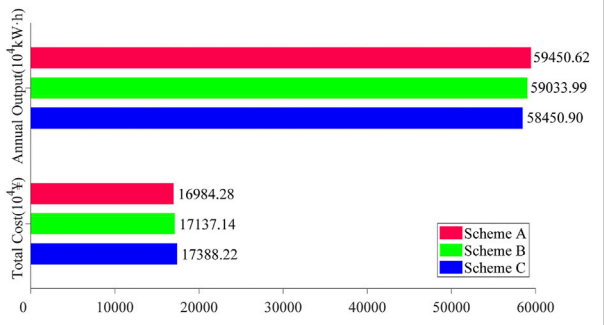


FIGURE 7
The deployment performance of three different methods.

utilization and creates higher economic and environmental value.

4 Conclusion

An optimization method for the deployment of PV panels in a centralized PV power plant under multiple meteorological and geographical factors is proposed. When deploying PV panels, the geographical and meteorological condition differences in various terrains of the planning area are fully considered, so as to obtain the maximum PV output power. Taking a planning area and an nearby existing PV power plant as examples, the data of interest are collected for case studies and the results of case studies are analyzed.

Considering that the PV panels in the centralized PV power plant are densely arranged during installation, the connectivity between the sub-areas is taken into account when dividing the different terrain in the PV power plant, so that the sub-areas in each terrain are not only similar in geographical characteristics, but also adjacent in geographical locations. The proposed terrain division method meets the needs of the PV power plant planning.

Then, Pearson correlation analysis succeeds in main meteorological selection. Moreover, DA-LSTM model has shown acceptable results in PV output power calculation, and successfully calculates the annual PV output power of various terrains in the planning area.

Finally, the advantages of the optimization method based on terrain division and PV output power calculated by DA-LSTM are proved. Compared with the comparison method, the proposed method has higher output power and less construction cost. The application of this method can improve the efficiency of PV resources utilization and create economic and environmental benefits, which promotes the use of PV technology.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

Author contributions

All authors have made substantial contributions to the conception, design of the work; analysis; and drafting of the paper. All authors read and approved the final manuscript.

Funding

This research is supported by the Science and Technology Project of State Grid Sichuan Economic Research Institute (Grant No. SGSCJY00GHJS2200046).

Conflict of interest

Author ZM was employed by the company State Grid Sichuan Electric Power Company.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Agoua, X. G., Girard, R., and Kariniotakis, G. (2018). Short-term spatio-temporal forecasting of photovoltaic power production. *IEEE Trans. Sustain. Energy* 9 (2), 538–546. doi:10.1109/tste.2017.2747765

Al-Rousan, N., Isa, N. A. M., and Desa, M. K. M. (2018). Advances in solar photovoltaic tracking systems: A review. *Renew. Sustain. Energy Rev.* 82, 2548–2569. doi:10.1016/j.rser.2017.09.077

Benmouiza, K., Tadj, M., and Cheknane, A. (2016). Classification of hourly solar radiation using fuzzy c-means algorithm for optimal stand-alone PV system sizing. *Int. J. Electr. Power & Energy Syst.* 82, 233–241. doi:10.1016/j.ijepes.2016.03.019

Cheng, L., Zang, H., Ding, T., Wei, Z., and Sun, G. (2021). Multi-meteorological-factor-based graph modeling for photovoltaic power forecasting. *IEEE Trans. Sustain. Energy* 12 (3), 1593–1603. doi:10.1109/tste.2021.3057521

Cotfas, D. T., and Cotfas, P. A. (2019). Multiconcept methods to enhance photovoltaic system efficiency. *Int. J. Photoenergy* 2019, 1–14. doi:10.1155/2019/1905041

Dong, C., Zhou, R., and Li, J. (2021). Rushing for subsidies: The impact of feed-in tariffs on solar photovoltaic capacity development in China. *Appl. Energy* 281, 116007. doi:10.1016/j.apenergy.2020.116007

Ettah, E., Ekah, U., Oyom, E., and Akonjom, N. (2021). Performance analysis of monocrystalline and polycrystalline solar panels in a semi-arid region. *Int. J. Eng. Sci. Invent.* 10 (7), 10–14. doi:10.35629/6734-1007011014

Gowid, S., and Massoud, A. (2020). A robust experimental-based artificial neural network approach for photovoltaic maximum power point identification considering electrical, thermal and meteorological impact. *Alexandria Eng. J.* 59 (5), 3699–3707. doi:10.1016/j.aej.2020.06.024

Greff, K., Srivastava, R. K., Koutnik, J., Steunebrink, B. R., and Schmidhuber, J. (2017). Lstm: A search space odyssey. *IEEE Trans. Neural Netw. Learn. Syst.* 28 (10), 2222–2232. doi:10.1109/tnnls.2016.2582924

Hachicha, A. A., Al-Sawafta, I., and Said, Z. (2019). Impact of dust on the performance of solar photovoltaic (PV) systems under United Arab Emirates weather conditions. *Renew. Energy* 141, 287–297. doi:10.1016/j.renene.2019.04.004

Hashemizadeh, A., Ju, Y., and Dong, P. (2019). A combined geographical information system and Best–Worst Method approach for site selection for photovoltaic power plant projects. *Int. J. Environ. Sci. Technol. (Tehran).* 17 (4), 2027–2042. doi:10.1007/s13762-019-02598-8

Huld, T. (2017). Pvmaps: Software tools and data for the estimation of solar radiation and photovoltaic module performance over large geographical areas. *Sol. Energy* 142, 171–181. doi:10.1016/j.solener.2016.12.014

Ingenhoven, P., Belluardo, G., Makrides, G., Georghiou, G. E., Rodden, P., Frearson, L., et al. (2019). Analysis of photovoltaic performance loss rates of six module types in five geographical locations. *IEEE J. Photovolt.* 9 (4), 1091–1096. doi:10.1109/jphotov.2019.2913342

Jarvis, A., Reuter, H. I., Nelson, A., and Guevara, E. (2008). "Hole-filled SRTM for the globe," in *CGIAR-CSI SRTM 90m database*, 4. [Dataset] Available: http://srtm.csi.cgiar.org.

Kim, G. G., Choi, J. H., Park, S. Y., Bhang, B. G., Nam, W. J., Cha, H. L., et al. (2019). Prediction model for PV performance with correlation analysis of environmental variables. *IEEE J. Photovolt.* 9 (3), 832–841. doi:10.1109/jphotov.2019.2898521

Li, Y., Wang, Y., and Chen, Q. (2020). Study on the impacts of meteorological factors on distributed photovoltaic accommodation considering dynamic line parameters. *Appl. Energy* 259, 114133. doi:10.1016/j.apenergy.2019.114133

Li, J., Chen, S., Wu, Y., Wang, Q., Liu, X., Qi, L., et al. (2021). How to make better use of intermittent and variable energy? A review of wind and photovoltaic power consumption in China. *Renew. Sustain. Energy Rev.* 137, 110626. doi:10.1016/j.rser.2020.110626

Mayer, M. J. (2021). Effects of the meteorological data resolution and aggregation on the optimal design of photovoltaic power plants. *Energy Convers. Manag.* 241, 114313. doi:10.1016/j.enconman.2021.114313

Mekhilef, S., Saidur, R., and Kamalisarvestani, M. (2012). Effect of dust, humidity and air velocity on efficiency of photovoltaic cells. *Renew. Sustain. Energy Rev.* 16 (5), 2920–2925. doi:10.1016/j.rser.2012.02.012

Mensour, O. N., El Ghazzani, B., Hlimi, B., and Ihlal, A. (2019). A geographical information system-based multi-criteria method for the evaluation of solar farms locations: A case study in souss-massa area, southern Morocco. *Energy* 182, 900–919. doi:10.1016/j.energy.2019.06.063

Nayak, J., Naik, B., and Behera, H. S. (2015). "Fuzzy C-means (FCM) clustering algorithm: A decade review from 2000 to 2014," in *Computational intelligence in data mining - volume 2*, 133–149.

Nogueira, C. E. C., Bedin, J., Niedzialkoski, R. K., de Souza, S. N. M., and das Neves, J. C. M. (2015). Performance of monocrystalline and polycrystalline solar panels in a water pumping system in Brazil. *Renew. Sustain. Energy Rev.* 51, 1610–1616. doi:10.1016/j.rser.2015.07.082

Qu, J., Qian, Z., and Pei, Y. (2021). Day-ahead hourly photovoltaic power forecasting using attention-based CNN-LSTM neural network embedded with multiple relevant and target variables prediction pattern. *Energy* 232, 120996. doi:10.1016/j.energy.2021.120996

Rodríguez, F., Galarza, A., Vasquez, J. C., and Guerrero, J. M. (2022). Using deep learning and meteorological parameters to forecast the photovoltaic generators intra-hour output power interval for smart grid control. *Energy* 239, 122116. doi:10.1016/j.energy.2021.122116

Sengupta, S., Sengupta, S., Chanda, C. K., and Saha, H. (2021). Modeling the effect of relative humidity and precipitation on photovoltaic dust accumulation processes. *IEEE J. Photovolt.* 11 (4), 1069–1077. doi:10.1109/jphotov.2021.3074071

Shewalkar, A., Nyavanandi, D., and Ludwig, S. A. (2019). Performance evaluation of deep neural networks applied to speech recognition RNN, LSTM and GRU. *J. Artif. Intell. Soft Comput. Res.* 9 (4), 235–245. doi:10.2478/jaiscr-2019-0006

Son, N., and Jung, M. (2020). Analysis of meteorological factor multivariate models for medium- and long-term photovoltaic solar power forecasting using long short-term memory. *Appl. Sci.* 11 (1), 316. doi:10.3390/app11010316

Yan, J., Hu, L., Zhen, Z., Wang, F., Qiu, G., Li, Y., et al. (2021). Frequency-domain decomposition and deep learning based solar PV power ultra-short-term forecasting model. *IEEE Trans. Ind. Appl.* 57 (4), 3282–3295. doi:10.1109/tia.2021.3073652

Zeng, Z., Jin, G., Xu, C., Chen, S., Zeng, Z., and Zhang, L. (2022). Satellite telemetry data anomaly detection using causal network and feature-attention-based LSTM. *IEEE Trans. Instrum. Meas.* 71, 1–21. doi:10.1109/tim.2022.3151930

Zhang, H., Wu, K., Qiu, Y., Chan, G., Wang, S., Zhou, D., et al. (2020). Solar photovoltaic interventions have reduced rural poverty in China. *Nat. Commun.* 11 (1), 1969. doi:10.1038/s41467-020-15826-4

Zhang, J., Zhang, B., Li, Q., Zhou, G., Wang, L., Li, B., et al. (2021). Integrated energy production unit: An innovative concept and design for energy transition toward low-carbon development. *CSEE J. Power Energy Syst.* 7 (6), 1133–1139. doi:10.17775/CSEEJPES.2021.05950

Zhu, K., Li, Y., Mao, W., Li, F., and Yan, J. (2022). LSTM enhanced by dual-attention-based encoder-decoder for daily peak load forecasting. *Electr. Power Syst. Res.* 208, 107860. doi:10.1016/j.epsr.2022.107860

Ziane, A., Necaibia, A., Sahouane, N., Dabou, R., Mostefaoui, M., Bouraiou, A., et al. (2021). Photovoltaic output power performance assessment and forecasting: Impact of meteorological variables. *Sol. Energy* 220, 745–757. doi:10.1016/j.solener.2021.04.004

frontiers | Frontiers in Energy Research

TYPE Original Research
PUBLISHED 23 January 2023
DOI 10.3389/fenrg.2022.1089921

OPEN ACCESS

EDITED BY
Youbo Liu,
Sichuan University, China

REVIEWED BY
Yue Yang,
Hefei University of Technology, China
Zihao Li,
State Grid Shanghai Electric Power Research Institute, China
Zhiyuan Tang,
Sichuan University, China

*CORRESPONDENCE
Weiye Zheng,
zhengwy13@tsinghua.org.cn

SPECIALTY SECTION
This article was submitted to Smart Grids, a section of the journal Frontiers in Energy Research

RECEIVED 04 November 2022
ACCEPTED 24 November 2022
PUBLISHED 23 January 2023

CITATION
Han Z, Li J, Wang Q, Lu H, Xu S, Zheng W and Zhang Z (2023), Deep learning-aided joint DG-substation siting and sizing in distribution network stochastic expansion planning.
Front. Energy Res. 10:1089921.
doi: 10.3389/fenrg.2022.1089921

Deep learning-aided joint DG-substation siting and sizing in distribution network stochastic expansion planning

Zhentao Han[1], Jianfeng Li[2], Qixiang Wang[1], Hao Lu[3], Siyu Xu[3], Weiye Zheng[3]* and Zixin Zhang[1]

[1]State Grid Liaoning Electric Power Company Limited, Economic Research Institute, Shenyang, China, [2]State Grid Liaoning Electric Power Supply Co., LTD., Shenyang, China, [3]School of Electrical Power Engineering, South China University of Technology, Guangzhou, China

The rapid growth of distributed generation (DG) and load has highlighted the necessity of optimizing their ways of integration, as their siting and sizing significantly impact distribution networks. However, little attention has been paid to the siting and sizing of new substations which are to be installed. This paper proposes deep learning-aided joint DG-substation siting and sizing in distribution network stochastic expansion planning. First, as the model depends on an accurate forecast, Long Short-Term Memory (LSTM) deep neural network is used to forecast DG output and load, where electricity growth rate, bidding capacity of the electric expansion, and industrial difference are all considered. Then, a two-stage stochastic mixed integer bilinear programming model was established for joint DG-substation siting and sizing under uncertainties, where multiple objective functions are comprehensively addressed. By using the Fortuny-Amat McCarl Linearization, the resultant bilinear model is equivalently transformed into a mixed integer linear program, which can be efficiently solved. Finally, stochastic power flow calculation in the IEEE 69-node system is conducted to analyze the influence of electric expansion and DG integration on the node voltage and power flow distribution of the power system. The effectiveness of the proposed method is also verified by simulation tests.

KEYWORDS

LSTM network, load forecasting, business and industrial expansion, renewable energy integration, two-stage stochastic programming, distribution network planning

1 Introduction

With the forthcoming shortage of fossil fuels, the accommodation of renewable energy is a critical topic in power systems. Although large-scale integration of DGs is favorable to promoting the development of the economy, environment, and society (Singh and Sharma, 2017), curtailment of renewable energy is still significant and remains a critical issue to date (Zheng et al., 2021; Zheng et al., 2022). On the other hand, the expansion capacity of different industries will also impact the demand side of the system.

It is left open how to reasonably plan the location and capacity of the renewables and the expanded industrial load to be integrated into the system, as the planning scheme has a huge impact on the operation of the power system.

Load forecasting and DG forecasting are important basis for power system decision-making and planning. In terms of load forecasting, due to the volatility of DG and load, how to accurately forecast the load in the presence of electric expansion is the focus of this research. The existing load forecasting research is mainly divided into statistics-based and learning-based methods, and the latter is the current mainstream method. Statistical methods mainly include multiple linear regression, autoregression, autoregressive moving average, and so on (Kim et al., 2018; Ahmad and Chen, 2019), but they can hardly deal with load data with random and dynamic development. (Yang et al., 2019). establishes a hybrid power load forecasting model by combining the autocorrelation function and least squares support vector machine in short-term power load forecasting. Compared with the benchmark model, experimental results show that this method can significantly improve forecasting accuracy. (Gul et al., 2021). adopts CNN-Bi-LSTM to process time series data sets for medium-term electricity prediction. However, the industrial difference in electricity consumption needs to be studied, while the quantitative relationship between industrial expansion capacity and the growth of load needs to be revealed. In this paper, the influence of industrial expansion will be considered in load forecasting, while the influence of direct irradiance and diffuse irradiance will be considered in DG sizing forecasting.

Although pioneering studies have investigated the siting and sizing of DG in the distribution network, little attention has been paid to installing new substations for industrial expansion. (Ho et al., 2016). proposes an optimal energy storage scheduling of DG distributed power generation system, which was formulated as a mixed integer linear program (MILP). (Vale et al., 2010). adopts the artificial neural network method to carry out distributed energy scheduling in isolated grids, and the construction of virtual power participants (VPP) can aggregate large-scale integration of DG and other distributed energy resources. (Daud et al., 2016). studies how to deploy the optimal location capacity of distributed photovoltaics. This paper considers multiple objectives such as power loss, voltage deviation, average voltage total harmonic distortion, and system average voltage decline to construct a multi-objective optimization problem, and the multi-objective optimization problem is converted into a single-objective optimization problem in a weighted way. In the research on industrial expansion, (Chen and Hsu, 1989). establishes an expert system for load allocation in the industrial expansion planning of the distribution network. The artificial language PROLOG is used to integrate the heuristic rules

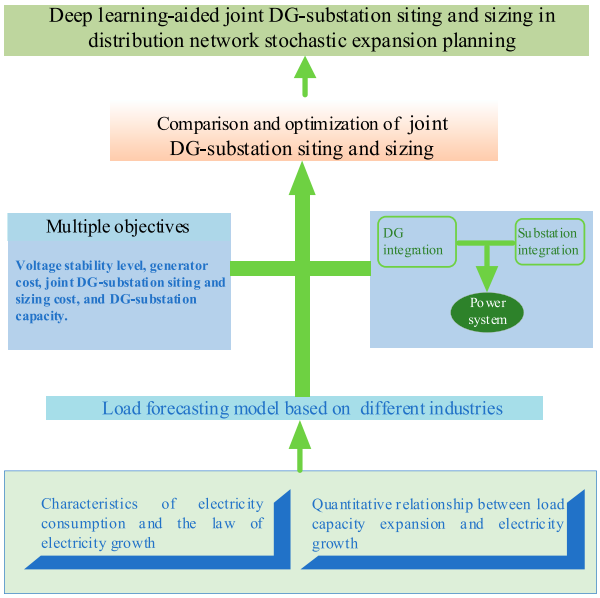


FIGURE 1
The research framework of this paper.

followed by the load allocation planner into the knowledge base, generating several appropriate load distribution schemes. (Aghaei et al., 2014). proposes a multi-stage distribution network expansion planning algorithm based on improved particle swarm optimization to ensure energy reliability and security, and realize the integration of distributed generation units into the distribution network. (Fan et al., 2020). considers the uncertainties of DG and electric vehicles and develops a comprehensive extended programming framework based on multi-objective mixed integer non-linear programming, where the Chebyshev decomposition is employed to solve the problem. However, heuristic algorithms can barely consider the uncertainty of DG and load, and their computational efficiency is generally low, which cannot satisfy the need for real-time dispatch.

Compared with the existing research on the siting and sizing of DGs, this paper tackles the scenario with industrial expansion by using the research framework in Figure 1. Utility companies process customers' applications for new substation installation and additional electricity capacity, referred to as industrial expansion and installation. Meanwhile, the main ways to achieve capacity growth include the installation of DG and substation. Therefore, this paper further explores the problem of joint DG-substation siting and sizing. The contributions are three-fold:

1) Industrial expansion data are fully employed in the LSTM network to forecast the increment load brought by the expansion.

2) A two-stage stochastic optimization model for joint DG-substation siting and sizing is established, which is reformulated into a mixed-integer linear program for an efficient solution.
3) Simulation tests are conducted on an IEEE system to prove the effectiveness of the research. Stochastic power flow is carried out to evaluate the impact of DG/substation integration on the system states, highlighting the merits of joint DG-substation siting and sizing.

Deep learning-based load forecasting and DG capacity forecasting

Load forecasting considering industrial expansion

As the industrial load is affected by industrial expansion and seasonal fluctuations, this paper improves the traditional LSTM load forecasting network, and applies the data of industrial expansion and electricity growth of different industries to the neural network, to more accurately forecast the load level under the influence of industrial expansion (Zheng et al., 2020).

In this section, the monthly load data of the pharmaceutical manufacturing industry, rubber and plastic products industry, and transportation, electrical and electronic equipment manufacturing industry in a province under the influence of industrial expansion are used to build an LSTM network, providing a basis for the load growth generated by industry expansion business.

The constructed model consists of the following steps:

1) Data selection

A large number of industrial expansion data are screened to eliminate the data caused by fault maintenance and line change and to ensure that the analyzed industrial expansion capacity generates actual load.

2) Data pre-processing

Assuming that the current time period is t, we select the industry monthly load data, annual load growth rate, and industrial expansion capacity of the past d time period for normalization and use them as the input of LSTM.

3) LSTM network structure

As shown in Figure 2, the network consists of an input layer, H sequence blocks, and an output layer. Firstly, the input layer is used to preprocess the load data, then the sequence blocks constructed by H custom LSTMs are used to extract the features of the input sequence, and finally, the output layer generates the load forecast for the industry.

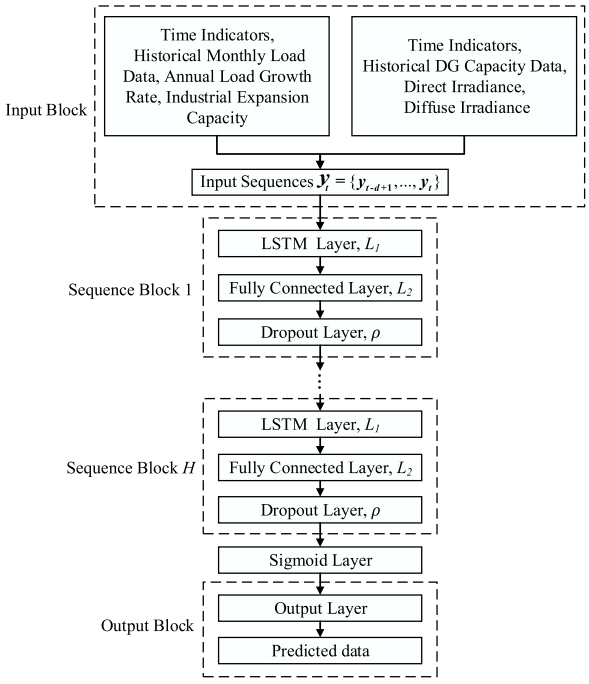


FIGURE 2
LSTM network structure diagram.

Each sequence block has the same structure, including an LSTM layer, a fully connected layer, and a dropout layer. LSTM network is a recurrent neural network that can establish the temporal correlation between previous information and the current environment, so LSTM is used as a basic component of sequence blocks. Each LSTM layer L_1 has multiple units, each of which has a memory unit $\boldsymbol{s}_\tau \in R^{L_1}$ and an input node $\boldsymbol{g}_\tau \in R^{L_1}$, an input gate $\boldsymbol{i}_\tau \in R^{L_1}$, a forgetting gate $\boldsymbol{f}_\tau \in R^{L_1}$, and an output gate $\boldsymbol{o}_\tau \in R^{L_1}$, where L_1 is a hyperparameter, and the output of each unit $\circ$ is in a hidden state. Taking the input sequence $\chi_t = \{x_{t-d+1}, ..., x_t\}$ as an example, the structure of LSTM is represented by (1)-(6) (Hochreiter and Schmidhuber, 1997), for the time period $\tau = t - \Delta + 1, t - \Delta + 2, ..., t$:

$$\boldsymbol{g}_\tau = \tanh\left(\boldsymbol{W}_{gy}\boldsymbol{x}_\tau + \boldsymbol{W}_{gh}\boldsymbol{h}_{\tau-1} + \boldsymbol{b}_g\right) \quad (1)$$

$$\boldsymbol{i}_\tau = \sigma\left(\boldsymbol{W}_{iy}\boldsymbol{x}_\tau + \boldsymbol{W}_{ih}\boldsymbol{h}_{\tau-1} + \boldsymbol{b}_i\right) \quad (2)$$

$$\boldsymbol{f}_\tau = \sigma\left(\boldsymbol{W}_{fy}\boldsymbol{x}_\tau + \boldsymbol{W}_{fh}\boldsymbol{h}_{\tau-1} + \boldsymbol{b}_f\right) \quad (3)$$

$$\boldsymbol{o}_\tau = \sigma\left(\boldsymbol{W}_{oy}\boldsymbol{x}_\tau + \boldsymbol{W}_{oh}\boldsymbol{h}_{\tau-1} + \boldsymbol{b}_o\right) \quad (4)$$

$$\boldsymbol{s}_\tau = \boldsymbol{g}_\tau \circ \boldsymbol{i}_\tau + \boldsymbol{s}_{\tau-1} \circ \boldsymbol{f}_\tau \quad (5)$$

$$\boldsymbol{h}_\tau = \tanh\left(\boldsymbol{s}_\tau\right) \circ \boldsymbol{o}_\tau \quad (6)$$

Where $\boldsymbol{h}_{\tau-1} \in R^{L_1}$ is the LSTM layer output of the historical time period $\tau - 1$, matrix $\boldsymbol{W}_{gy} \in R^{L_1 \times n}$, $\boldsymbol{W}_{gh} \in R^{L_1 \times n}$, $\boldsymbol{W}_{iy} \in R^{L_1 \times n}$,

$\boldsymbol{W_{ih}} \in R^{L_1 \times n}$, $\boldsymbol{W_{fy}} \in R^{L_1 \times n}$, $\boldsymbol{W_{fh}} \in R^{L_1 \times n}$, $\boldsymbol{W_{oy}} \in R^{L_1 \times n}$ and $\boldsymbol{W_{oh}} \in R^{L_1 \times n}$ is the weight of the network activation function input, $\boldsymbol{b_g} \in R^{L_1}$, $\boldsymbol{b_i} \in R^{L_1}$, $\boldsymbol{b_f} \in R^{L_1}$ and $\boldsymbol{b_o} \in R^{L_1}$ is the bias vector, σ is the sigmoid activation function, and $\circ$ represents Hadamard product. In deep learning, the sigmoid function is often used as the activation function of neural networks due to its monotonically increasing nature and the monotonically increasing inverse function, which maps variables between [0,1] and normalizes the output of each neuron. Since the probabilities range from 0 to 1, the sigmoid function is a good fit for models that take predicted probabilities as output. Hadamard product is a kind of matrix operation. If $A = (a_{ij})$ and $B = (b_{ij})$ are two matrices of the same order, and if $c_{ij} = a_{ij} \times b_{ij}$, then the matrix $C = (c_{ij})$ is the Hadamard product of A and B, i.e. $C = A \circ B$.

The input node, input gate, forgetting gate, and output gate are shown in Eq. 1–4. Different activation functions are used to calculate the activation of the weighted sum of input $\boldsymbol{x_\tau}$ and LSTM hidden output $\boldsymbol{h_{\tau-1}}$. These three doors regulate the flow of information to and from the storage unit $\boldsymbol{s_\tau}$. As shown in Eq. 5, the storage unit $\boldsymbol{s_\tau}$ is the Hadamard product of $\boldsymbol{g_\tau}$ and $\boldsymbol{i_\tau}$ plus the Hadamard product of $\boldsymbol{s_{\tau-1}}$ and $\boldsymbol{f_\tau}$. It remembers values in any time period, which are controlled by input nodes $\boldsymbol{g_\tau}$, input gates $\boldsymbol{s_\tau}$, and forget gates $\boldsymbol{f_\tau}$, to determine which elements in the LSTM cell should be updated, maintained, or deleted. The hidden LSTM output $\boldsymbol{h_\tau}$ is shown in Eq. 6, which is obtained by the memory unit $\boldsymbol{s_\tau}$ and the output gate $\boldsymbol{o_\tau}$. Finally, in order to improve the generalization ability of LSTM, the fully connected layer and dropout layer are added after the LSTM layer. The main purpose of the dropout layer is to prevent the overfitting of the neural network. In the forward propagation process of training, the neural network is sampled and the activation value of neurons is randomly set to 0, while the dropout layer is not used in the test process.

4) The setting of training parameters

In the above network structure, each Bernoulli random variable has probability $\rho = 1$. The fully connected layer in Eq. 6 is used to capture useful features in the target domain (Zhang et al., 2018). Meanwhile, in order to prevent overfitting, the output of the fully connected layer is randomly set to zero in the dropout layer with probability ρ.

Finally, in the training of the network, the time algorithm is used for back-propagation (Gers et al., 1999) to minimize the loss between the predicted output of the neural network $\tilde{\boldsymbol{y}}_\tau$ and the corresponding output label $\boldsymbol{y_\tau}$. During network initialization, set $\boldsymbol{s_0} = \boldsymbol{0}$ and $\boldsymbol{h_0} = \boldsymbol{0}$ (Shi et al., 2015), the weight matrix is initialized by Glorot initializer (Glorot and Bengio, 2010), the deviation of the forgetting gate is initialized to 1, and other deviations are initialized to 0.

The historical data of training is input into the above network, and the predicted load value of the current time period t is obtained through forward propagation. Then the loss of the predicted load and the true value of the label is backpropagated through the network, and then the Adam optimizer is used to update the weight and deviation of the network. The detailed training process can be found in (Gers et al., 1999).

DG capacity forecasting

Given that the output of DG is affected by some factors, such as environment, time, and so on, this section similarly uses the LSTM network to learn the historical data output by DG to predict the output capacity of DG more accurately. Therefore, as shown in Figure 2, the DG capacity data, direct irradiance, and diffuse irradiance of the same period in the province are selected for normalization and used as the input of LSTM. The network structure and training process are the same as in the previous section, and finally, the DG output data at this moment is obtained.

A multi-objective two-stage stochastic optimization model for joint DG-substation siting and sizing

The previous section forecasts the load yielded by industrial expansion and DG output, which provides the data basis for the modeling in this section. This section will use multi-objective two-stage stochastic programming to deal with the uncertainties of DGs and loads, which will be reformulated into an MILP for an efficient solution.

Objective function

A multi-objective two-stage stochastic optimization model is established. The system voltage stability and generator cost are considered in the objective function. The location and capacity of new DGs and substations are formulated as the first-stage variables, and the other variables are the second-stage variables:

$$\min f_1 = \sum_{s=1}^{N_s} \omega_s \sum_{i=1}^{N_{bus}} \left| U_{i,s} - 1 \right| \tag{7}$$

$$\min f_2 = \sum_{s=1}^{N_s} \omega_s \sum_{i=1}^{N_g} C_{Gi} P_{Gi,s} \tag{8}$$

$$\min f_3 = \sum_{i=1}^{N_{bus}} \sum_{k=1}^{N_{ext}} \sum_{s=1}^{N_s} C_{ki}^{ext} \alpha_{ki} P_{extk,s} + \sum_{i=1}^{N_{bus}} \sum_{k=1}^{N_{DG}} \sum_{s=1}^{N_s} C_{ki}^{DG} \beta_{ki} \left(P_{DGk,s} + Q_{DGk,s} \right) \tag{9}$$

$$\min\ f_4 = -\left(\sum_{s=1}^{N_s}\sum_{k=1}^{N_{ext}}\alpha_{ki}P_{extk,s} + \sum_{s=1}^{N_s}\sum_{k=1}^{N_{DG}}\beta_{ki}\left(P_{DGk,s} + Q_{DGk,s}\right)\right) \tag{10}$$

where s is the number of scenarios, ω_s is the probability of scenario s, and N_{bus} is the number of nodes in the system, N_s is the number of scenarios. $U_{i,s}$ is the square of the voltage amplitude of node i in scenario s. N_g is the number of generators in the system, $P_{Gi,s}$ is the output of the ith generator in scenario s, C_{Gi} is the cost coefficient of the ith generator. N_{ext} is the number of substation installation, C_{ki}^{ext} is the cost of the kth substation installation task at node i, α_{ki} is a binary variable, α_{ki} is 1 if the kth substation installation task is installed at node i, otherwise, α_{ki} is 0. N_{DG} is the number of DGs to be integrated, C_{ki}^{DG} is the cost of the kth DG siting task at node i, β_{ki} is a binary variable, if the kth DG siting task is at node i, β_{ki} is 1, otherwise it is 0. $P_{extk,s}$ is the expected load supplied by the kth new substation under scenario s, and $P_{DGk,s}, Q_{DGk,s}$ is the active/reactive power flowing from kth DG in node i to other adjacent nodes under scenario s. In the above objective function, f_1 reflects the voltage stability level of the system, f_2 is the generator cost, f_3 is the joint DG-substation siting and sizing cost, and f_4 is the DG-substation capacity.

The original problem can be transformed into a single-objective optimization problem by weighting the multi-objective, which can be directly solved by the mainstream solver. Therefore, this paper converts the above multi-objective optimization problem into the following single-objective optimization problem:

$$\min\ F = \sum_{i=1}^{4}\gamma_i f_i \tag{11}$$

In Eq. 11, γ_i is the corresponding weight of the ith objective function, $\sum_{i=1}^{4}\gamma_i = 1$ and the weight coefficient of each objective function can be set according to the actual demand of the project.

Network constraint

Since the joint DG-substation siting and sizing are usually in the distribution network, considering the distribution network is a radial network, LinDistFlow model is used to describe the power flow (Šulc et al., 2014):

$$P_{ji,s} + P_{DG,i,s} - P_{di} - P_{di,s}^{ext} + P_{Gi,s} - \sum_{w} P_{iw,s} = 0 \tag{12}$$

$$Q_{ji,s} + Q_{DG,i,s} - Q_{di} + Q_{Gi,s} - \sum_{w} Q_{iw,s} = 0 \tag{13}$$

$$U_{i,s} = U_{j,s} - 2R_{ij}P_{ji,s} - 2X_{ij}Q_{ji,s} \tag{14}$$

$$P_{ji} \le P_{ji,s} \le \bar{P}_{ji} \tag{15}$$

$$Q_{ji} \le Q_{ji,s} \le \bar{Q}_{ji} \tag{16}$$

$$U_i \le U_{i,s} \le \bar{U}_i \tag{17}$$

$$0 \le P_{DG,i,s} \le \bar{P}_{DG,i} \tag{18}$$

$$0 \le Q_{DG,i,s} \le \bar{Q}_{DG,i} \tag{19}$$

$$0 \le P_{di,s}^{ext} \le \bar{P}_{di}^{ext} \tag{20}$$

Among them, Equations 12, 13 are the active and reactive power balance constraints of node i, $P_{ji,s}/Q_{ji,s}$ is the active and reactive power flow between nodes j and i in scenario s, $P_{DG,i,s}/Q_{DG,i,s}$ is the active and reactive power capacity of DG integration at node i in scenario s. P_{di}/Q_{di} is the active/reactive power load of node i, $P_{di,s}^{ext}$ is the expected load supplied by the new substation at node i, and $P_{iw,s}/Q_{iw,s}$ is the active/reactive power flowing from node i to other adjacent nodes under scenario s, where w is the set of downstream nodes with respect to node i. Equation 14 is the voltage relationship between nodes j and i. Equations 15, 16 line power flow upper and lower bound constraints, $\bar{P}_{ji}/P_{ji}$ is the upper/lower bound of active power between nodes j and i, and $\bar{Q}_{ji}/Q_{ji}$ is the upper/lower bound of reactive power between nodes j and i. Eq. 17 is the voltage constraint of node, and $\bar{U}_i/U_i$ is the upper/lower bound of the square voltage amplitude of node i. Equations 18, 19 are the constraints on the active and reactive power capacity of the DG of node i integrated in the system, and $\bar{P}_{DG,i}/\bar{Q}_{DG,i}$ are the upper bound of the capacity of DG at node i. Eq. 20 is the constraint on the capacity of the substation at node i, and $\bar{P}_{di}^{ext}$ is the corresponding upper bound.

In scenario s, the additional load at node i is shown in Eqs 21, 22.

$$\begin{aligned} P_{di,s}^{ext} &= \sum_{k=1}^{N_{ext}}\alpha_{ki}P_{extk,s} \\ \alpha_{ki} &= \begin{cases} 0, & substation\ installation\ task\ k\ is\ not\ at\ node\ i \\ 1, & substation\ installation\ task\ k\ is\ not\ at\ node\ i \end{cases} \end{aligned} \tag{21}$$

$$\sum_{i=1}^{N_{bus}}\alpha_{ki} = 1 \tag{22}$$

In scenario s, the capacity of DG integrated at node i is shown in Eqs 23, 24.

$$\begin{aligned} S_{DG,i,s} &= \sum_{k=1}^{N_{DG}}\beta_{ki}\left(P_{DGk,s} + Q_{DGk,s}\right) \\ \beta_{ki} &= \begin{cases} 0, & DG\ integration\ task\ k\ is\ not\ at\ node\ i \\ 1, & DG\ integration\ task\ k\ is\ at\ node\ i \end{cases} \end{aligned} \tag{23}$$

$$\sum_{i=1}^{N_{bus}}\beta_{ki} = 1 \tag{24}$$

Among all variables, α_{ki}, β_{ki} are the first-stage variables, $U_{i,s}$, $P_{Gi,s}$, $P_{ji,s}/Q_{ji,s}$, P_{di}/Q_{di}, $P_{DG,i,s}/Q_{DG,i,s}$, $P_{extk,s}$ are the second-stage variables. Since α_{ki}/β_{ki} are binary variables, while $P_{extk,s}, P_{DGk,s}, Q_{DGk,s}$ are continuous variables, and the multiplication of the two is non-linear, this paper introduces auxiliary variables $W_{ks}^{ext}, W_{ks}^{DG}$ and uses Fortuny-Amat McCarl Linearization (Fortuny-Amat and McCarl, 1981) to deal with the problem. $W_{ks}^{ext} = \alpha_{ki}P_{dl.s}^{ext}$ is an auxiliary variable, which represents the substation capacity of scenario s integrating the system, and $W_{ks}^{DG} = \beta_{ki}\left(P_{DG,i,s} + Q_{DGi,s}\right)$ is an auxiliary variable,

which represents the active and reactive power capacity of DG integrating the system at scenario s.

$$0 \le W_{ks}^{ext} \le M \tag{25}$$

$$0 \le W_{ks}^{DG} \le M \tag{26}$$

$$W_{ks}^{ext} \le M\alpha_{ki} \tag{27}$$

$$W_{ks}^{DG} \le M\beta_{ki} \tag{28}$$

$$W_{ks}^{ext} \le P_{extk,s} \tag{29}$$

$$W_{ks}^{DG} \le S_{DG,i,s} \tag{30}$$

$$W_{ks}^{ext} \ge P_{extk,s} - M(1-\alpha_{ki}) \tag{31}$$

$$W_{ks}^{DG} \ge S_{DG,i,s} - M(1-\beta_{ki}) \tag{32}$$

Among them, constraints (25), (27), (29) and (31) are to deal with the non-linearity arising from the multiplication of binary variables and continuous variables in Eq. 21, and constraints (26), (28), (30) and (32) are to deal with the non-linear problem in Eq. 23, and the parameter M is set to a large number.

Finally, our final model using Fortuny-Amat McCarl Linearization is as follows:

$$\min\ F = \sum_{i=1}^{4} y_i f_i \tag{33}$$

$$\begin{cases} f_1 = \sum_{s=1}^{N_s} \omega_s \sum_{i=1}^{N_{bus}} \left| U_{i,s} - 1 \right| \\ f_2 = \sum_{s=1}^{N_s} \omega_s \sum_{i=1}^{N_g} C_{Gi} P_{Gi,s} \\ f_3 = \sum_{i=1}^{N_{bus}} \sum_{k=1}^{N_{ext}} \sum_{s=1}^{N_s} C_{ki}^{ext} W_{ks}^{ext} + \sum_{i=1}^{N_{bus}} \sum_{k=1}^{N_{DG}} \sum_{s=1}^{N_s} C_{ki}^{DG} W_{ks}^{DG} \\ f_4 = -\left(\sum_{s=1}^{N_s} \sum_{k=1}^{N_{ext}} W_{ks}^{ext} + \sum_{s=1}^{N_s} \sum_{k=1}^{N_{DG}} W_{ks}^{DG} \right) \end{cases} \tag{34}$$

Constraints include Eqs. 12-32. The reformulated problem is an MILP.

Case study

Case studies are conducted on the IEEE 69-node system. The deep learning part is implemented by using tensorflow 1.14.0. The MILP model is established by Yalmip, and solved by the commercial solver Gurobi. In this section, firstly, stochastic power flow is used to measure the impact of DG-substation siting and sizing on the distribution network, highlighting the merits of this research. Then, the accuracy of the load forecast under industrial expansion is tested. Finally, based on two-stage stochastic programming, the optimization results of joint DG-substation siting and sizing are analyzed.

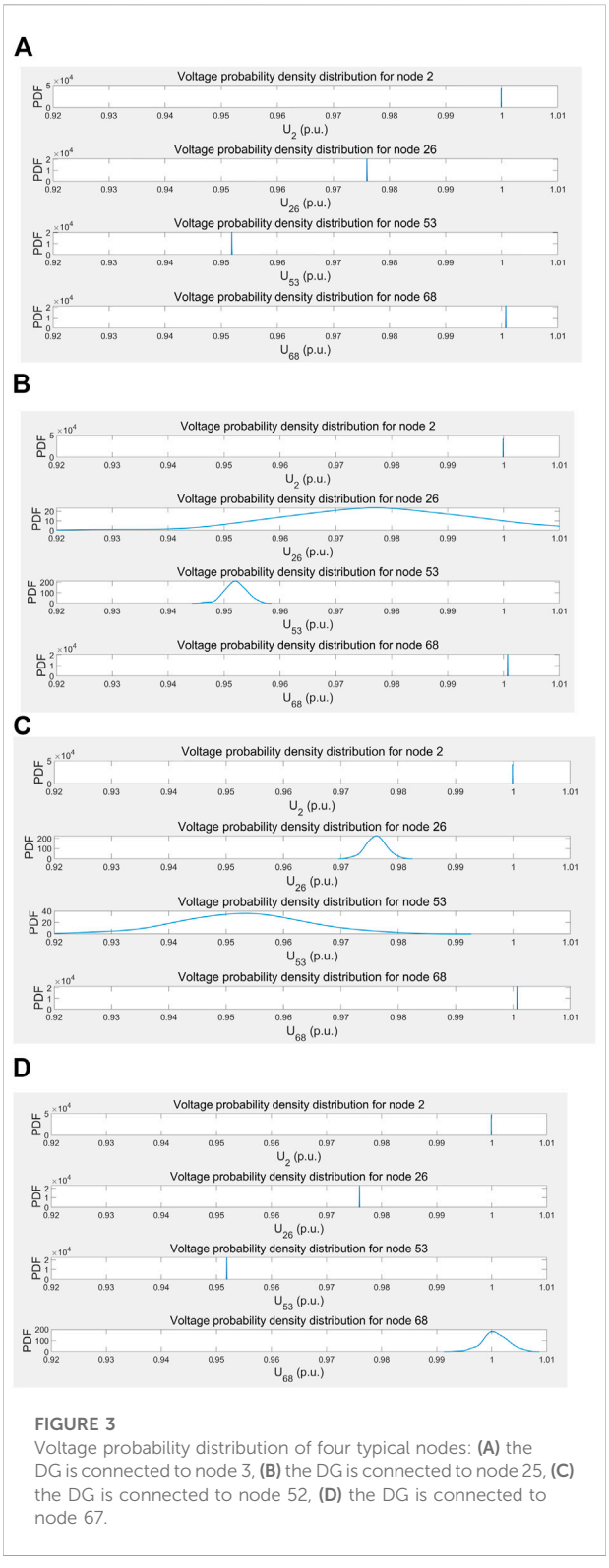


FIGURE 3
Voltage probability distribution of four typical nodes: **(A)** the DG is connected to node 3, **(B)** the DG is connected to node 25, **(C)** the DG is connected to node 52, **(D)** the DG is connected to node 67.

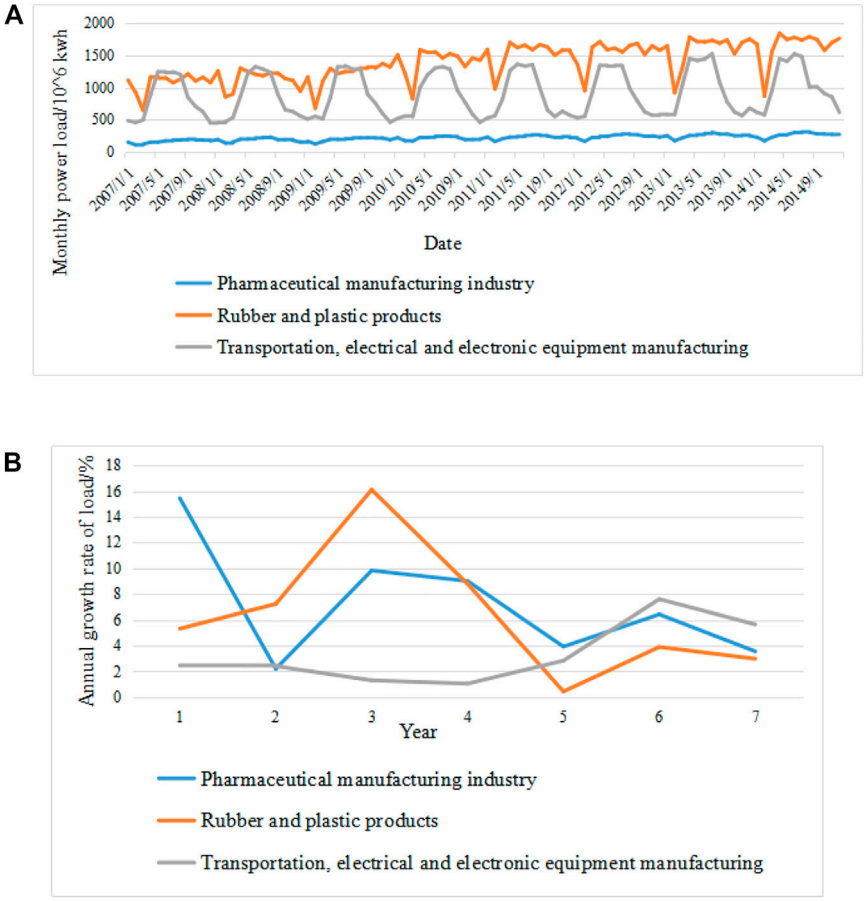


FIGURE 4
The input of LSTM: **(A)** monthly load curves for three industries, **(B)** Annual load growth curves for three industries.

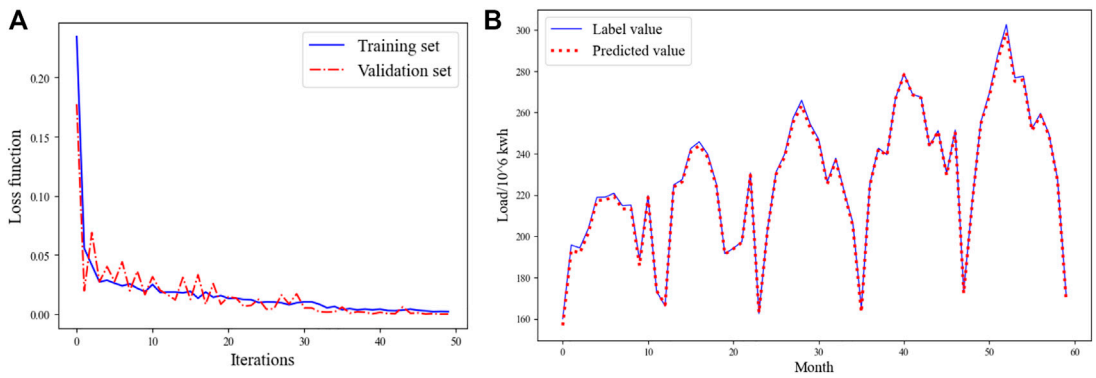


FIGURE 5
The result of pharmaceutical manufacturing industry load forecasting: **(A)** loss function curve, **(B)** comparison curve between the predicted value and the label value.

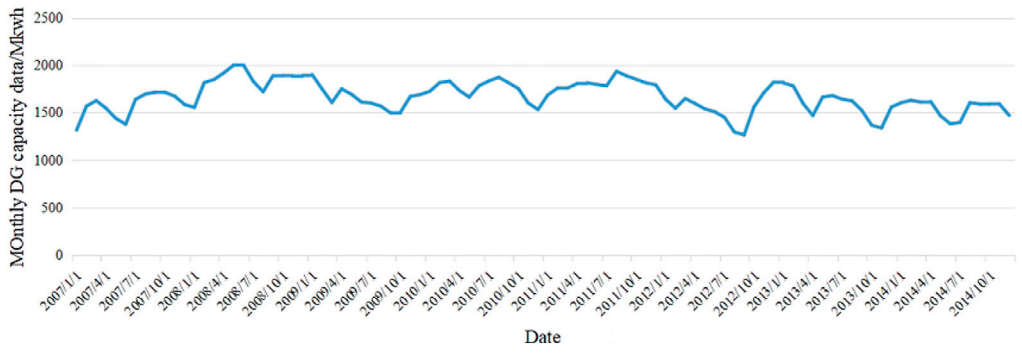


FIGURE 6
Monthly DG capacity data curve.

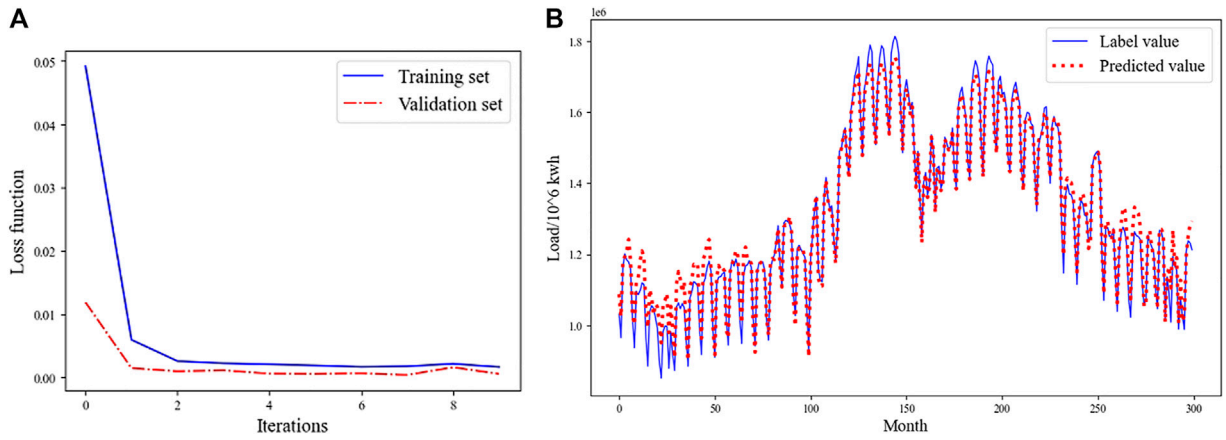


FIGURE 7
The result of DG capacity data forecasting: **(A)** loss function curve, **(B)** comparison curve between the predicted value and the label value.

TABLE 1 Multi-objective optimization of joint DG-substation siting and sizing.

	Integration point	Hosting capacity (p.u.)	Maximum capacity (p.u.)
Substation 1	Node 30	1.6478	3.0000
Substation 2	Node 32	3.0000	3.0000
Substation 3	Node 36	0.0500	0.0500
DG1	Node 47	0.3000	0.3000
DG2	Node 56	0.2000	0.2000

Impact of DG-substation siting on the distribution network under uncertainties

In this paper, the Monte Carlo method is used to measure the impact of DG integration on the distribution system. Firstly, it generated several groups of data through the probability distribution of DG integration to reflect the uncertainty of DG, then used these data to carry out Monte Carlo stochastic power flow simulation respectively. Finally, it statistically analyzed the voltage probability distribution of four typical nodes, node 2, node 26, node 53, and node 68, and analyzed the results. The sample size of the Monte Carlo simulation is set

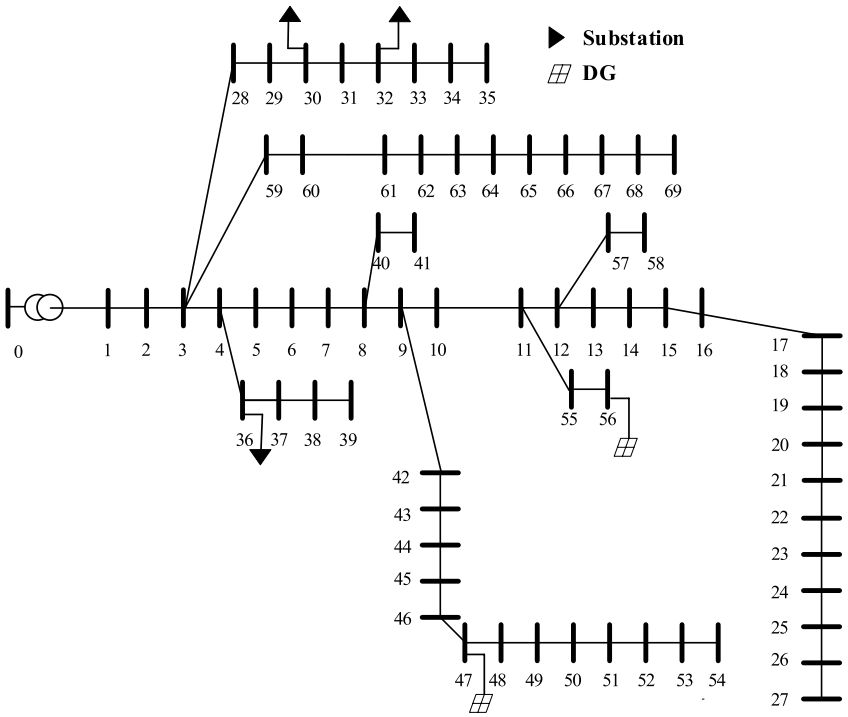


FIGURE 8
Optimal sites of DG and substation based on multiple objective optimization.

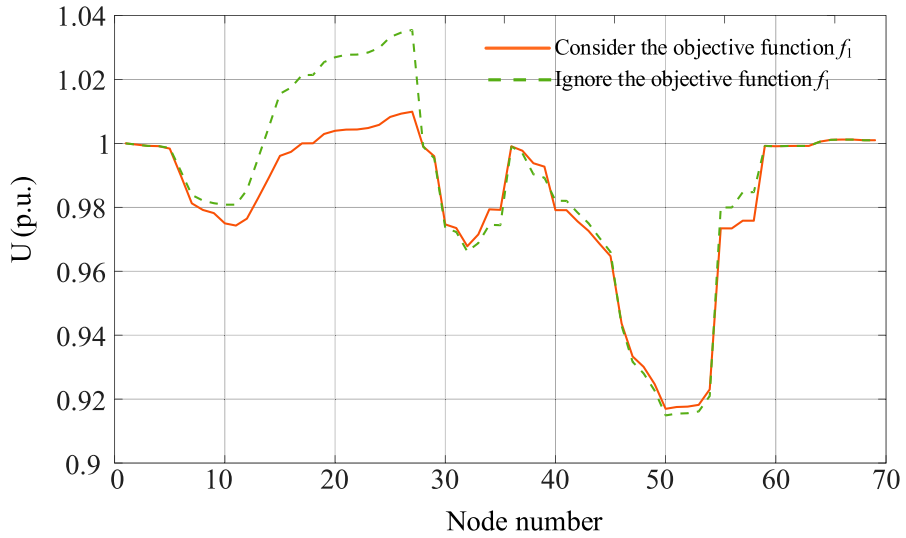


FIGURE 9
Squared node voltage in scenario one based on multiple objective optimization.

to 1000, the fluctuation of DG integration is set to the Gaussian distribution of mathematical expectation is $\mu = 0$, and the standard deviation is $\sigma = 0.3$. Figures 3A–D respectively set the DG integration locations at node 3, node 25, node 52, and node 67, and the branch impedance of node 67 is significantly smaller than that of other branches.

TABLE 2 Comparison of objective function values between the proposed joint DG-substation optimization and other methods.

Obj	Joint DG-substation	Method 1	Method 2
f_1	0.1190	0.1099	0.2417
f_2	0.1452	0.1292	0.1022
f_3	2.2020	2.3584	15.4204
f_4	−5.8166	−3.9607	−3.3087
F	−3.3503	−1.3632	12.3372

TABLE 3 Correlation analysis of the objective function in weight $\gamma = (0.25, 0.25, 0.4, 0.1)$.

		f_1	f_2	f_3	f_4
f_1	Pearson Correlation	1	0.883	−0.876	0.904
	Sig. (2-tailed)	—	0.000	0.000	0.000
	N	300	300	300	300
f_2	Pearson Correlation	0.883	1	−0.973	0.979
	Sig. (2-tailed)	0.000	—	0.000	0.000
	N	300	300	300	300
f_3	Pearson Correlation	−0.876	−0.973	1	−0.991
	Sig. (2-tailed)	0.000	0.000	—	0.000
	N	300	300	300	300
f_4	Pearson Correlation	0.904	0.979	−0.991	1
	Sig. (2-tailed)	0.000	0.000	0.000	—
	N	300	300	300	300

TABLE 4 Correlation analysis of the objective function in weight $\gamma = (0.3, 0.3, 0.3, 0.1)$.

		f_1	f_2	f_3	f_4
f_1	Pearson Correlation	1	0.855	−0.861	0.871
	Sig. (2-tailed)	—	0.000	0.000	0.000
	N	300	300	300	300
f_2	Pearson Correlation	0.855	1	−0.980	0.986
	Sig. (2-tailed)	0.000	—	0.000	0.000
	N	300	300	300	300
f_3	Pearson Correlation	−0.861	−0.980	1	−0.988
	Sig. (2-tailed)	0.000	0.000	—	0.000
	N	300	300	300	300
f_4	Pearson Correlation	0.871	0.986	−0.988	1
	Sig. (2-tailed)	0.000	0.000	0.000	—
	N	300	300	300	300

TABLE 5 Correlation analysis of the objective function in weight $\gamma = (0.2, 0.2, 0.3, 0.3)$.

		f_1	f_2	f_3	f_4
f_1	Pearson Correlation	1	0.872	−0.852	0.895
	Sig. (2-tailed)	—	0.000	0.000	0.000
	N	300	300	300	300
f_2	Pearson Correlation	0.872	1	−0.952	0.955
	Sig. (2-tailed)	0.000	—	0.000	0.000
	N	300	300	300	300
f_3	Pearson Correlation	−0.852	−0.952	1	−0.972
	Sig. (2-tailed)	0.000	0.000	—	0.000
	N	300	300	300	300
f_4	Pearson Correlation	0.895	0.955	−0.972	1
	Sig. (2-tailed)	0.000	0.000	0.000	—
	N	300	300	300	300

The simulation results are shown in Figures 3A–D, whose scenarios are explained as follows. Figure 3A: When DG is connected to node 3, there is no voltage fluctuation everywhere. Figure 3B: When DG is connected to node 25, the voltage fluctuation of node 26 is relatively obvious, the voltage of node 53 fluctuates slightly, and the voltage of the other two places has no fluctuation. Figure 3C: When DG is connected to node 52, the voltage fluctuation of node 53 is relatively significant, the voltage of node 26 fluctuates slightly, and the voltage of the other two places has no fluctuation. Figure 3D: When DG is connected to node 67, the voltage at node 68 fluctuates slightly, while the voltage at other places does not fluctuate.

The sensitivity of each node to DG integration is different. Because the location of node 2 is very close to the root node, the voltage of node 2 is almost not affected by the location of DG. The locations of nodes 26, 53, and 68 are all at the end of their branches, and their voltages are greatly affected by the location of DG.

Different locations of DG integration have different effects on the voltage of each node in the system. It can be seen from Figures 3B–D that when the DG integration causes the voltage fluctuation of a node, the voltage of the node closer to the node will be more affected by it. Therefore, the DG is connected to a location far away from the important load, which can reduce the adverse impact on the voltage of the important load.

The amplitude of node voltage fluctuation is significantly affected by the branch impedance sum of the branch where the node is located. Due to the small sum of branch impedances of node 68, the voltage in Figures 3A–C is not affected by the integration to DG basically, and it is also less affected by the integration to DG in Figure 3D. The voltage of node 26 and node 53 have obvious fluctuations when the DG is connected to node 25 (Figure 3B) and node 52 (Figure 3C), respectively. Therefore,

TABLE 6 Correlation analysis of the objective function in weight $\gamma = (0.25, 0.25, 0.25, 0.25)$.

		f_1	f_2	f_3	f_4
f_1	Pearson Correlation	1	0.843	−0.855	0.861
	Sig. (2-tailed)	—	0.000	0.000	0.000
	N	300	300	300	300
f_2	Pearson Correlation	0.843	1	−0.960	0.970
	Sig. (2-tailed)	0.000	—	0.000	0.000
	N	300	300	300	300
f_3	Pearson Correlation	−0.855	−0.960	1	−0.974
	Sig. (2-tailed)	0.000	0.000	—	0.000
	N	300	300	300	300
f_4	Pearson Correlation	0.861	0.970	−0.974	1
	Sig. (2-tailed)	0.000	0.000	0.000	—
	N	300	300	300	300

DG is connected to the branch impedance and the small branch, which can maximize the absorption of DG.

Load forecasting and DG capacity forecasting results

In load forecasting, this paper chooses a province pharmaceutical manufacturing industry, rubber and plastic products industry, and transportation, electrical and electronic equipment manufacturing industry for 3 months in load data, load growth, industry reporting for expanding capacity, and load time series as input of LSTM, as shown in Figure 4. The detailed data are available in (Han et al., 2022).

Taking the pharmaceutical manufacturing industry as an example, the training loss function curve is shown in Figure 5A. It can be seen that the loss value of the training set is reduced from 0.20 to 0.00045, the loss value of the validation set is reduced from 0.175 to 0.00032, and the iteration can converge.

Further, the model is used to generate the predicted value of the test set and compare it with the actual label value, as shown in Figure 5B. The average absolute error of the statistical data is 782.2870.

Finally, the average relative percentage error of the predicted value of the three industry loads is no more than 2.9820%, and the error value is no more than 10%, which means the accuracy meets the system's requirements.

In the prediction of DG capacity, the historical DG capacity data, direct irradiance, and diffuse irradiance of the same period are selected for normalization and used as the input of LSTM as shown in Figure 6. Since the historical data of the DG capacity, direct irradiance and diffuse irradiance in the IEEE 69-node system are not available, we use the real data from the platform Open Power System Data (Open Power System). The data in France in a period of 96 months from Jan. 2007 to Dec. 2014 are used, for their data integrity is relatively better. The loss function curve of DG capacity data forecasting is shown in Figure 7A, while the comparison curve between the predicted value and the label value is shown in Figure 7B. Their average relative error does not exceed 0.5451%, which is acceptable.

Analysis of joint DG-substation siting and sizing

The load forecasting and DG capacity forecasting results obtained in the previous section under different scenarios are normalized. In order to consider the error of the forecasting results, Gaussian distribution error is added to the forecasting results, and k-means algorithm is used to generate the load and DG capacity of three groups of typical scenarios and the probability of the scenario. The load and DG capacity of the three sets of scenarios are taken as the input of the substation and DG integration model and are denoted as the maximum value of the system. It can be known that there are three joint DG-substation siting and sizing tasks, and at the same time, two DG integration tasks are set with the same probability to optimize the joint DG-substation siting and sizing.

In order to test the effect of comprehensive consideration of the multi-objective of the proposed method, $\gamma_1 \sim \gamma_4$ in the objective functions are set to 0.25, 0.25, 0.4, and 0.1, respectively. Table 1 shows the integration points of substation and DG obtained by solving (33). The corresponding objective function values f_1, f_2, f_3 and f_4 are 0.1190, 0.1452, 2.2020 and −5.8166, respectively. Due to the comprehensive consideration of various factors, its joint DG-substation siting and sizing, DG integration locations are shown in Table 1 and Figure 8, and the solution time is 0.4533 s. It can be seen that the running time of this method is short, which can meet the real-time application in engineering.

Figure 9 shows the influence of whether the objective function f_1 is considered on the voltage of each node in the system in scenario 1. It can be seen that when f_1 is not considered, many nodes deviate from the rated voltage significantly. However, after considering the objective function f_1, the node voltage level of the system is significantly improved.

To verify the effectiveness of joint DG-substation siting and sizing, we set the control group which only considers DG siting and sizing, a total of three groups are compared, and the obtained optimization results are shown in Table 2.

1) Joint DG-substation: joint DG-substation siting and sizing are optimized.
2) Method 1: The location of the substation is fixed and the capacity is optimized, and DG siting and sizing are optimized.
3) Method 2: The location of the substation is fixed and the capacity is fixed, and DG siting and sizing are optimized.

It can be found that the voltage stability level of the system f_1 and the generator cost f_2 are not far apart in method 1 and joint DG-substation, but in method 2, the voltage stability level of the

system f_1 is relatively worse. However, the integration cost f_3 is less in the case of joint DG-substation, and the expansion capacity f_4 is larger, so the total objective function value is smaller in the end.

Further, 300 scenarios under four different weights $\gamma = (0.25, 0.25, 0.4, 0.1)$, $\gamma = (0.3, 0.3, 0.3, 0.1)$, $\gamma = (0.2, 0.2, 0.3, 0.3)$, and $\gamma = (0.25, 0.25, 0.25, 0.25)$ are selected; that is, the load fluctuated randomly at 90%–110%, and the objective function value is obtained. The correlation of each objective function value $f_1 - f_4$ under four different weights is shown in Tables 3–6. It can be found from the table that f_1, f_2 and f_4 are significantly positively correlated with each other, while f_3 is significantly negatively correlated with f_1, f_2 and f_4.

Therefore, the scheme of joint DG-substation siting and sizing determined by this method can make the power flow distribution of the distribution network reasonable and the voltage level close to the rated voltage by optimizing the integration location, and also reduce the system operation cost and the joint DG-substation siting and sizing cost to some extent.

Conclusion

To answer the call of industrial expansion, this paper proposes deep learning-aided joint DG-substation siting and sizing in distribution network stochastic expansion planning. Industrial expansion data are fully employed in the LSTM network to forecast the increment load brought by the expansion. A two-stage stochastic optimization model for joint DG-substation siting and sizing is established, which is reformulated into a mixed-integer linear program for an efficient solution. Simulation tests are conducted on an IEEE system to prove the effectiveness of the research. Stochastic power flow is carried out to evaluate the impact of DG/substation integration on the system states, highlighting the merits of joint DG-substation siting and sizing. Case studies show that the forecasting results meet the accuracy requirements, and the proposed siting and sizing method is computationally efficient. It can reduce the total cost of system operation as well as alleviate the system voltage fluctuation. In future work, we will investigate an objective manner to determine the weights for multiple objective functions.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/supplementary material.

Author contributions

ZH: Conceptualization, Methodology, Resources, Writing–Original Draft; JL: Conceptualization, Methodology, Writing–Review and Editing; QW: Methodology, Data Curation, Writing–Review and Editing; HL: Software, Writing–Original Draft; SX: Validation, Writing–Original Draft; WZ: Funding acquisition, Project administration, Writing–Review and Editing; ZZ: Writing–Review and Editing.

Funding

This work is supported by the Management Scientific and Technological Project of State Grid Liaoning Electric Power Supply Co. LTD. under Grant No. 2022YF-57.

Conflict of interest

Author JL was employed by the company State Grid Liaoning Electric Power Supply Co., LTD.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The authors declare that this study received funding from State Grid Liaoning Electric Power Supply Co. LTD. The funder had the following involvement in the study: study design.

Publisher's note

References

Aghaei, J., Muttaqi, K. M., Azizivahed, A., and Gitizadeh, M. (2014). Distribution expansion planning considering reliability and security of energy using modified PSO (Particle Swarm Optimization) algorithm. *Energy* 65, 398–411. doi:10.1016/j.energy.2013.10.082

Ahmad, T., and Chen, H. (2019). Nonlinear autoregressive and random forest approaches to forecasting electricity load for utility energy management systems. *Sustain. Cities Soc.* 45, 460–473. doi:10.1016/j.scs.2018.12.013

Chen, J. L., and Hsu, Y. Y. (1989). An expert system for load allocation in distribution expansion planning. *IEEE Power Eng. Rev.* 9 (7), 77–78. doi:10.1109/mper.1989.4310835

Daud, S. b., Kadir, A. F. A., Gan, C. K., Mohamed, A., and Khatib, T. J. (2016). A comparison of heuristic optimization techniques for optimal placement and sizing of photovoltaic based distributed generation in a distribution system. *Sol. Energy* 140, 219–226. doi:10.1016/j.solener.2016.11.013

Fan, V. H., Dong, Z., and Meng, K. (2020). Integrated distribution expansion planning considering stochastic renewable energy resources and electric vehicles. *Appl. Energy* 278, 115720. doi:10.1016/j.apenergy.2020.115720

Fortuny-Amat, J., and McCarl, B. (1981). A representation and economic interpretation of a two-level programming problem. *J. Operational Res. Soc.* 32 (9), 783–792. doi:10.2307/2581394

Gers, F. A., Schmidhuber, J., and Cummins, F. (1999). "Learning to forget: Continual prediction with LSTM," in 1999 Ninth International Conference on Artificial Neural Networks ICANN 99, Edinburgh, UK, 07-10 September 1999, 850–855. (Conf. Publ. No. 470). doi:10.1049/cp:19991218

Glorot, X., and Bengio, Y. (2010). "Understanding the difficulty of training deep feedforward neural networks," in Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics, PMLR, Sardinia, Italy, May 13–May 15, 2010, 249–256.

Gul, M. J. J., Urfa, G. M., Paul, A., Moon, J., Rho, S., and Hwang, E. (2021). Mid-term electricity load prediction using CNN and Bi-LSTM. *J. Supercomput.* 77, 10942–10958. doi:10.1007/s11227-021-03686-8

Han, Z., Li, J., Wang, Q., Lu, H., Xu, S., Zheng, W., et al. (2022). Detailed data of the test system. Available at: https://pan.baidu.com/s/1174VbXVrqHA0ICoj5EnwOQ?pwd=uo3b.

Ho, W. S., Macchietto, S., Lim, J. S., Hashim, H., Muis, Z. A., and Liu, W. H. (2016). Optimal scheduling of energy storage for renewable energy distributed energy generation system. *Renew. Sustain. Energy Rev.* 58, 1100–1107. doi:10.1016/j.rser.2015.12.097

Hochreiter, S., and Schmidhuber, J. (1997). Long short-term memory. *Neural Comput.* 9 (8), 1735–1780. doi:10.1162/neco.1997.9.8.1735

Kim, J., Cho, S., Ko, K., and Rao, R. R. (2018). "Short-term electric load prediction using multiple linear regression method," in 2018 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm), Aalborg, Denmark, 9-31 October 2018, 1–6. doi:10.1109/SmartGridComm.2018.8587489

Open Power System (n.d.). Open power system data. Available at: https://open-power-system-data.org (Accessed October 6, 2022).

Shi, X., Chen, Z., Wang, H., Yeung, D.-Y., Wong, W.-K., and Woo, W.-c. (2015). Convolutional LSTM network: A machine learning approach for precipitation nowcasting. NIPS. Available at: https://arxiv.org/abs/1506.04214.

Singh, B., and Sharma, J. (2017). A review on distributed generation planning. *Renew. Sustain. Energy Rev.* 76, 529–544. doi:10.1016/j.rser.2017.03.034

Šulc, P., Backhaus, S., and Chertkov, M. (2014). Optimal distributed control of reactive power via the alternating direction method of multipliers. *IEEE Trans. Energy Convers.* 29 (4), 968–977. doi:10.1109/tec.2014.2363196

Vale, Z. A., Faria, P., Morais, H., Khodr, H. M., Silva, M., and Kadar, P. (2010). "Scheduling distributed energy resources in an isolated grid — an artificial neural network approach," in IEEE PES General Meeting, Minneapolis, MN, USA, 25-29 July 2010, 1–7. doi:10.1109/PES.2010.5589701

Yang, A., Li, W., and Yang, X. (2019). Short-term electricity load forecasting based on feature selection and Least Squares Support Vector Machines. *Knowledge-Based Syst.* 163, 159–173. doi:10.1016/j.knosys.2018.08.027

Zhang, C.-L., Luo, J.-H., Wei, X.-S., and Wu, J. (2018). "Defense of fully connected layers in visual representation transfer," in *Advances in multimedia information processing – pcm 2017* (Cham: Springer International Publishing), 807–817.

Zheng, W., Hou, Y., and Li, Z. (2021). A dynamic equivalent model for district heating networks: Formulation, existence and application in distributed electricity-heat operation. *IEEE Trans. Smart Grid* 12 (3), 2685–2695. doi:10.1109/tsg.2020.3048957

Zheng, W., Huang, W., and Hill, D. J. (2020). A deep learning-based general robust method for network reconfiguration in three-phase unbalanced active distribution networks. *Int. J. Electr. Power & Energy Syst.* 120, 105982. doi:10.1016/j.ijepes.2020.105982

Zheng, W., Zhu, J., and Luo, Q. (2022). Distributed dispatch of integrated electricity-heat systems with variable mass flow. *IEEE Trans. Smart Grid*, in press. doi:10.1109/tsg.2022.3210014

frontiers | Frontiers in Energy Research

TYPE Original Research
PUBLISHED 02 February 2023
DOI 10.3389/fenrg.2022.1035193

OPEN ACCESS

EDITED BY
Youbo Liu,
Sichuan University, China

REVIEWED BY
Liansong Xiong,
Xi'an Jiaotong University, China
Rui Wang,
Northeastern University, China

*CORRESPONDENCE
Tao Chen,
gwxmtgy2022@163.com

SPECIALTY SECTION
This article was submitted to Smart Grids, a section of the journal Frontiers in Energy Research

RECEIVED 02 September 2022
ACCEPTED 21 October 2022
PUBLISHED 02 February 2023

CITATION
Chen T, Feng D, Wu X, Lu C and Xu B (2023), Modeling and stability analysis of interaction between converters in AC-DC distribution systems.
Front. Energy Res. 10:1035193.
doi: 10.3389/fenrg.2022.1035193

Modeling and stability analysis of interaction between converters in AC-DC distribution systems

Tao Chen*, Depin Feng, Xiaobin Wu, Changlu Lu and Bing Xu

State Grid Shandong Electric Power Company Linyi Power Supply Company, Linyi, China

With the development and promotion of AC-DC distribution systems, the stability of multi-terminal systems has drawn much interest. This paper analyzes the instability phenomena in AC-DC distribution systems by considering the interactions between converters. The typical structure and control strategy of double-terminal AC-DC distribution systems under master-slave control is introduced. Then an AC-DC distribution system model is developed taking into account the interaction of the converter control system. The resonant spikes generated by the interaction between the converters are used as the reference object of the interaction to analyze the influence of the converter control parameters and the system electrical parameters on the system stability and interaction. Finally, a double-terminal AC/DC distribution system model is built using PLECS and RTBOX for simulation and semi-physical experimental verification.

KEYWORDS

DC power distribution system, interaction, system stability, hardware-in-the-loop experiment, master-slave control

1 Introduction

With the large-scale integration of photovoltaic, electric vehicles, LED lighting and other DC equipment, the traditional AC distribution system is facing a series of complex problems, such as diversification of power supply types and load demands, multi-level transformation links, etc. The power supply quality, reliability and efficiency are increasingly affected. The AC/DC hybrid distribution systems based on flexible DC technology can more flexibly integrate distributed generations and DC loads, reduce power conversion links and improve power supply efficiency (Li et al., 2019). In particular, the flexible DC technology can realize independent decoupling control of active and reactive power, which can effectively improve power supply quality and ensure power supply reliability. It is more suitable for the development of modern distribution systems (Zhang et al.,2017).

The flexible DC distribution system realizes interconnection with AC system and integration of distributed generations and loads mainly by power electronic devices. The AC-DC distribution systems may have impedance mismatch between different converter port impedances (Huang et al., 2020), between converter port impedances and line impedances (Tao et al., 2017; Wu et al., 2017), and between converter port impedances

and grid impedances (Nian et al., 2020), i.e., The AC-DC distribution systems has stability problems on both DC sides and AC sides. Many analytical methods have been proposed to study the stability of grid-connected systems with multiple VSCs, including small signal analysis based on state space, frequency domain analysis based on impedance model and single-input-single-output (SISO) transfer function-based analysis.

The small signal analysis method based on state space can obtain the main variables affecting the stability of grid-connected systems by calculating the characteristic roots and participation factors of the state space. But with the increase of VSC, the system order increases, resulting in the difficulty of modeling. The literature (Qiu et al., 2022) used eigenvalue analysis to derive the quantitative evaluation index of the participation factor, and analyzed the relationship between oscillation modes and state variables in the wind farm with a MMC-HVDC transmission system. The correlation between each oscillation mode, wind farm, and participation factor of the MMC system are studied, which laid the foundation for the development of broadband oscillation suppression strategy. In (Wu et al., 2018), a current in-loop transfer function considering the effects of grid impedance and PLL was developed, and the destabilization mechanism of the system when the PLL and current in-loop bandwidths were close to each other under a weak grid was analyzed. The literature (Yuan et al., 2016) analyzed the effect of control parameters as well as main circuit parameters on stability based on state space eigenvalues. The literature (Deng and Fang, 2022) derived a detailed state space model of GFMCs and performed a sensitivity analysis to reveal the instability mechanism, by proposing a dominant compensator to improve the system inertia and frequency regulation. In the literature (Zhao et al., 2022), an active voltage regulation control strategy was proposed for large-scale renewable energy grid-connected power systems using a small-signal model, and sag control characteristics were proposed to improve the system's ability to consume distributed generations. The literature (Xie et al., 2022) proposed a unified modeling scheme (UMS) for MMC systems in a synchronous rotation (dq) reference frame, which used a small-signal modeling approach and could obtain a nonlinear state-space model of the overall MMC system by configuring and connecting the inputs and outputs of the state-space models of each subsystem. It could avoid the direct derivation of the matrix elements of the whole system.

The impedance analysis method can be considered as a frequency domain small signal analysis method, and the system is equated to a positive and negative sequence impedance or dq domain impedance-based model through the port input and output information of the VSC grid-connected system, and the Nyquist criterion is used to determine the system stability. However, for multi-VSC systems, multiple input and multiple output Nyquist criterion is required to analyze the system stability, which loses the advantage of analyzing stability based on single input and single output Nyquist criterion (Yin et al., 2021). In the literature (Feng et al., 2022), a full-order small-signal impedance model of the CTPS-based DAB converter is derived, and the influence of DAB converter circuit parameters on system stability is fully revealed by Bode plots and Nyquist plots, and then an optimization criterion for DAB circuit parameters is proposed. The literature (Liu et al., 2020) analyzed the DC oscillation problem of a double-terminal flexible system using the impedance method. The literature (Xue et al., 2020) analyzed the degree of influence of each parameter on the DC port impedance under different control modes. The literature (Li et al., 2022) used small signal analysis and impedance method to analyze the influence of PLL and voltage loop controller on the stability of the system in grid-connected inverter, and proposed a voltage feedforward PLL interference compensation method and an adaptive bandwidth frequency low-pass filter (LPF) to improve the system stability.

In addition, literature (Wu et al., 2019) proposed the stability analysis method of grid-connected VSC based on SISO open-loop transfer function, both of which embed the q-axis dynamics into the d-axis dynamics. Due to the large number of variables in the VSC system, the SISO models obtained by such methods are more complicated and not suitable for stability mechanism analysis. In the literature (Zhang et al.,2018; Zhang et al., 2019; Wang et al., 2020), a SISO model was established by decoupling the d-axis control and q-axis control into a unit negative feedback model containing the same control structure, which can intuitively explain the mechanism of the influence of each control link of the VSC on stability while simplifying the stability modeling and analysis process.

In general, the current research has studied the stability of AC-DC microgrids and low-voltage AC-DC systems from several aspects, mainly focusing on sag control and AC-side stability, using analysis methods based on small-signal analysis and impedance model of frequency domain analysis. However, for AC-DC distribution systems under master-slave control, the interaction dynamics, key parameters, destabilization mechanisms and effects on DC voltage stability between converters and DC systems still lack clear explanations. In this paper, through an in-depth analysis of the influence of each control parameter and electrical parameter on the interaction and stability in AC-DC distribution systems under master-slave control, the main contributions are summarized as follows.

(1)A model considering the interaction between converter control systems is developed by decomposing the control block diagram of the main converter under master-slave control into a DC voltage control part and a DC feedback part. On the premise of ensuring the applicability of the Nyqusit criterion of single input and single output system, the impedance analysis method is used for the DC side analysis, which reduces the difficulty of system modeling

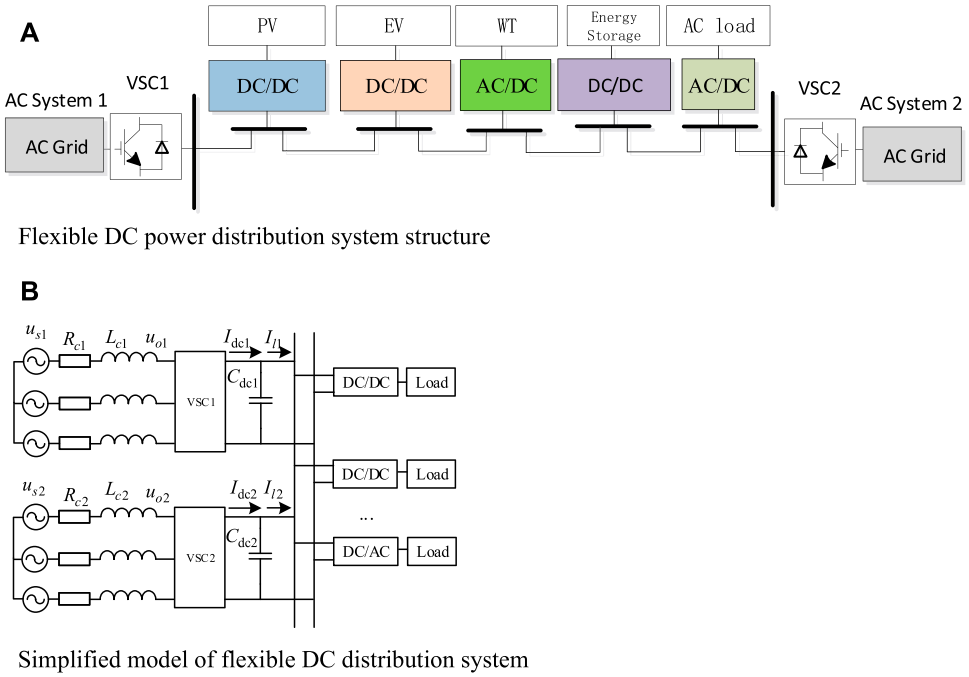


FIGURE 1
Flexible DC power distribution system structure. **(A)** Flexible DC power distribution system structure. **(B)** Simplified model of flexible DC distribution system.

and facilitates the extension of the analysis of system stability under different conditions.

(2) The stability assessment is performed by the established system transfer function, and the differences between the original model transfer function and the interaction model transfer function are compared. The resonant spikes generated by the interaction between converters are used as reference objects to analyze the interactions existing in the AC-DC distribution system. And the effects of each control parameter and electrical parameter on the spike amplitude, resonant frequency and system stability are obtained.

(3) A simulation model of a double-terminal AC-DC distribution system under master-slave control is established in PLECS, and hardware-in-the-loop experiments are conducted by RT Box, which proves the validity of theoretical analysis and simplified analysis. It also has important reference significance for engineering practice.

The structure of the paper is as follows: Section 2 presents the typical structure, control strategy and overall transfer function of the AC-DC distribution system. The system model for counting and inter-converter interactions is developed in Section 3. Section 4 analyzes the effects of different system parameters and electrical parameters on the interaction and system stability based on the proposed interaction model. The corresponding experimental validation is given in Section 5. Conclusions are drawn in the last section.

2 AC-DC distribution system model

2.1 Topology of AC-DC power distribution system

In an AC-DC distribution network, different areas of the AC system can be interconnected by a DC network. This structure can realize the multi-terminal interconnection of AC feeders from different substations, and can be further extended to systems of higher voltage levels. In terms of network topology, the current AC-DC distribution system is mainly powered by double-terminal power supply and multi-terminal power supply. In this paper, a two-terminal power supply system (Jiang et al., 2021) is used as the object of analysis. Figure 1A gives a two-terminal topology of an AC-DC distribution system, in which two AC systems are interconnected through VSC1 and VSC2, and each AC system is set as an ideal three-relative symmetric system.

The operation control mode of the AC-DC distribution system can be divided into peer-to-peer control and master-slave control, and the master-slave control mode is adopted in this paper. VSC1 is selected as the master station and adopts the constant DC voltage

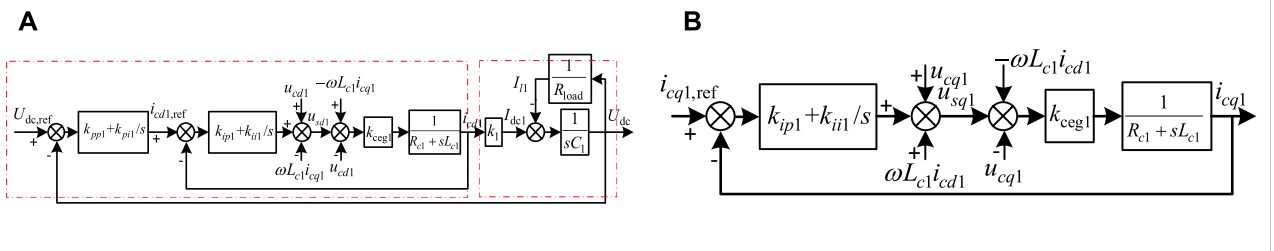


FIGURE 2
Block diagram of VSC1 control strategy. **(A)** Block diagram of double-closed control for constant DC voltage. **(B)** q-axis control block diagram of current loop.

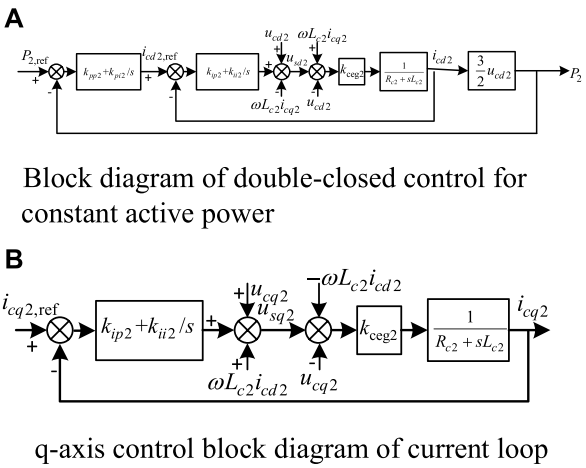


FIGURE 3
Block diagram of VSC2 control strategy. **(A)** Block diagram of double-closed control for constant active power. **(B)** q-axis control block diagram of current loop.

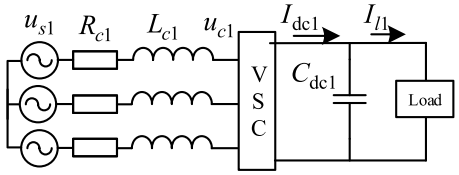


FIGURE 4
Simplified model of AC-DC distribution system.

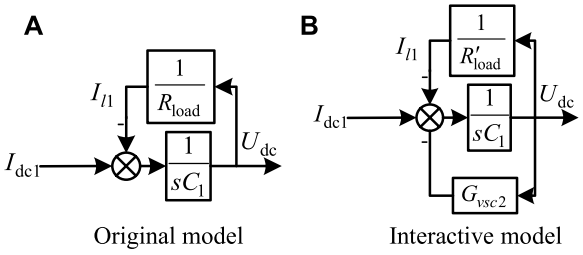


FIGURE 5
DC equivalent section. **(A)** Original model. **(B)** Interactive model.

control strategy to ensure the DC grid voltage stability. The slave station VSC2 adopts the constant active power control strategy. The load side uses the converter to realize the bidirectional flow of energy and voltage transformation. The power loss of each power electronic converter is neglected, and the equivalent circuit structure of the double-terminal AC-DC distribution system is shown in Figure 1B.

2.2 Control mode of AC-DC distribution system

The control block diagrams of VSC1 constant DC voltage control and VSC2 constant power control are shown in Figures 2, 3, respectively, and their control strategies can be seen as consisting of two parts: internal control system and DC feedback. Under normal circumstances, the load of AC-DC distribution system has constant power characteristics unlike the commonly used AC constant impedance load. For VSC2, which uses constant active power control from the converter station, it also has constant power load characteristics and is generally regarded as a constant power load. Therefore, the rest of

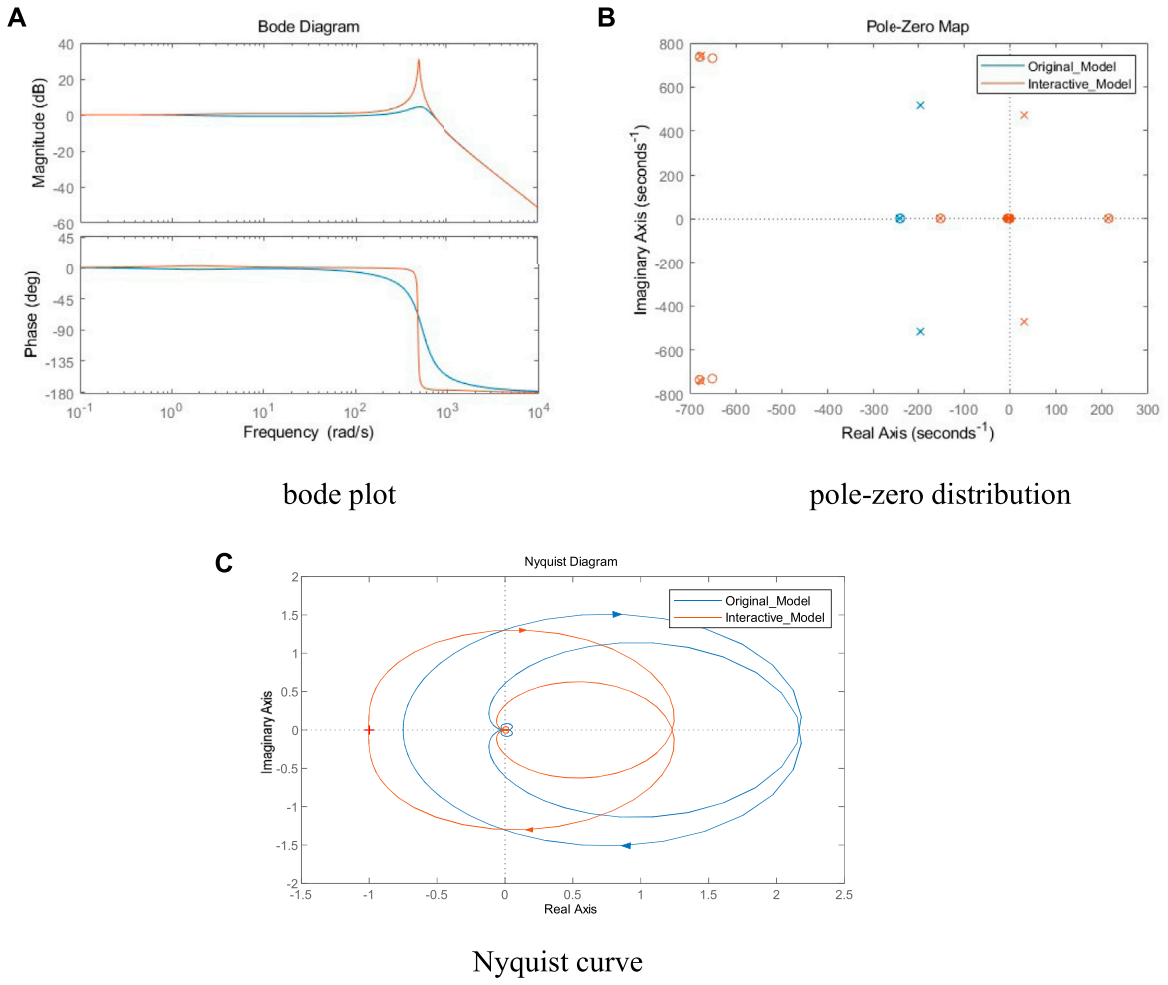


FIGURE 6
Comparison of original model and interactive model. **(A)** bode plot. **(B)** pole-zero distribution. **(C)** Nyquist curve.

the constant power control ports except the master station are generally equated as constant power loads, which are considered as series-parallel combinations of multiple components, as shown in the simplified three-phase equivalent model in Figure 4 (Peng et al., 2021).

In this paper, the main converter station control system is used as the framework to establish the SISO transfer function model of a double-terminal AC-DC distribution system, and only a single input and single output Nyqusit criterion is required to analyze the stability of the grid-connected system. In order to simplify the calculation and facilitate the extension, the impedance analysis method is used for the DC side, which effectively reduces the difficulty of system modeling. Based on the proposed SISO model, the effects of parameter changes on system interactions are analyzed by transfer function Bode plots, zero-pole distributions and Nyqusit curve variations.

2.3 Transfer function of AC-DC distribution system

For the model in Figure 4, the dynamic characteristics of the AC side voltage and current of the converter can be obtained according to KVL as shown in Eq. 1.

$$\begin{bmatrix} u_{sa1} \\ u_{sb1} \\ u_{sc1} \end{bmatrix} = L_{c1}\frac{d}{dt}\begin{bmatrix} i_{ca1} \\ i_{cb1} \\ i_{cc1} \end{bmatrix} + R_{c1}\begin{bmatrix} i_{ca1} \\ i_{cb1} \\ i_{cc1} \end{bmatrix} + \begin{bmatrix} u_{ca1} \\ u_{cb1} \\ u_{cc1} \end{bmatrix} \tag{1}$$

Where: u_{s1} is VSC1 equivalent three-phase supply voltage; R_{c1} is VSC1 AC side equivalent resistance; L_{c1} is VSC1 AC side equivalent inductance; u_{c1} is VSC1 AC side voltage.

Since all three phases are time-varying AC quantities in the stationary coordinate system, which are difficult to control, the Park transform of Eq. 1 is used to obtain the dynamic characteristics of the voltage and current on the AC side of

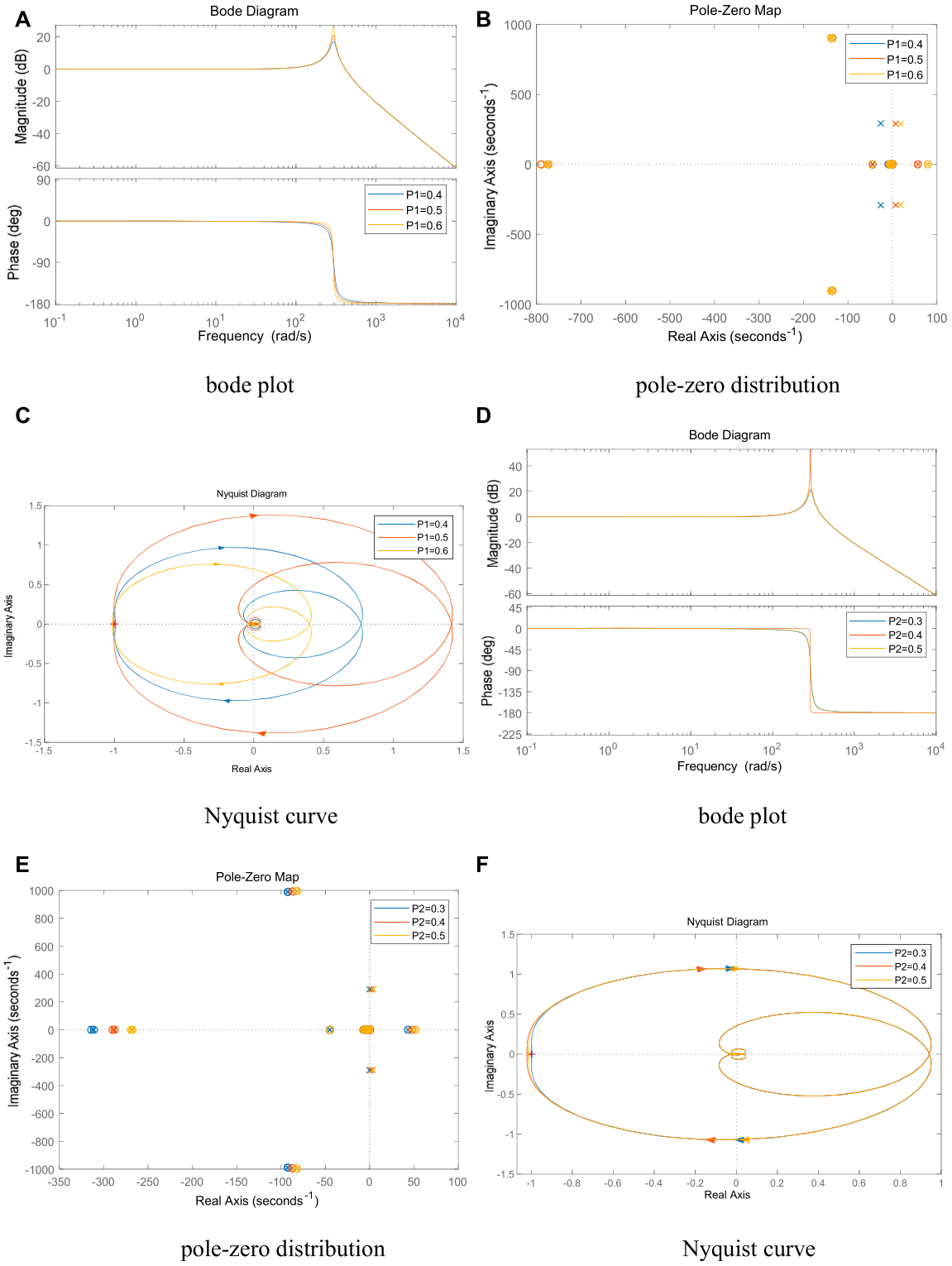


FIGURE 7
Effect of reference power variation of VSC on system stability. **(A)** bode plot, **(B)** pole-zero distribution, **(C)** Nyquist curve , **(D)** bode plot, **(E)** pole-zero distribution and **(F)** Nyquist curve.

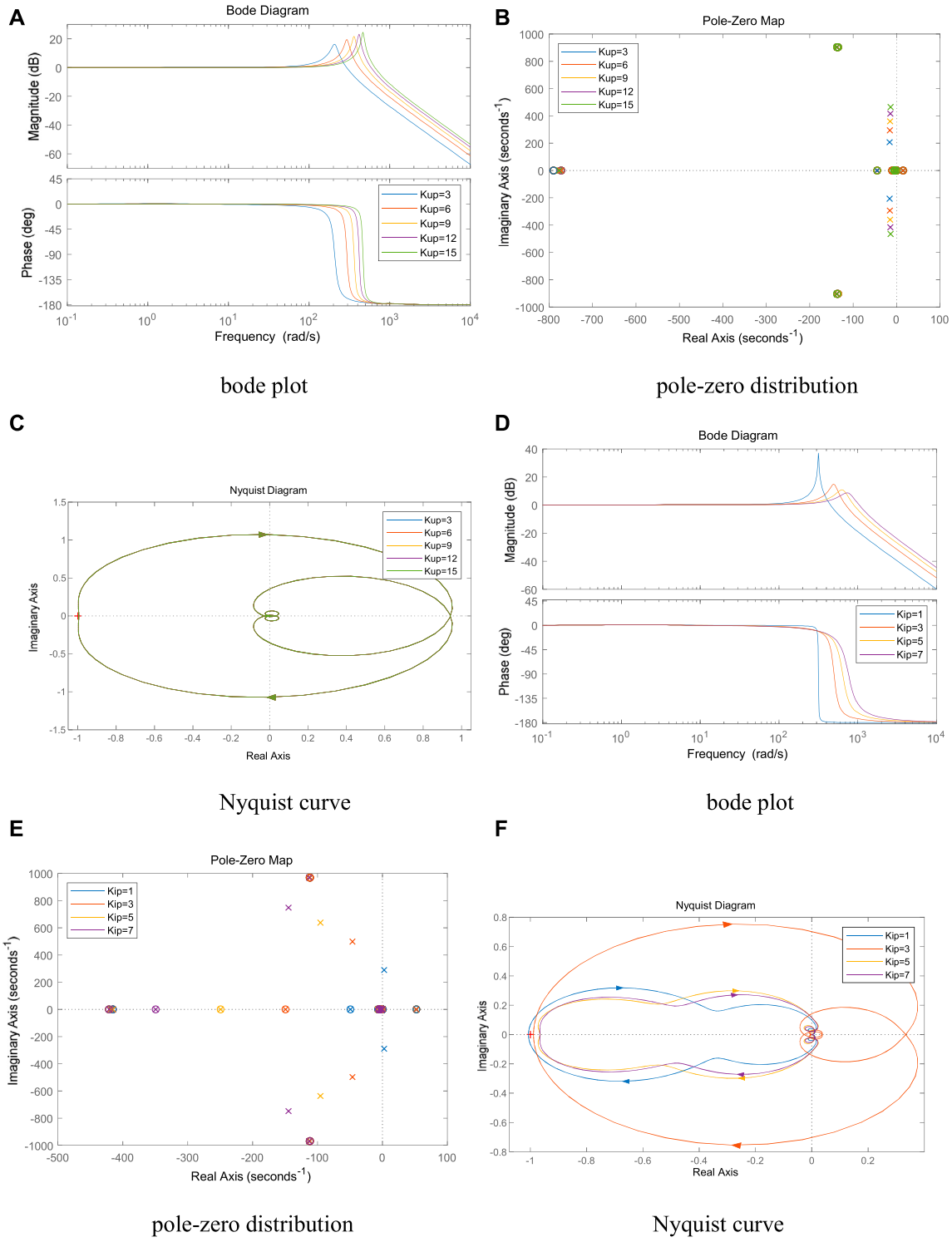


FIGURE 8
Effect of the control parameters of VSC1 on system stability. **(A)** bode plot, **(B)** pole-zero distribution, **(C)** Nyquist curve, **(D)** bode plot, **(E)** pole-zero distribution and **(F)** Nyquist curve.

the converter in the dq coordinate system, as shown in Eq. 2 and Eq. 3.

$$\begin{cases} u_{sd1} - u_{cd1} = L_{c1}\dfrac{di_{cd1}}{dt} + R_{c1}i_{cd1} - \omega L_{c1}i_{cq1} \\ u_{sq1} - u_{cq1} = L_{c1}\dfrac{di_{cq1}}{dt} + R_{c1}i_{cq1} + \omega L_{c1}i_{cd1} \end{cases} \tag{2}$$

The form of Eq. 2 and Eq. 3 in the s-domain is:

$$\begin{cases} u_{sd1} - u_{cd1} = i_{cd1}\left(R_{c1} + sL_{c1}\right) - \omega L_{c1}i_{cq1} \\ u_{sq1} - u_{cq1} = i_{cq1}\left(R_{c1} + sL_{c1}\right) + \omega L_{c1}i_{cd1} \end{cases} \tag{3}$$

Where u_{sd1} and u_{sq1} are the dq-axis components of the VSC1 AC supply; u_{c1d} and u_{c1q} are the dq-axis components of the AC-side voltage of VSC1; i_{cd1} and i_{cq1} are the dq-axis components of the VSC1 AC-side current; ω is the AC supply fundamental angular frequency; and s is the Laplace operator.

In order to eliminate the coupling between variables, a feedforward decoupling control strategy is used for the current inner loop (Wu et al., 2020), which yields the voltage outer loop control equation:

$$i_{cd1,ref} = \left(K_{up1} + \frac{K_{ui1}}{s}\right)\left(U_{dc,ref} - U_{dc}\right) \tag{4}$$

Current inner-loop control equation:

$$\begin{cases} u_{sd1} = \left(K_{ip1} + \dfrac{K_{ii1}}{s}\right)\left(i_{cd1,ref} - i_{cd1}\right) - \omega L_{c1}i_{cq1} + u_{cd1} \\ u_{sq1} = \left(K_{ip1} + \dfrac{K_{ii1}}{s}\right)\left(i_{cq1,ref} - i_{cq1}\right) + \omega L_{c1}i_{cd1} + u_{cq1} \end{cases} \tag{5}$$

Where: $i_{cd1,\ ref}$ and $i_{cq1,\ ref}$ are the dq-axis reference currents on the AC side; U_{dc} and $U_{dc,\ ref}$ are the actual DC voltage and DC voltage reference values, respectively; K_{vp1} and K_{vi1} are the proportional and integral coefficients of the DC voltage outer-loop PI regulator, respectively; K_{ip1} and K_{ii1} are the proportional and integral coefficients of the AC current inner-loop PI regulator, respectively.

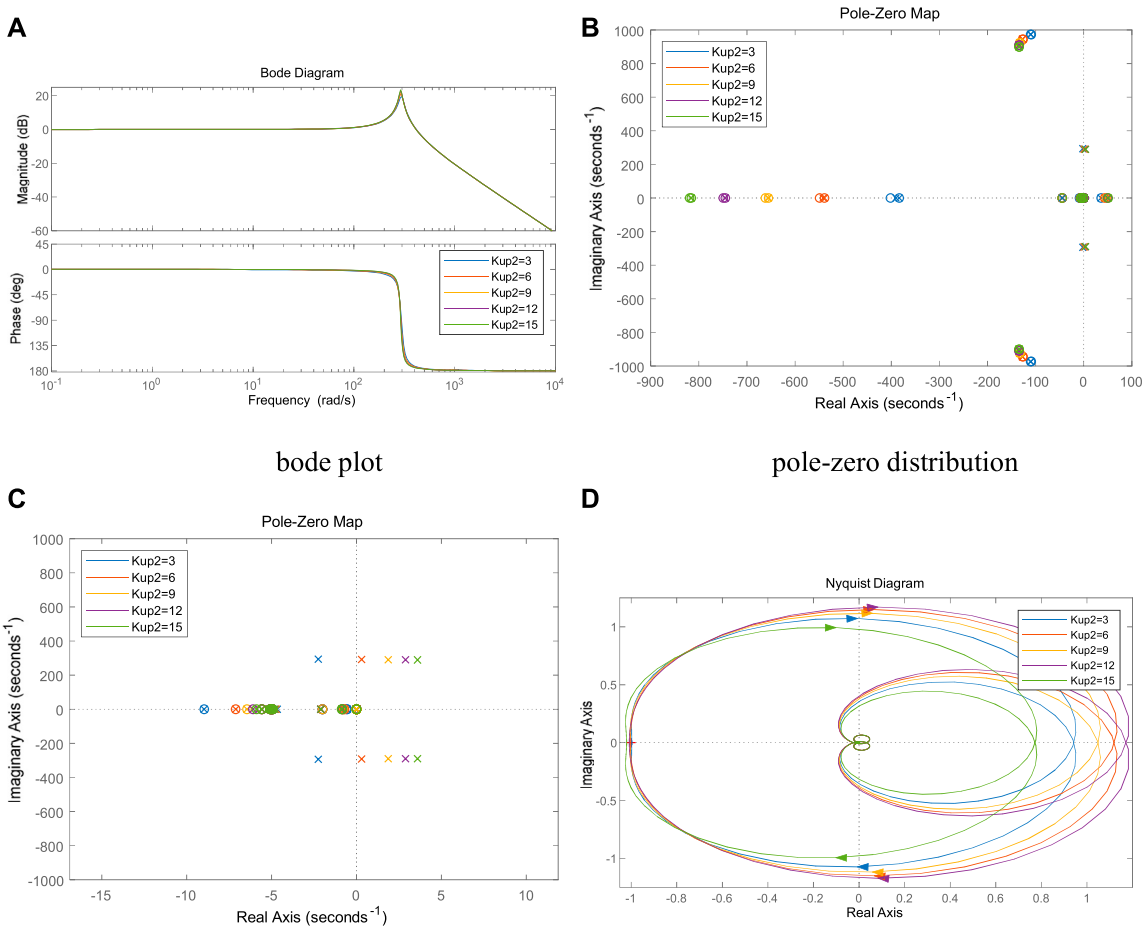


FIGURE 9
(Continued).

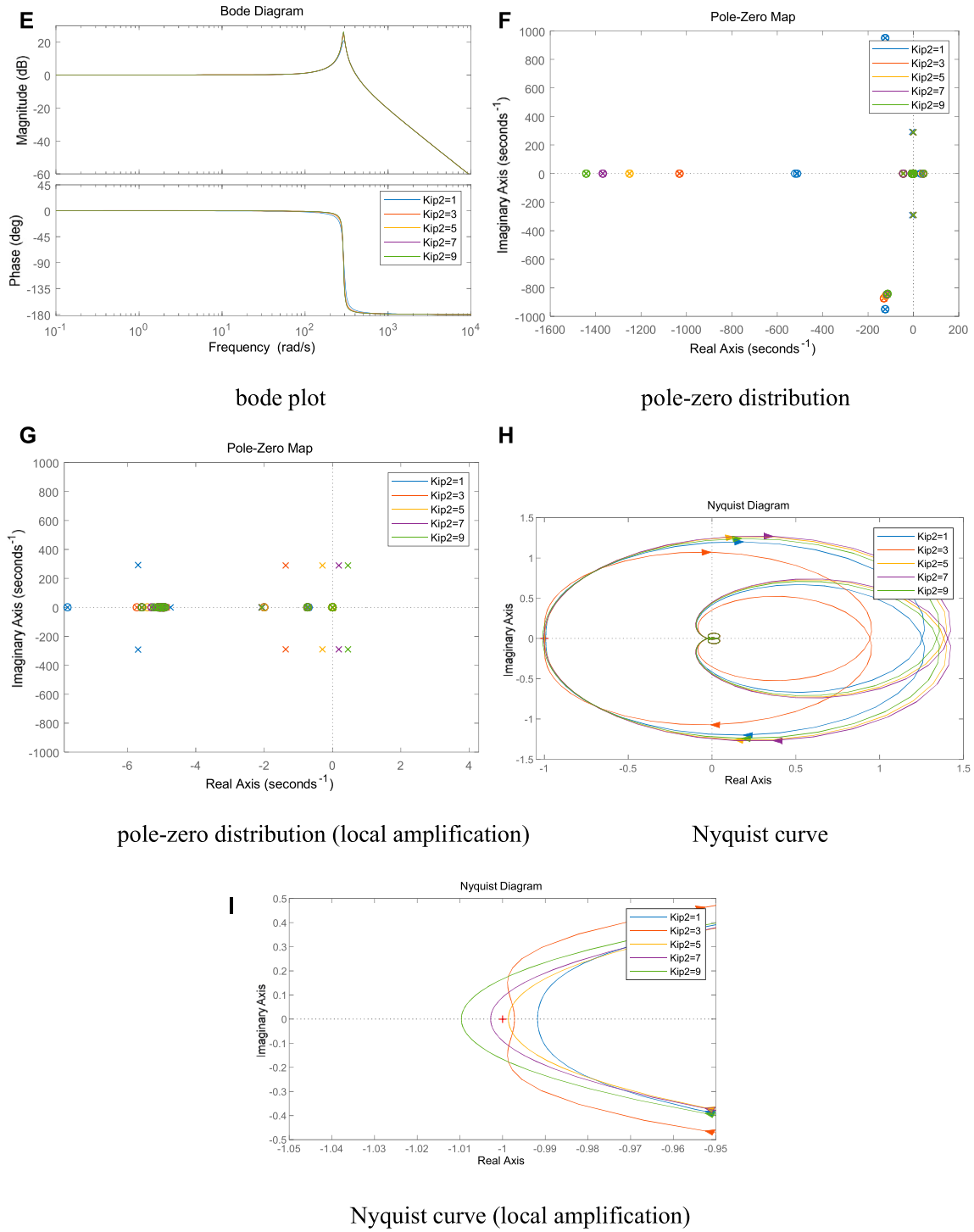


FIGURE 9
(Continued). Effect of the control parameters of VSC2 on system stability. **(A)** bode plot, **(B)** pole-zero distribution, **(C)** pole-zero distribution (local amplification), **(D)** Nyquist curve, **(E)** bode plot, **(F)** pole-zero distribution, **(G)** pole-zero distribution (local amplification), **(H)** Nyquist curve, **(I)** Nyquist curve (local amplification).

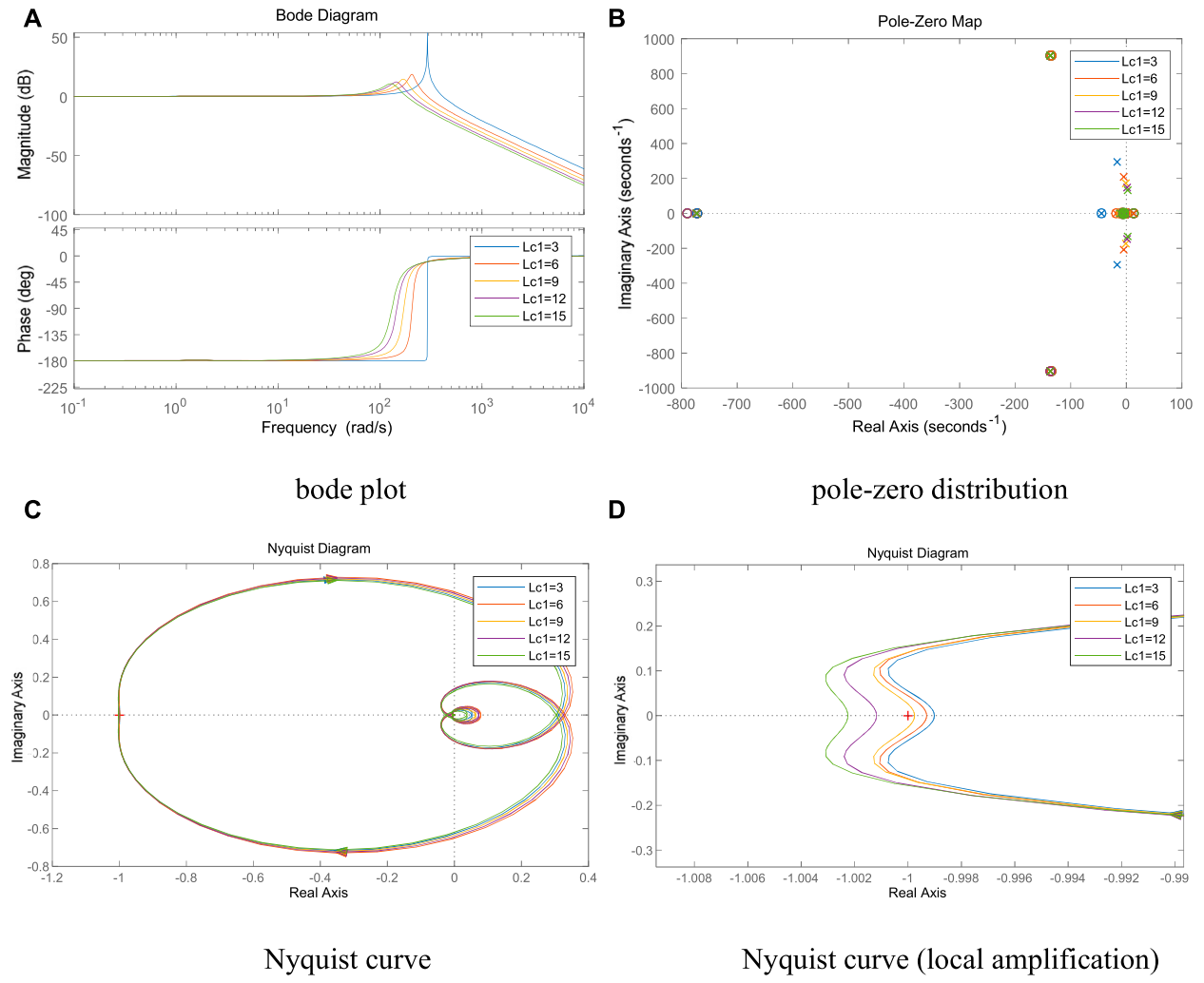


FIGURE 10
Effect of AC side inductance of VSC1 on system stability. **(A)** bode plot **(B)** pole-zero distribution, **(C)** Nyquist curve, **(D)** Nyquist curve (local amplification).

The fully decoupled inner-loop control equations are obtained by combining Eqs 2–5.

$$\begin{cases} \left(K_{ip1}+\dfrac{K_{ii1}}{s}\right)\left(i_{cd1,ref}-i_{cd1}\right)=i_{cd1}\left(R_{c1}+sL_{c1}\right) \\ \left(K_{ip1}+\dfrac{K_{ii1}}{s}\right)\left(i_{cq1,ref}-i_{cq1}\right)=i_{cq1}\left(R_{c1}+sL_{c1}\right) \end{cases} \tag{6}$$

Combining Eqs. 6, 9 and 10 and according to the circuit structure of Figure 1B, the dual-loop control block diagram can be obtained as shown in Figure 2, where: R_{load} is equivalent DC load; I_{l1} is VSC1 equivalent DC load current; C_{dc1} is VSC1 DC side filter capacitor; I_{dc1} is VSC1 output DC current; K_{ceg1} is the equivalent gain of the converter; K_1 is the proportionality factor (Yao et al., 2020).

The main station control block diagram can be seen as composed of two parts, which are the VSC1 control strategy part and the DC system feedback part, so the joint Eqs 3–6, then the d-axis open-loop transfer function of VSC1 can be obtained as:

$$G_{VSC1}=\frac{K_1K_{ceg}\left(K_{ip}+\frac{K_{ii}}{s}\right)\left(K_{up1}+\frac{K_{ui1}}{s}\right)G_{dc}}{\left(R_{c1}+sR_{c1}\right)+K_1K_{ceg}\left(K_{ip1}+\frac{K_{ii1}}{s}\right)} \tag{7}$$

Where G_{dc} is the equivalent transfer function for the DC system. When the slave converter station with constant power control is considered as a constant power load, G_{dc} can be expressed by Eq. 8.

$$G_{dc}=sC_1-\frac{P_{dc}}{u_{dc(0)}^2} \tag{8}$$

3 System model considering interaction

The above modeling approach treats the slave converter station with constant power control as a constant power load and brings it into the system equation as part of the DC load,

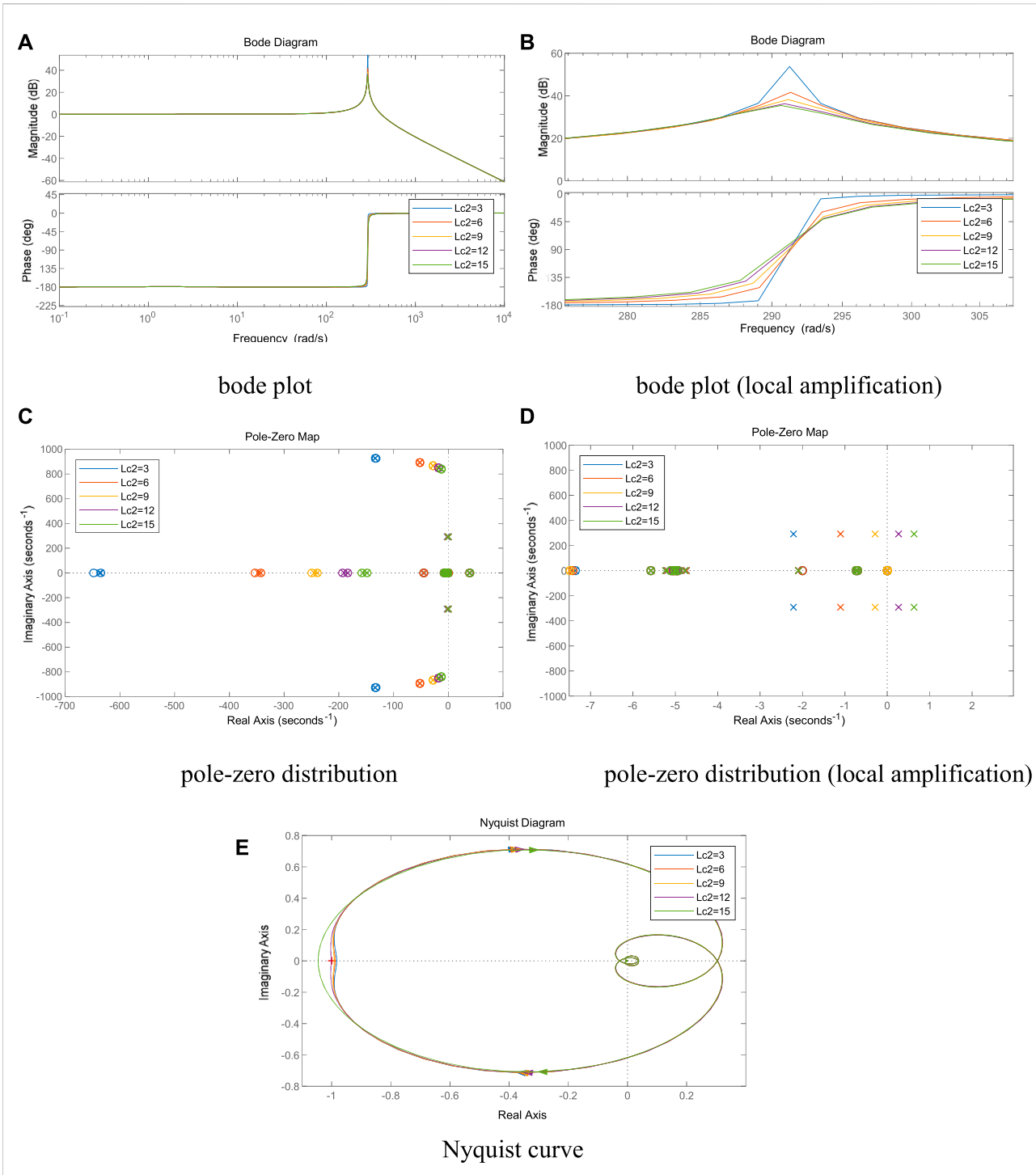


FIGURE 11
Effect of AC side inductance of VSC2 on system stability. **(A)** bode plot, **(B)** bode plot (local amplification), **(C)** pole-zero distribution, **(D)** pole-zero distribution (local amplification), **(E)** Nyquist curve.

without considering the effect of the slave control system on the control margin and system stability of the master station. This can lead to power oscillations between the master and slave stations in some specific cases, which in turn can lead to DC bus voltage oscillations and system instability.

Therefore, in order to consider the influence of the slave control system on the master control capability and system

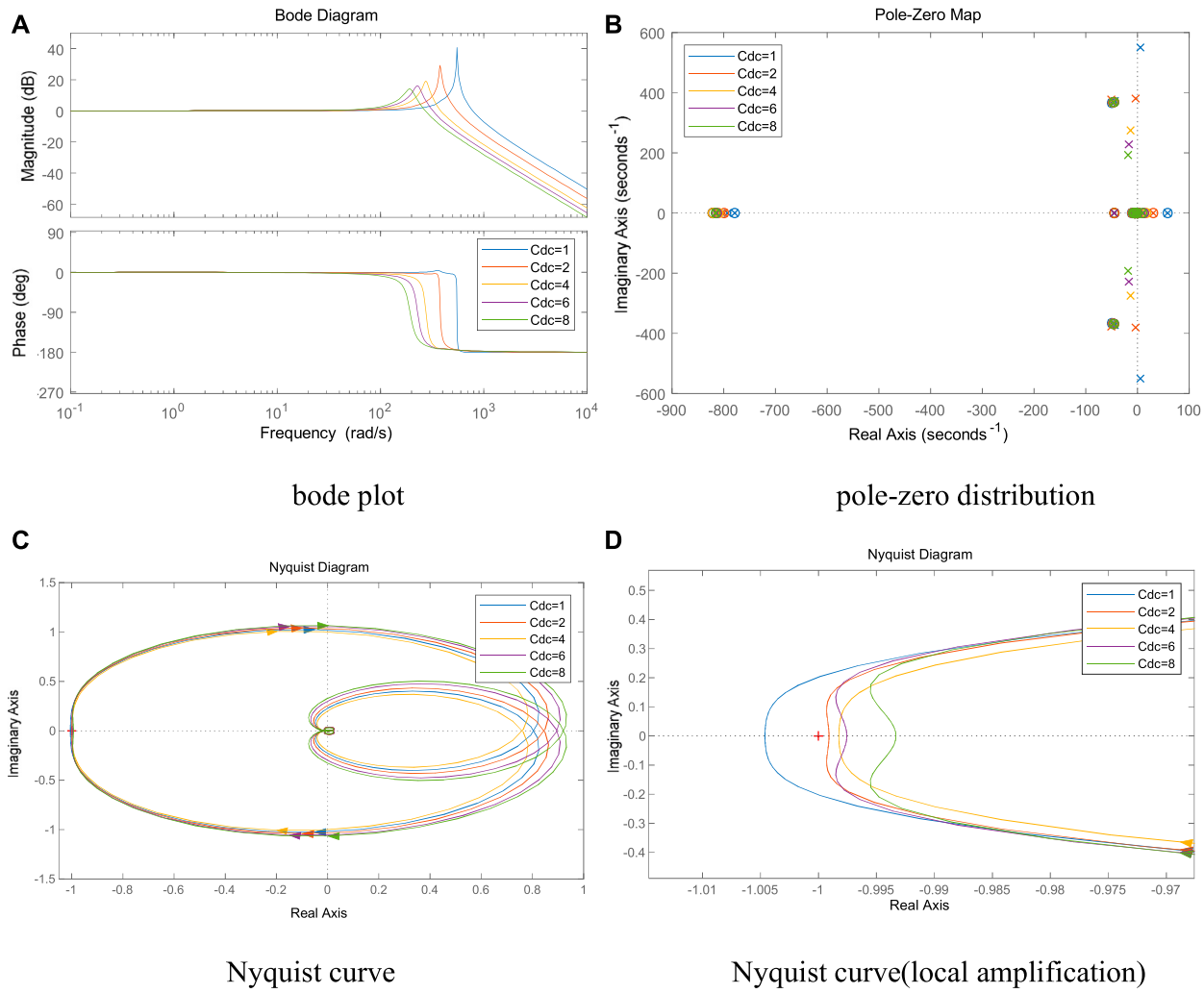


FIGURE 12
Effect of DC-side capacitance on system stability., **(A)** bode plot, **(B)** pole-zero distribution, **(C)** Nyquist curve **(D)** Nyquist curve(local amplification).

stability, this paper improves on the original model, models the DC system more finely, introduces the slave control system into the master transfer function, establishes the AC-DC distribution system model (hereinafter referred to as the interaction model) that counts and converters control system interaction, and further analyzes the influence of the slave control system and line parameters on the stability margin and system stability of the master control system.

3.1 Topology of AC-DC power distribution system

On the basis of the original DC voltage double closed-loop control block diagram of the main converter station, the

TABLE 1 Basic parameters of AC/DC power distribution system.

Symbols	Description	Values
U_s	Reference AC Voltage	380 V
U_{dc}	Reference DC voltage	800 V
L_{c1}	VSC1 AC side inductance	2 mH
L_{c2}	VSC2 AC side inductance	2 mH
C_1	DC side capacitance	4 uF
f	Switch Frequency	10 kHz
K_{ip1}/K_{ii1}	VSC1 inner-loop proportional/integral coefficient	3/15
K_{up1}/K_{ui1}	VSC1 outer-loop proportional/integral coefficient	2/10
K_{ip2}/K_{ii2}	VSC2 inner-loop proportional/integral coefficient	3/15
K_{up2}/K_{ui2}	VSC2 outer-loop proportional/integral coefficient	1/5
K_1	Converters scale factor	0.775
K_{ceg1}	Converters Equivalent Gain	0.8

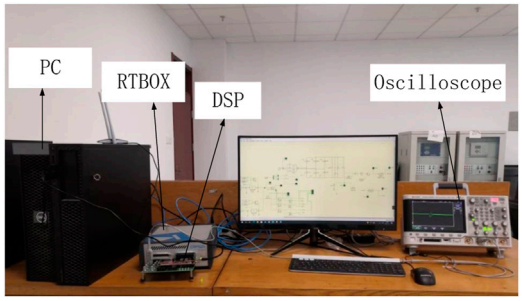


FIGURE 13
Hardware-in-the-loop experiment platform.

influence of the slave control system on the master control system is considered in re-establishing the equivalent transfer function of the DC system, as shown in Figure 5.

For the slave station with constant power control, the same has.

$$\begin{bmatrix} u_{sa2} \\ u_{sb2} \\ u_{sc2} \end{bmatrix} = L_{c2}\frac{d}{dt}\begin{bmatrix} i_{ca2} \\ i_{cb2} \\ i_{cc2} \end{bmatrix} + R_{c2}\begin{bmatrix} i_{ca2} \\ i_{cb2} \\ i_{cc2} \end{bmatrix} + \begin{bmatrix} u_{ca2} \\ u_{cb2} \\ u_{cc2} \end{bmatrix} \tag{9}$$

Where: u_{s2} is the equivalent three-phase supply voltage on the AC side of VSC2; R_{c2} is the equivalent resistance on the AC side of VSC2; L_{C2} is the equivalent inductance on the AC side of VSC2; u_{c2} is the inlet voltage on the AC side of VSC2.

A Park transformation of Eq. 9 has.

$$\begin{cases} u_{sd2} - u_{cd2} = i_{cd2}\left(R_{c2} + sL_{c2}\right) - \omega L_{c2} i_{cq2} \\ u_{gq2} - u_{sq2} = i_{cd2}\left(R_{g2} + sL_{g2}\right) + \omega L_{cg} i_{cd2} \end{cases} \tag{10}$$

According to the control block diagram of the slave converter station with constant power control shown in Figure 4, it is obtained that:

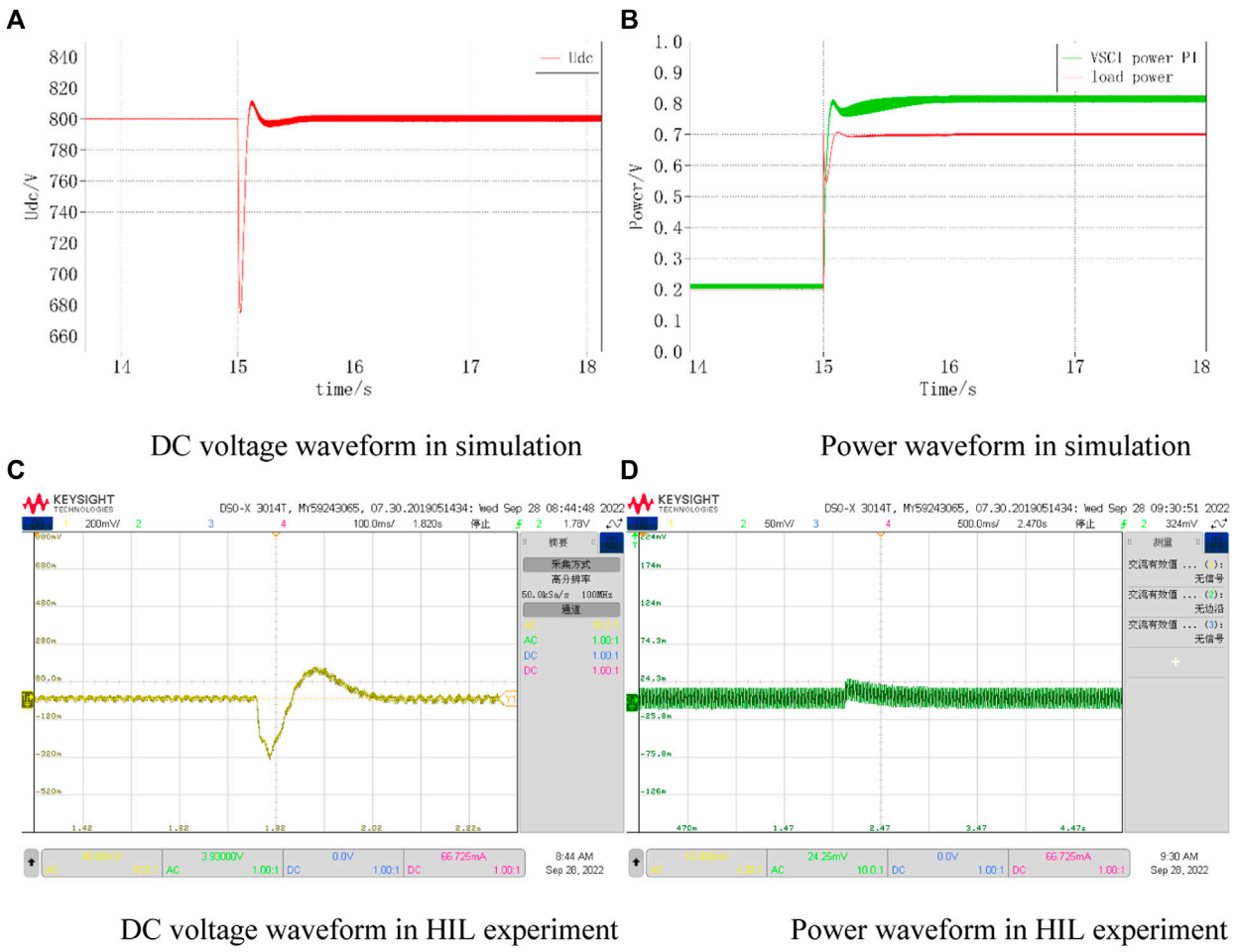


FIGURE 14
Single-terminal model simulation and HIL experimental waveforms. **(A)** DC voltage waveform in simulation **(B)** Power waveform in simulation, **(C)** DC voltage waveform in HIL experiment **(D)** Power waveform in HIL experiment.

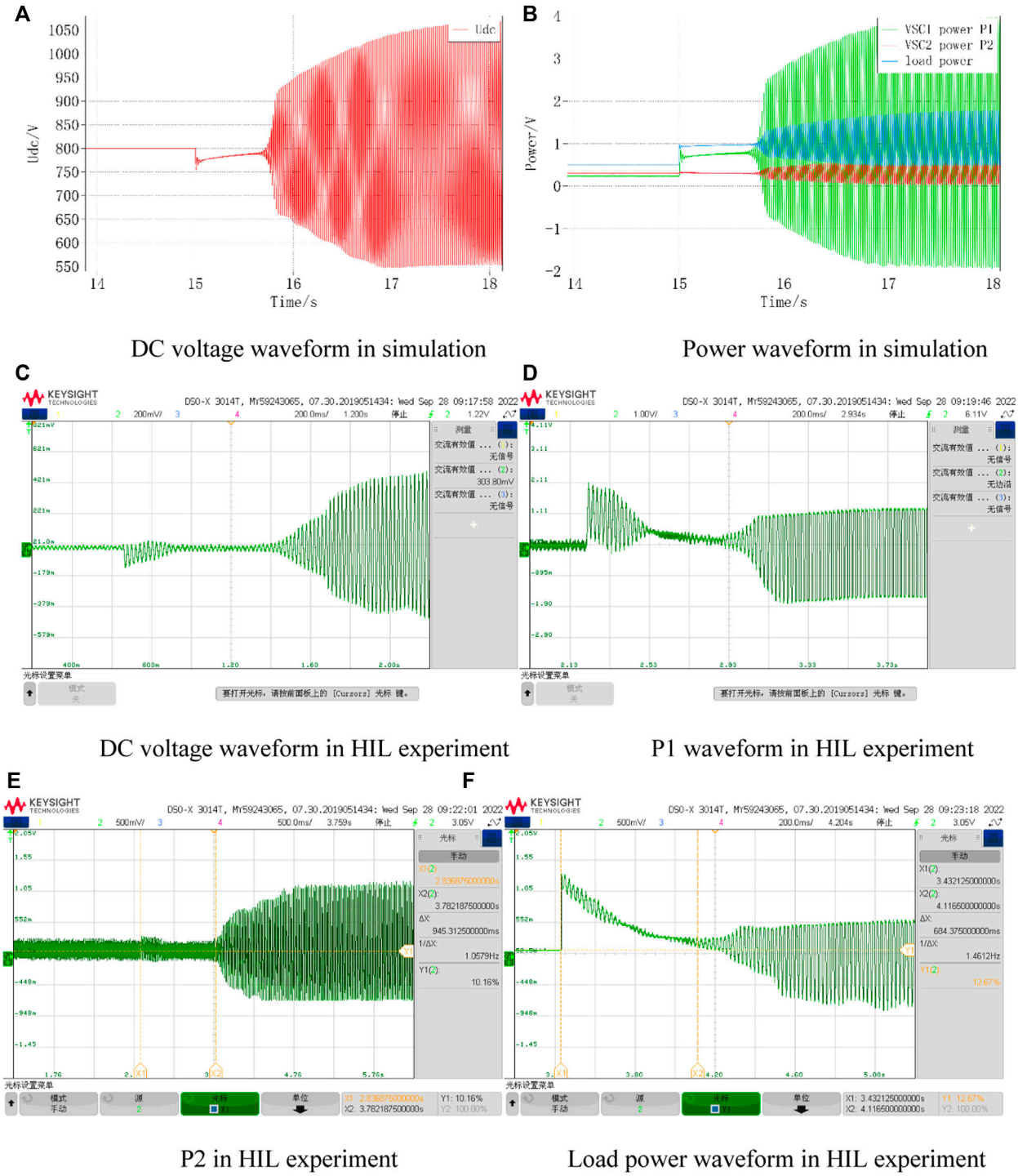


FIGURE 15
Interactive model simulation and HIL experimental waveforms. **(A)** DC voltage waveform in simulation, **(B)** Power waveform in simulation, **(C)** DC voltage waveform in HIL experiment, **(D)** P1 waveform in HIL experiment, **(E)** P2 in HIL experiment, **(F)** Load power waveform in HIL experiment

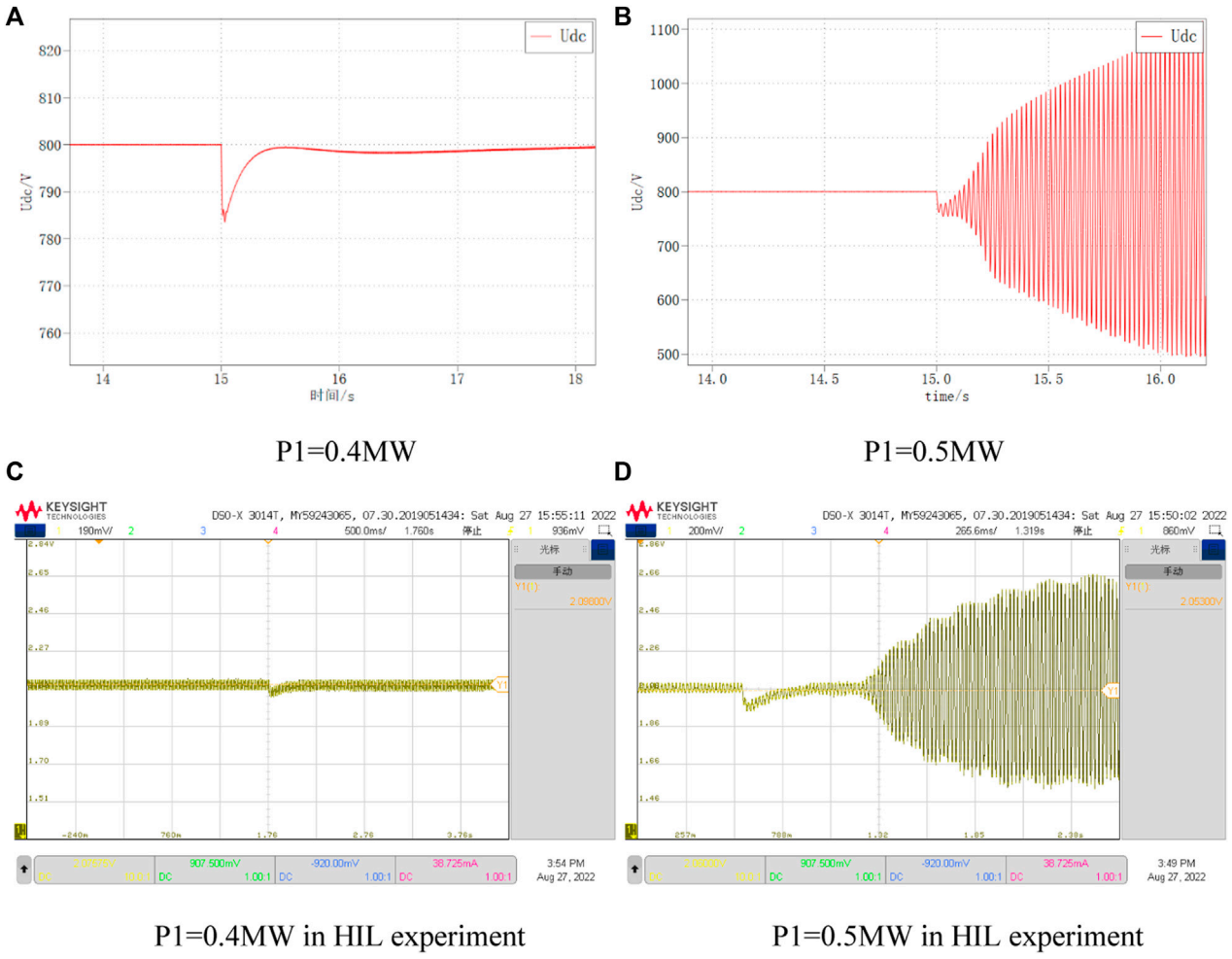


FIGURE 16
DC voltage variation at different P1. **(A)** P1=0.4 MW, **(B)** P1=0.5 MW, **(C)** P1=0.4 MW in HIL experiment, **(D)** P1=0.5 MW in HIL experiment.

$$i_{cd2,ref} = \left(K_{up2} + \frac{K_{ui1}}{s}\right)\left(P_{2,ref} - P_2\right) \tag{11}$$

$$u_{sd2} = \left(K_{ip2} + \frac{K_{ii2}}{s}\right)\left(i_{cd2,ref} - i_{cd2}\right) - \omega L_{c2} i_{cq2} + u_{cd2} \tag{12}$$

$$\left(K_{ip2} + \frac{K_{ii2}}{s}\right)\left(i_{cd2,ref} - i_{cd2}\right) = i_{cd2}\left(R_{c2} + sL_{c2}\right) \tag{13}$$

Where: $i_{cd2,\ ref}$ and $i_{cq2,\ ref}$ are the dq-axis reference currents on the AC side; P_2 and $P_{2,\ ref}$ are the actual active power and active power reference values, respectively; K_{vp2} and K_{vi2} are the proportional and integral coefficients of the active power outer-loop PI regulator, respectively; K_{ip2} and K_{ii2} are the proportional and integral coefficients of the AC current inner-loop PI regulator, respectively.

Coupling Eqs. 10–13, the d-axis transfer function of VSC2 can be obtained as

$$G_{VSC2} = \frac{3\left[u_{sd2} - u_{dc} i_{dc}\left(K_{up2} + \frac{K_{ui1}}{s}\right)\left(K_{ip2} + \frac{K_{ii2}}{s}\right)\right]\left[\frac{u_{sd2} + i_{cd2}\left(R_{c2} + sL_{c2}\right)}{R_{g2} + sL_{g2}} + i_{cd2}\right] - 2u_{dc} i_{dc}\left[\frac{u_{sd2} + i_{cd2}\left(R_{c2} + sL_{c2}\right)}{R_{g2} + sL_{g2}} + 1\right]}{u_{dc}^2\left[2\frac{\left(R_{c2} + sL_{c2}\right) + \left(K_{ip2} + \frac{K_{ii2}}{s}\right)}{R_{g2} + sL_{g2}} + 3\frac{u_{sd2} + i_{cd2}\left(R_{c2} + sL_{c2}\right)}{R_{g2} + sL_{g2}}\left(K_{up2} + \frac{K_{ui1}}{s}\right)\left(K_{ip2} + \frac{K_{ii2}}{s}\right) + 3\frac{i_{cd2}}{R_{g2} + sL_{g2}}\left(K_{up2} + \frac{K_{ui1}}{s}\right)\left(K_{ip2} + \frac{K_{ii2}}{s}\right) + 2\right]} \tag{14}$$

Therefore, when the system interaction is considered, the DC feedback part of the transfer function changes from Figures 6A,B, i.e., the transfer function of the DC system is changed from the original two parts of DC capacitance and equivalent load to three parts consisting of DC capacitance, equivalent load and VSC2 equivalent transfer function. Therefore, the DC system transfer function of the control system from the station is considered as

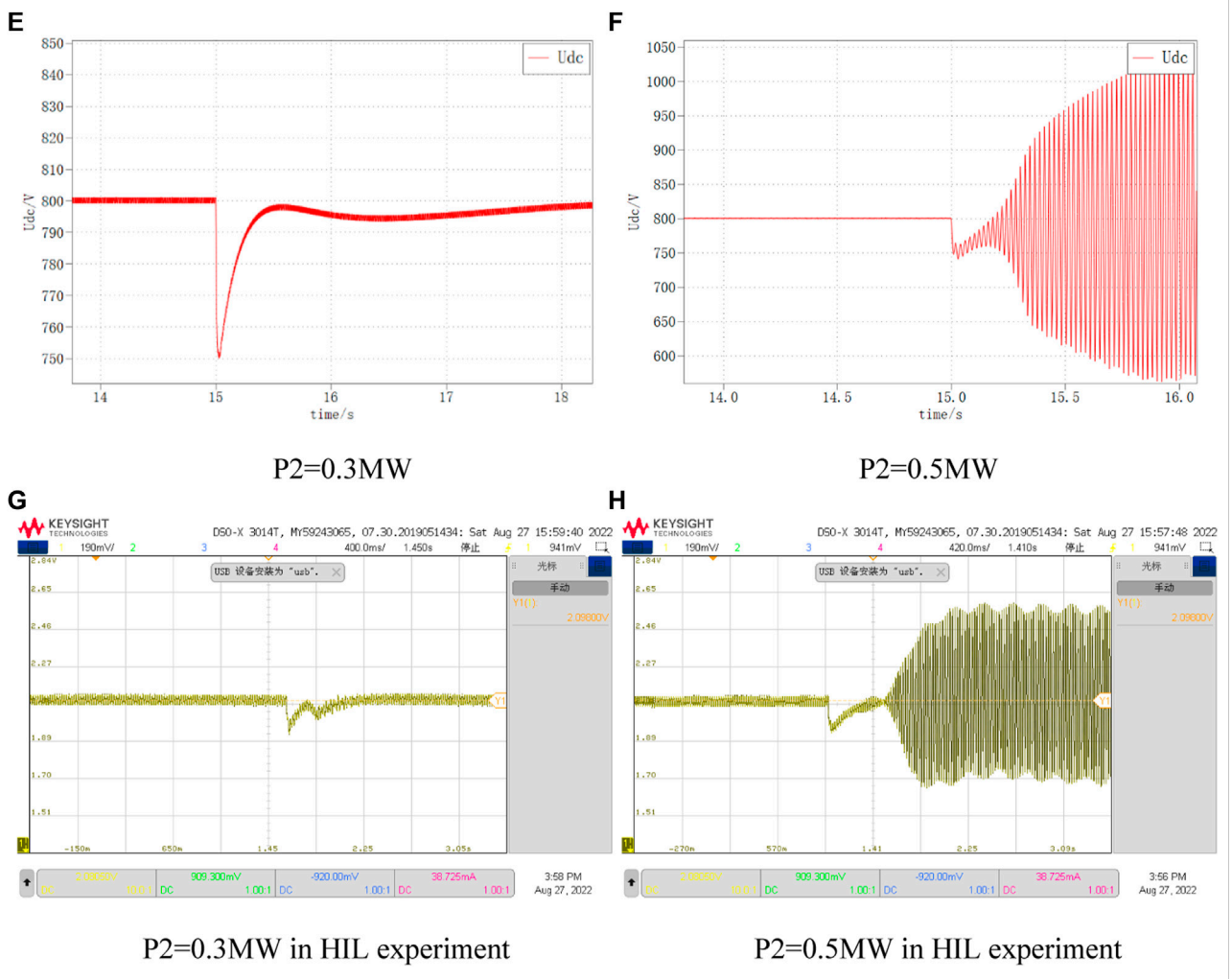


FIGURE 17
DC voltage variation at different P2. **(E)** P2=0.3 MW, **(F)** P2=0.5 MW, **(G)** P2=0.3 MW in HIL experiment, **(H)** P2=0.5 MW in HIL experiment

$$G_{dc}^{'} = sC_1 - \frac{P_{dc}^{'}}{u_{dc(0)}^2} + G_{VSC2} \tag{15}$$

The overall transfer function of VSC1 considering system interactions is

$$G_{VSC1}^{'} = \frac{K_1 K_{ceg1}\left(K_{ip1} + \frac{K_{ii1}}{s}\right)\left(K_{up1} + \frac{K_{ui1}}{s}\right)G_{dc}^{'}}{(R_{c1} + sR_{c1}) + K_1 K_{ceg1}\left(K_{ip1} + \frac{K_{ii1}}{s}\right)} \tag{16}$$

Comparing Eqs. 7, 16, the main difference is the change in the equivalent transfer function of the DC part. Due to the introduction of the VSC2 control part, the order of the system increases and the complexity is higher. Subsequently, the influence of the interaction on the system is analyzed by comparing the impedance characteristics of Eqs. 7, 16, and the influence of each control and electrical parameter on the interaction and system stability based on Eq. 16.

3.2 Analysis of interaction

The classical control theory is used for stability analysis on Eqs. 7, 16. By comparing the impedance characteristics in Figure 6A, there is a resonant spike in the VSC DC port impedance in the middle frequency band range when the influence of the slave control system is considered. The amplitude of the VSC DC port impedance increases in the resonant band, and the phase at the resonance point has a 180° jump, and when the phase angle difference >180°, the system will not satisfy the impedance ratio stability criterion and become unstable. And the same is reflected in Figure 6B and Figure 6C, where the zero pole distribution and Nyquist curve of the system change considerably under the influence of the interaction, and the interaction model has a right half-plane pole whose Nyquist curve encloses the (-1,0) point, which means that the system is in an unstable state.

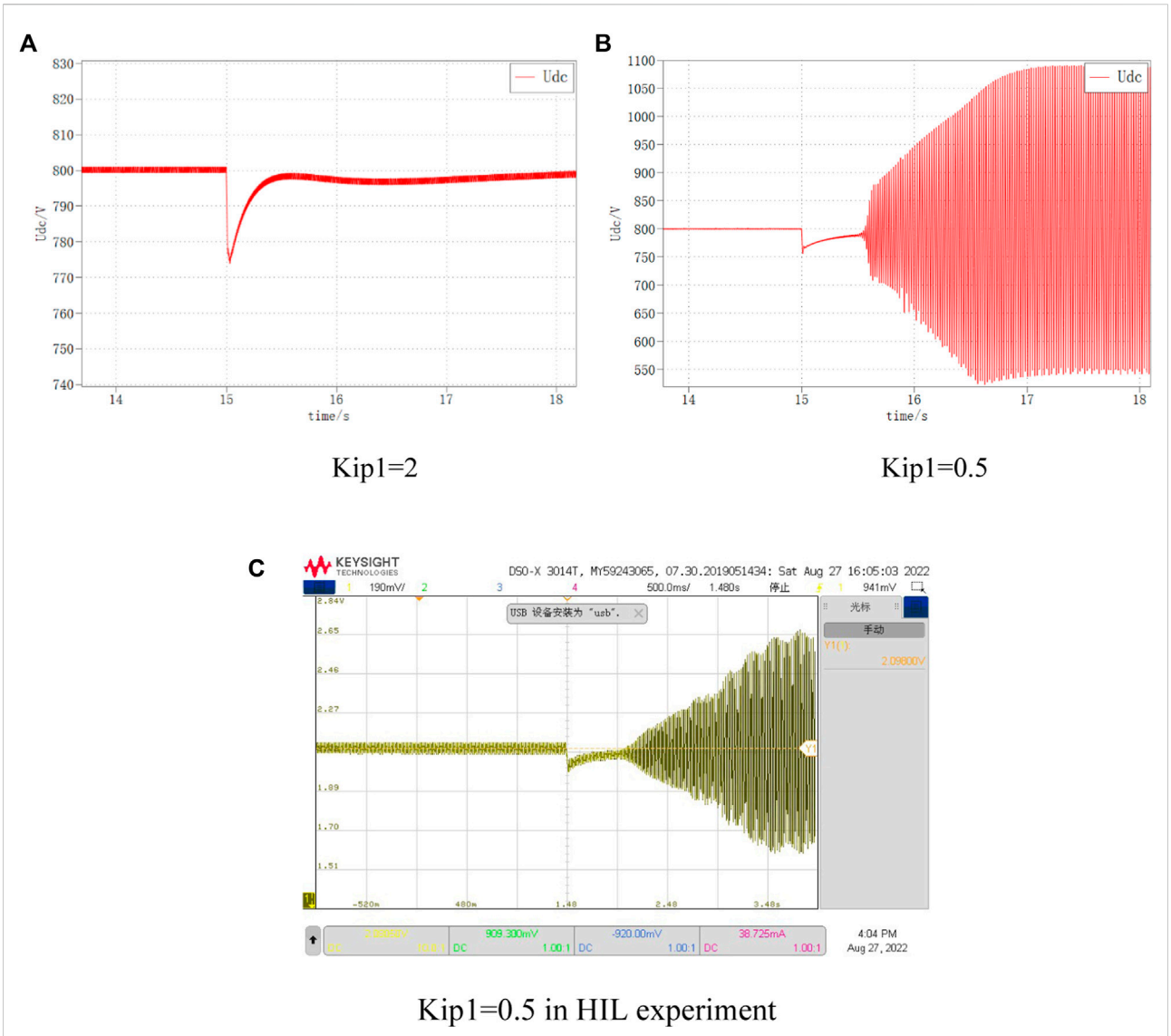


FIGURE 18
DC voltage variation at different Kip1. **(A)** Kip1=2, **(B)** Kip1=0.5, **(C)** Kip1=0.5 in HIL experiment

This is caused by the interaction between the impedance characteristics of VSC2 and the DC network, and between the master converter stations, after it is connected to the DC network. Even when the DC bus voltage is constant, this interaction behavior still weakens the DC network damping and the increase of the transmitted power from the slave station may result in a phase margin of 0° or less than 0°. That is, the resonant peaks of the VSC DC port impedance in the mid-frequency range and the negative damping characteristics of the CPL wide band are the main reasons for the reduced stability of the DC side of the AC-DC distribution system. Therefore, the influence law of control parameters and system parameters on the VSC DC port impedance characteristics needs to be studied to lay the foundation for the subsequent stability analysis of the AC-DC distribution system.

Combining Eq. 15 and Eq. 16, it can be seen that the VSC1 dc impedance depends on three parts: the dc bus capacitance C_{dc}, the VSC1 output conduction and the VSC2 output conduction. According to the analysis of the RLC equivalent circuit principle, the resonant peak in the VSC1 dc port impedance is generated by the impedance interaction between C_{dc} and the converter output conductance (VSC1, VSC2) in the corresponding frequency band. Therefore, the resonant frequency corresponding to this resonant peak will be affected by the change of the dc-side

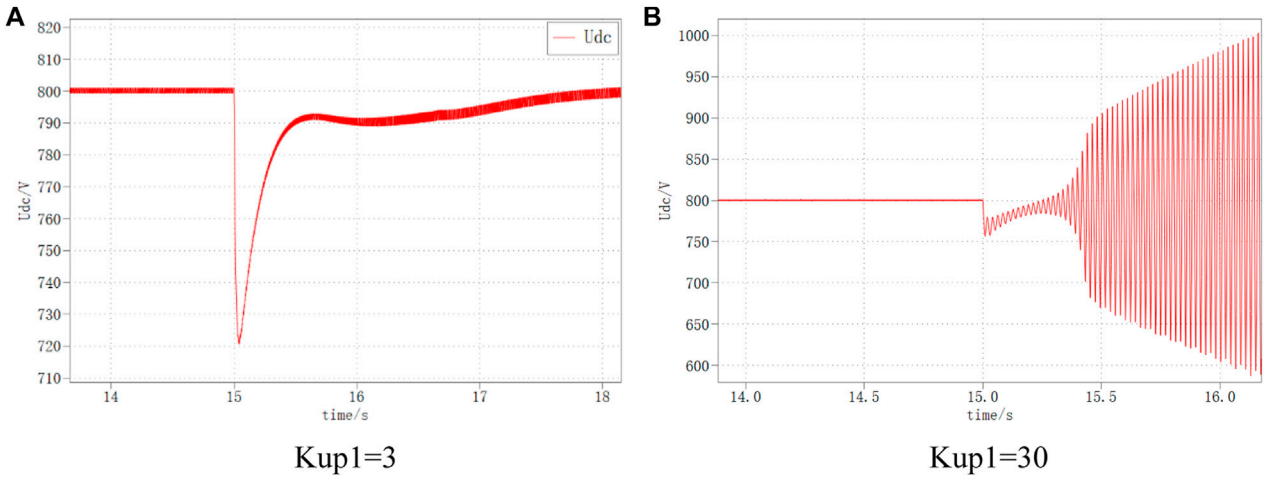


FIGURE 19
DC voltage variation at different Kup1. **(A)** Kup1=3. **(B)** Kup1=30

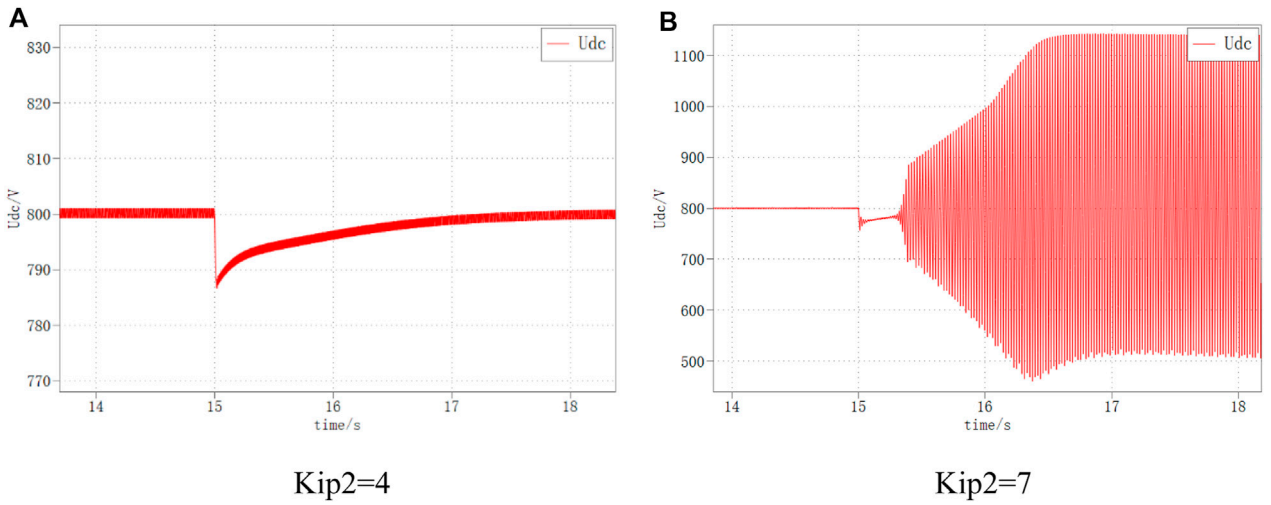


FIGURE 20
DC voltage variation at different Kip2. **(A)** Kip2=4. **(B)** Kip2=7

capacitor Cdc and the slave VSC2 parameters, in addition to the control parameters of VSC1 itself.

4 Stability analysis

Based on the above AC-DC distribution system interaction model, the effects of parameters such as rated power, control parameters, AC-side line impedance and DC capacitance of each converter station on system interaction strength and system stability are analyzed in detail below.

In order to avoid the influence of improperly selected control parameters on the system stability analysis, the control parameters of each converter station are first designed according to the control bandwidth and stability margin. Typically, the bandwidth of the current inner loop is 1/10 to 1/5 of the switching frequency, and the bandwidth of the voltage outer loop is 1/10 to 1/5 of the current inner loop bandwidth. Specific control parameter design values are shown in Table 1. The effect of a single parameter change on the system interaction strength and system stability is discussed during the subsequent stability analysis, and the

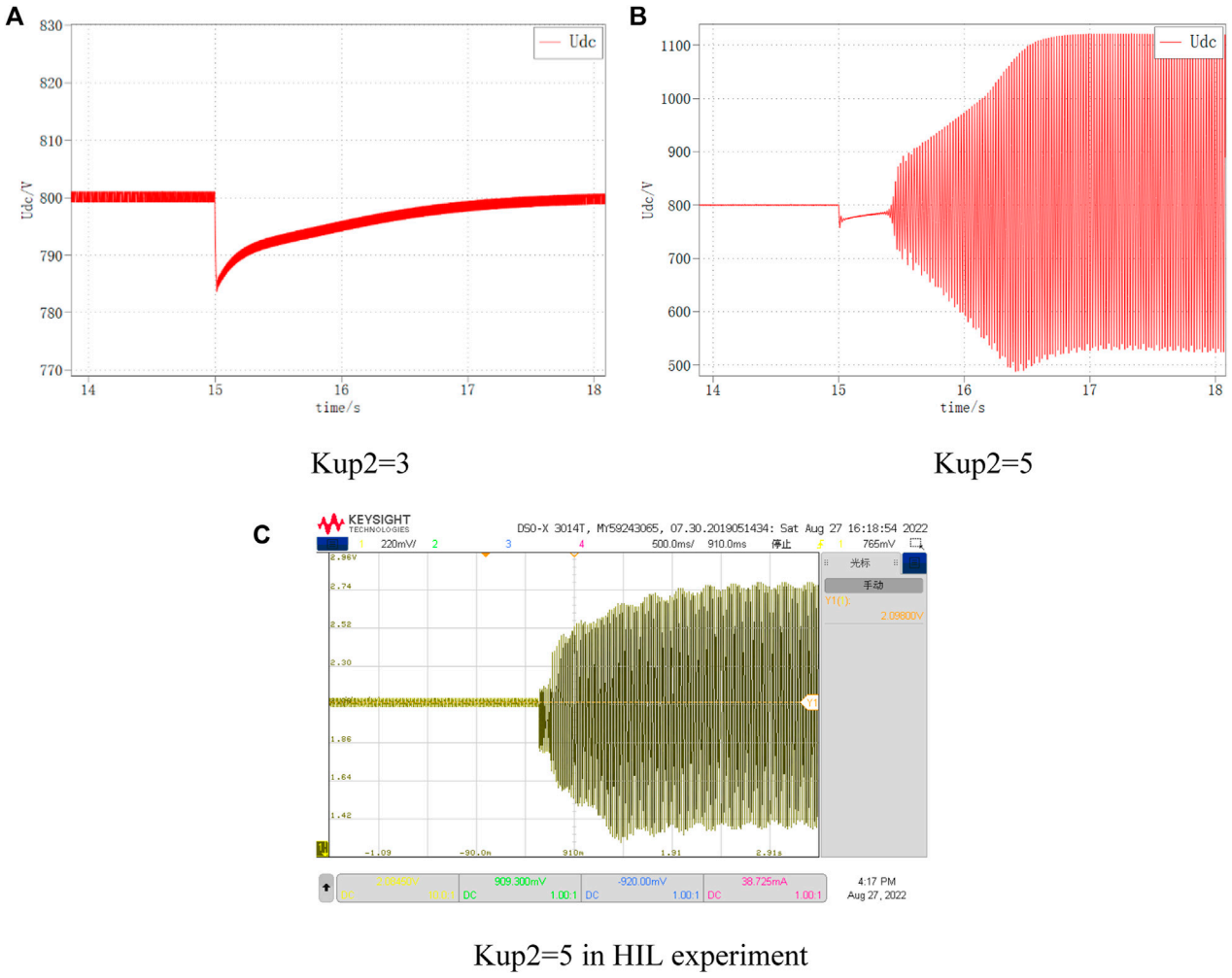


FIGURE 21
DC voltage variation at different Kup2. **(A)** Kup2=3. **(B)** Kup2=5. **(C)** Kup2=5 in HIL experiment.

rest of the parameters are the same as those in the table, so they are not repeated.

4.1 Influence of the rated power of the system

4.1.1 Reference power P1 of converter VSC1

It can be seen from Figures 7A–C, as the reference power P_1 of VSC1 gradually increases, the spike amplitude gradually increases, the system pole gradually shifts to the right, and the intersection point of Nyquist curve and negative real axis gradually moves away from the imaginary axis. When $P_1 \geq 0.5$, the Nyquist curve contains the point (-1, 0), the system pole enters the right half-plane, and the system is destabilized.

4.1.2 Reference power P2 of converter VSC2

It can be seen from Figures 7D–F, as the reference power P_2 of VSC2 gradually increases, the spike amplitude gradually increases, the system pole gradually shifts to the right, and the intersection of the Nyquist curve and the negative real axis gradually moves away from the imaginary axis. When $P_2 \geq 0.4$, the Nyquist curve contains the (-1, 0) point, the system pole enters the right half-plane, and the system is destabilized.

4.2 Influence of VSC1 control parameters

4.2.1 The outer loop control parameters of VSC1

It can be seen from Figures 8A–C, as the outer loop control parameter K_{up1} of VSC1 gradually increases, the spike amplitude

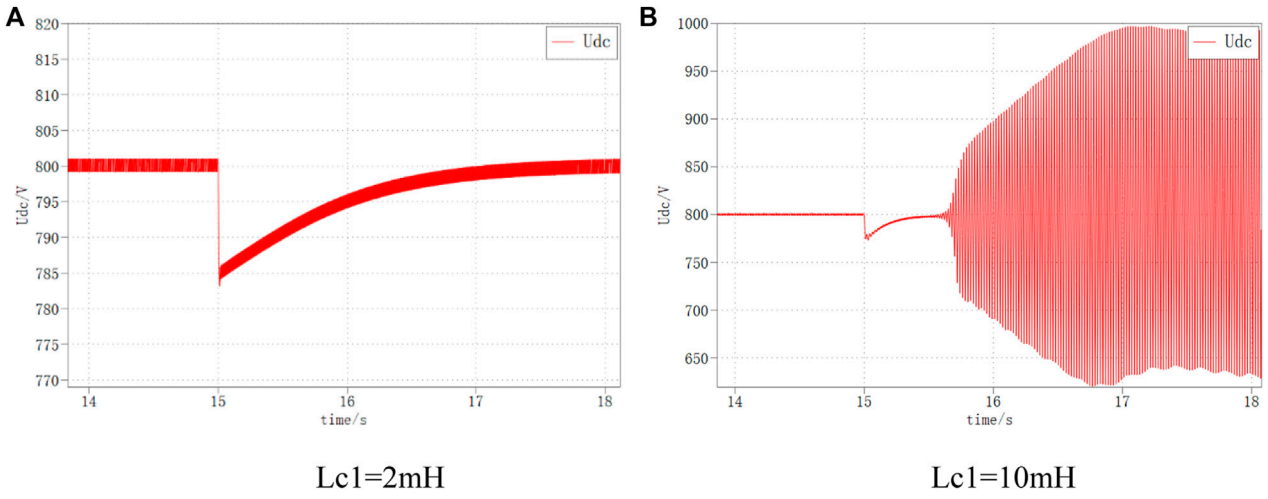


FIGURE 22
DC voltage variation at different Lc1. **(A)** Lc1=2 mH. **(B)** Lc1=10 mH.

gradually increases, the spike frequency keeps increasing, the system pole slowly shifts to the right, and the Nyquist curve does not change much. When K_{up1} was varied within the appropriate parameter interval, the system did not appear to be unstable.

4.2.2 The internal loop control parameters of VSC1

It can be seen from Figures 8D–F, as the inner loop control parameter K_{ip1} of VSC1 gradually increases, the spike amplitude gradually decreases, the spike frequency keeps increasing, the system pole gradually shifts to the right, the intersection point of Nyquist curve and negative real axis gradually approaches the imaginary axis, and the system stability improves. When K_{ip1}≤1, the Nyquist curve contains (-1, 0) points and the system is destabilized.

4.3 Influence of VSC2 control parameter

4.3.1 The outer loop control parameters of VSC2

It can be seen from Figures 9A–D, as the outer loop control parameter K_{up2} of VSC2 gradually increases, the spike amplitude gradually increases, the spike frequency remains basically unchanged, the system pole gradually shifts to the right, and the intersection of the Nyquist curve and the negative real axis gradually moves away from the imaginary axis. When K_{up2}≥3, the Nyquist curve contains the (-1, 0) point, the system pole enters the right half-plane, and the system is destabilized.

4.3.2 Internal loop control parameters of VSC2

It can be seen from Figures 9E–I, as the inner loop control parameter K_{ip2} of VSC2 gradually increases, the spike amplitude gradually increases, the spike frequency basically remains the same, the system pole gradually shifts to the right, and the intersection point of Nyquist curve and negative real axis gradually moves away from the imaginary axis. When K_{ip2}≥5, the Nyquist curve contains the (-1, 0) point, the system pole enters the right half-plane, and the system is destabilized.

4.4 Influence of line parameters

4.4.1 AC side inductance of VSC1

It can be seen from Figures 10A–D, as the AC side line inductance L_{c1} of VSC1 gradually increases, the spike amplitude and spike frequency gradually decrease, the system pole gradually shifts to the right, and the intersection of the Nyquist curve and the negative real axis gradually moves away from the imaginary axis. When L_{c1}≥9 mH, the Nyquist curve contains the (-1, 0) point, the system pole enters the right half-plane, and the system is destabilized.

4.4.2 AC side inductance of VSC2

It can be seen from Figures 11A–E, as the AC side line inductance L_{c2} of VSC2 gradually increases, the spike amplitude gradually decreases, the spike frequency does not change significantly, the system pole gradually shifts to the right, and the intersection of the Nyquist curve and the negative real axis gradually moves away from the imaginary axis. When L_{c2}≥9 mH, the Nyquist curve contains (-1, 0) points, the system pole enters the right half-plane, and the system is destabilized.

4.4.3 Effect of DC-side capacitance

It can be seen from Figures 12A–D, as the dc capacitance C_{dc} gradually increases, the spike amplitude and spike

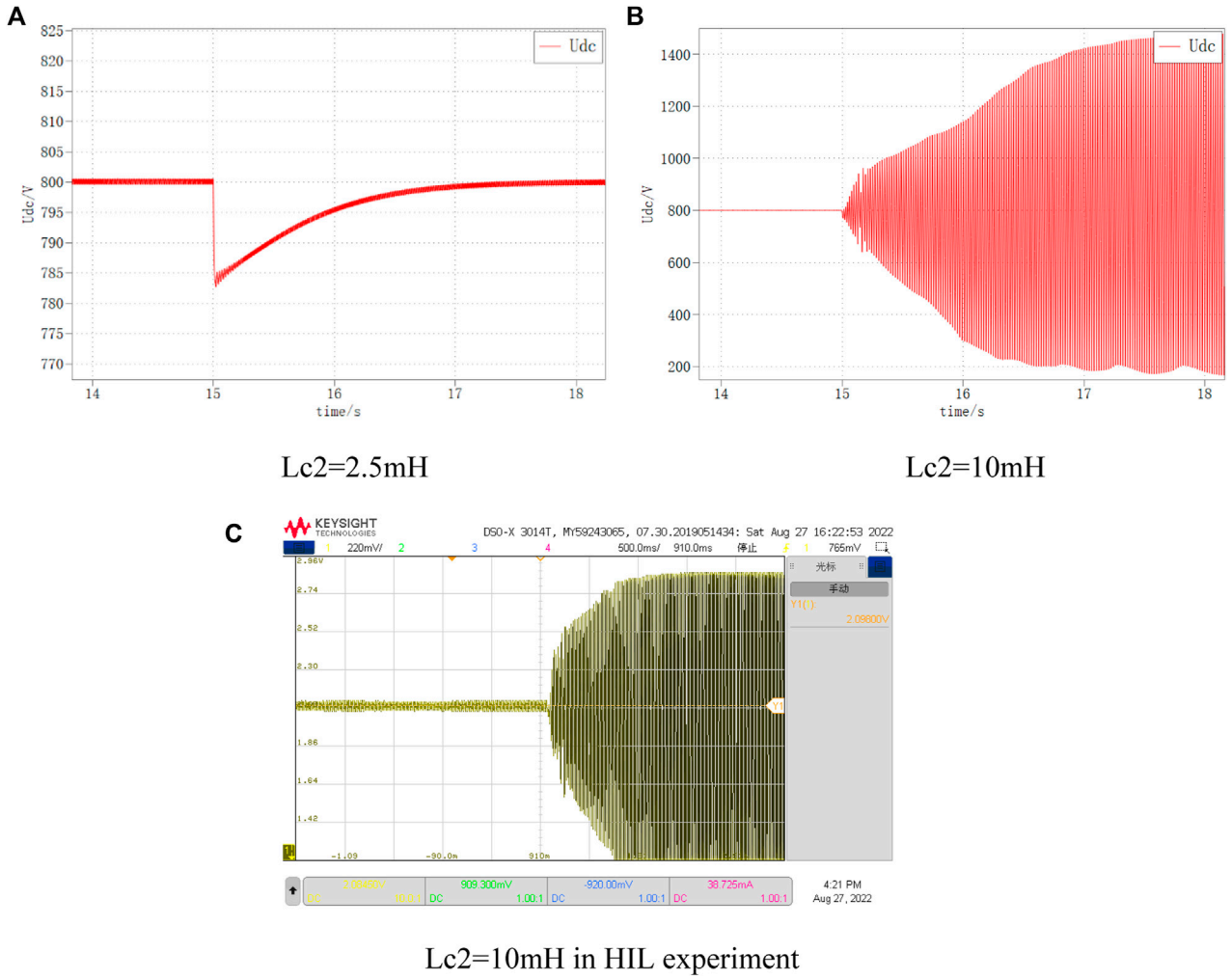


FIGURE 23
DC voltage variation at different Lc2. **(A)** Lc2=2.5 mH, **(B)** Lc2=10 mH, **(C)** Lc2=10 mH in HIL experiment.

frequency gradually decrease, the system pole gradually shifts to the left, and the intersection of the Nyquist curve and the negative real axis gradually approaches the imaginary axis. When C_{dc}≤2 mF, the Nyquist curve contains the (-1, 0) point, the system pole enters the right half-plane, and the system is destabilized.

5 Simulation verification and hardware-in-the-loop experiment

Since the DC voltage is one of the important indicators to measure the power balance and stability of the DC distribution system, the DC side bus voltage is selected as the observed quantity in the subsequent simulation to determine the system stability. 100% increase of DC load at 15s and end of simulation at 30s.

In order to verify the correctness and effectiveness of the proposed control strategy, a controller-level hardware-in-the-loop test platform is built as shown in Figure 13. The platform consists of an RT-LAB real-time simulation experiment bench, a DSP control board, a host computer and an oscilloscope. The controller adopts TI's TM320F28069 main control chip, and the RTBOX has 16 analog input and output ports and 32 digital input and output ports, which fully meet the requirements of this experiment. The output voltage and frequency waveforms are output through the analog output channels of the RTBOX and displayed on the oscilloscope.

First, the original single -terminal model and the interactive model will be compared. The system power of the simulated model is 0.2 for the single -terminal model, 0.5 for the interactive model, and 0.3 for the reference power of VSC2. The load increases by 0.5 at 15 s. The DC voltage

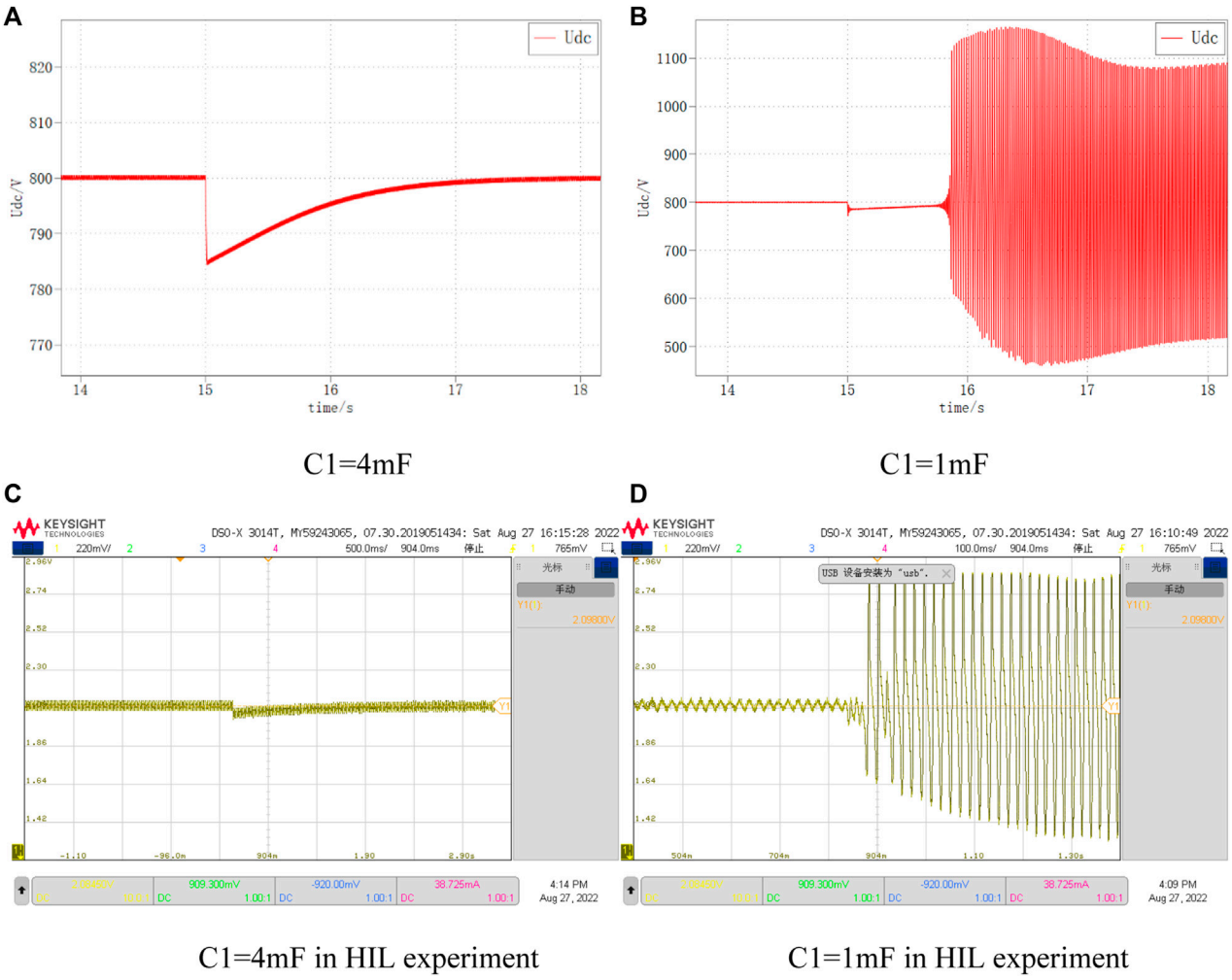


FIGURE 24
DC voltage variation at different C1. **(A)** C1=4 mF, **(B)** C1=1 mF, **(C)** C1=4 mF in HIL experiment, **(D)** C1=1 mF in HIL experiment

variation and the output power of each converter for the single -terminal model and the interactive model are given in Figures 14, 15, respectively, and are tested in hardware-in-the-loop experiments. From the figures, it can be obtained that the interactive model experienced power oscillations in VSC1 during power recovery when the output power of VSC1 increased by the same value, which further led to load power and VSC2 output power oscillations and system collapse. This result verifies the previous theoretical analysis that for the same control parameters and circuit parameters, the system will be in an unstable state after some specific perturbation when considering the influence of the control system from the converter on the system, which is not captured by the original single -terminal equivalent model.

5.1 Reference power of converter station

5.1.1 Changing the reference power P_1 of VSC1

The DC bus voltage waveforms for VSC1 reference power of 0.4 MW and 0.5 MW, respectively, are shown in Figure 16. It can be seen that when P1 is 0.4, the DC bus voltage returns to the stable state smoothly; while when P_1 is 0.5, the DC bus voltage oscillates and diverges under the same disturbance, and the system is unstable.

5.1.2 Changing the reference power P_2 of VSC2

The DC bus voltage waveforms for VSC1 reference power of 0.3 MW and 0.5 MW, respectively, are shown in Figure 17. It can be seen that when P2 is 0.3, the DC bus voltage returns to the stable state smoothly; while when P_2 is 0.5, the DC bus voltage oscillates and diverges under the same disturbance, and the system is unstable.

5.2 Reference power of converter station

5.2.1 Changing the inner loop parameters of VSC1

The DC bus voltage waveforms for VSC1 the inner loop parameters of 2 and 0.5, respectively, are shown in Figure 18. It can be seen that when K_{ip1} is 2, the DC bus voltage returns to the stable state smoothly; while when K_{ip1} is 0.5, the DC bus voltage oscillates and diverges under the same disturbance, and the system is unstable.

5.2.2 Changing the outer loop parameters of VSC1

The DC bus voltage waveforms for VSC1 the inner loop parameters of 3 and 30, respectively, are shown in Figure 19. It can be seen that when K_{up1} is 3, the DC bus voltage returns to the stable state smoothly; while when K_{up1} is 30, the selection of K_{up1} is out of the normal range at this time, the DC bus voltage oscillates and diverges under the same disturbance, and the system is unstable.

5.3 VSC2 control parameters

5.3.1 Changing the inner loop parameters of VSC2

The DC bus voltage waveforms for VSC1 the inner loop parameters of 4 and 7, respectively, are shown in Figure 20. It can be seen that when K_{ip2} is 4, the DC bus voltage returns to the stable state smoothly; while when K_{ip2} is 7, the DC bus voltage oscillates and diverges under the same disturbance, and the system is unstable.

5.3.2 Changing the outer loop parameters of VSC2

The DC bus voltage waveforms for VSC1 the inner loop parameters of 3 and 5, respectively, are shown in Figure 21. It can be seen that when K_{up2} is 3, the DC bus voltage returns to the stable state smoothly; while when K_{up2} is 5, the DC bus voltage oscillates and diverges under the same disturbance, and the system is unstable.

5.4 Line parameters

5.4.1 Changing AC side inductance of VSC1

The DC bus voltage waveforms for VSC1 AC side inductance of 2 mH and 10 mH, respectively, are shown in Figure 22. It can be seen that when L_{c1} is 2 mH, the DC bus voltage returns to the stable state smoothly; while when L_{c1} is 10 mH, the DC bus voltage oscillates and diverges under the same disturbance, and the system is unstable.

5.4.2 Changing AC side inductance of VSC2

The DC bus voltage waveforms for VSC2 the AC side inductance of 2.5 mH and 10 mH, respectively, are shown in Figure 23. It can be seen that when L_{c2} is 2.5 mH, the DC bus voltage returns to the stable state smoothly; while when L_{c2} is 10 mH, the DC bus voltage oscillates and diverges under the same disturbance, and the system is unstable.

5.4.3 Changing DC-side capacitance

The DC bus voltage waveforms for DC-side capacitance of 4 mF and 1mF, respectively, are shown in Figure 24. It can be seen that when C_{dc} is 4mF, the DC bus voltage returns to the stable state smoothly; while when C_{dc} is 1mF, the DC bus voltage oscillates and diverges under the same disturbance, and the system is unstable.

6 Conclusion

This paper establishes an AC-DC distribution system model that takes into account the interaction of each converter control system, compares the system frequency response curves of the interaction model and the original model, takes the resonance spikes generated by the interaction between converters as the reference object, and further analyzes the influence of converter control parameters and system electrical parameters on system stability and interaction. The main conclusions are as follows.

(1) The interaction between each converter and the CPL characteristics make the system impedance characteristics have a resonance spike, the amplitude of VSC DC port impedance increases in the resonance frequency band, and the phase at the resonance point has a 180° jump, which reduces the system damping and stability and easily induces system instability.
(2) The resonant spike amplitude generated by the system interaction and CPL characteristics is positively correlated with the rated power of each converter station, the outer loop control parameters of each converter station, the inner loop control of the main converter station, and negatively correlated with the inner loop control parameters of the slave converter station, the AC side inductance and DC side capacitance of each converter station.
(3) The resonant spike frequency generated by system interaction and CPL characteristics is positively correlated with the control parameters of the main converter station, negatively correlated with the AC-side inductance and DC-side capacitance of the main converter station, and has little relationship with the rated power of the converter station, the control parameters of the slave converter station and the AC-side inductance of the slave converter station.

In this paper, a transfer function model of a double-ended AC-DC distribution system is developed to analyze the effects of parameter changes on system interactions, using the main converter station control system as a reference. Since the impedance modeling approach is adopted for the DC side, the DC side is equated to a feedback link in the control block diagram of the main converter station, which greatly simplifies the calculation process while ensuring that the input single-output Nyqusit criterion is still applicable and avoiding the problems of increasing system complexity with the introduction of converter stations, distributed generation and loads. In the future, the model will be optimized to analyze the impact of different distributed generation, different types of loads, and different DC-side topologies on the system stability.

The above analysis only qualitatively analyzes the parameters and draws relevant conclusions, but fails to quantitatively analyze the parameters; the model used in this paper does not consider the influence of the DC line on the system stability; it fails to optimize the control for the interaction of the system, which needs to be further explored in the subsequent work.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

TC: Responsible for resources, software, formal analysis, program compilation and writing original draft. DF: Responsible for methodology, project administration and funding acquisition. XW: Responsible for writing review and editing. BX: Responsible for visualization and supervision. CL: Responsible for investigation and resources. BX: Responsible for obtaining the experimental data.

Funding

The authors declare that this study received funding from Science and Technology Project of "Research on dynamic interaction characteristics and power optimization control technology of AC/DC distribution systems with high proportion of distributed generation" (SGSDLY00DKJS2200405), Linyi power supply company of Shandong electric power company, State Grid. The funder had the following involvement in the study:Provided data support and HIL experimental design.

Conflict of interest

Authors TC, DF, XW, CL, and BX were employed by the company State Grid Shandong Electric Power Company Linyi Power Supply Company.

The authors declare that this study received funding from Shandong electric power company, State Grid. The funder had the following involvement in the study: Provided data support and HIL experimental design.

Publisher's note

References

Deng, H., and Fang, J. (2022). State-space modeling, stability analysis, and controller design of grid-forming converters with distributed virtual inertia. *Front. Energy Res.* 08 (26). doi:10.3389/fenrg.2022.833387

Feng, F., Fang, J., Manandhar, U., Gooi, H., and Wang, L. (2022). Impedance modeling and stability-oriented parameter optimization of isolated dual active bridge-based two-stage AC-DC-DC converter. *Front. Energy Res.* 30 (06). doi:10.3389/fenrg.2022.874467

Huang, X., Liu, Y., Chen, Y., He, Z. X., Wu, W., Ning, Q., et al. (2020). Impedance-based modeling, stability analysis and virtual damping approach in DC grid. *Power Syst. Prot. Control* 48 (7), 108–117. doi:10.19783/j.cnki.pspc.190533

Jiang, S., Peng, K., Xu, B., Zhang, X., and Liu, Y. (2021). Current situation and prospect of demonstration projects of DC distribution system. *Electr. Power Autom. Equip.* 41 (05), 219–231. doi:10.16081/j.epae.202105039

Li, C., Liu, X., Wang, R., Zhang, Y., and Zhang, L. (2022). An improved dual-loop feedforward control method for the enhancing stability of grid-connected PV and energy storage system under weak grids. *Front. Energy Res.* 22 (06). doi:10.3389/fenrg.2022.939376

Li, X., Guo, L., Huang, D., Zhao, Y., and Wang, C. (2019). Research review on operation and control of DC distribution networks. *High. Volt. Eng.* 45 (10), 3039–3049. doi:10.13336/j.1003-6520.hve.20190924002

Liu, K., Li, Y., Liao, X., Zhu, S., and Huai, Q. (2020). Virtual resistor control strategy for improving stability of VSC-hvdc operating in bidirectional power flow mode. *High. Volt. Eng.* 46 (5), 1665–1675. doi:10.13336/j.1003-6520.hve.20200515021

Nian, H., Yang, J., Chen, L., Xu, Y., Hu, B., and Li, M. (2020). DC impedance modeling and stability analysis of AC/DC hybrid power supply system. *High. Volt. Eng.* 46 (10), 3477–3490. doi:10.13336/j.1003-6520.hve.20200634

Peng, K., Wang, L., Li, X., Liu, F., Zhang, H., and Jia, S. (2021). Analysis and control of high-frequency oscillation for DC distribution system with droop control. *Automation Electr. Power Syst.* 45 (17), 43–50. doi:10.7500/AEPS20210329005

Qiu, P., Song, J., Chen, Q., Zhou, W., Xiong, B., and Tang, A. (2022). Research on oscillation characteristics of wind farm sending system based on participation factor. *Front. Energy Res.* 09 (08). doi:10.3389/fenrg.2022.997782

Tao, X., Xiang, J., and Wei, W. (2017). "Impedance modeling and analysis of AC-DC modular cascade system," in 2017 29th Chinese Control and Decision

Conference, Chongqing, China, 28-30 May 2017 (Chongqing: CCDC), 2316–2321. doi:10.1109/CCDC.2017.7978901

Wang, Y., Zhao, C., and Guo, C. (2020). Transfer function model and low-frequency stability analysis for PMSG-based wind farm interconnected with flexible-HVDC system. *Proc. CSEE* 40 (05), 1485–1498. doi:10.13334/j.0258-8013.pcsee.190657

Wu, G., Wangzhou, X., Zhao, B., Liang, J., Li, Y., et al. (2019). Analytical analysis on the active power control stability of the weak grids-connected VSC. *Proc. CSEE* 39 (21), 6163–2633. doi:10.13334/j.0258-8013.pcsee.190232

Wu, G., Zhou, X., Wang, S., Liang, J., Wang, T., Li, Y., et al. (2018). Nalytical research on the mechanism of the interaction between PLL and inner current loop when VSC-HVDC connected to weak grid. *Proc. CSEE* 38 (9), 2622–2633. doi:10.13334/j.0258-8013.pcsee.172502

Wu, W., Luo, A., Chen, Y., Dong, Y., Zhou, L., and Zhou, X. (2017). "DC impedance modeling, stability analysis and active stabilization of the VSC-HVDC system," in IECON 2017 - 43rd Annual Conference of the, Beijing, China, 29 October 2017 - 01 November 2017 (IEEE Industrial Electronics Society), 4891–4896. doi:10.1109/IECON.2017.8216844

Wu, W., Xie, Z., Chen, Y., Liu, J., Guo, J., Xu, Y., et al. (2020). Analysis and suppression of high-frequency oscillation between converter-based source and loads in an island power system. *Int. J. Electr. Power & Energy Syst.* 117, 105616.10–105616.13. doi:10.1016/j.ijepes.2019.105616

Xie, X., Wu, Z., Hu, Q., Quan, X., Dou, X., and Cao, X. (2022). A unified modeling scheme of modular multilevel converter for hybrid AC/DC power grids. *Front. Energy Res.* 08 (13). doi:10.3389/fenrg.2022.844713

Xue, D., Liu, J., Liu, Z., Tu, Y., and Liu, T. (2020). Modeling and analysis of DC terminal impedance of voltage source converters with different control modes. *IEEE Trans. Power Electron.* 35 (6), 5883–5896. doi:10.1109/TPEL.2019.2953118

Yao, G., Peng, K., Li, H., Wei, Z., and Li, X. (2020). Reduced-order model and mechanism analysis of high-frequency oscillation in flexible DC distribution system. *Automation Electr. Power Syst.* 44 (20), 29–36. doi:10.7500/AEPS20200412002

Yin, R., Sun, Y., Wang, S., Zhao, B., Wu, G., Yang, P., et al. (2021). The modeling and stability analysis of dual-infeed VSC based on the single-input and single-output transfer function model. *Proc. CSEE* 41 (19), 6724–6739. doi:10.13334/j.0258-8013.pcsee.201359

Yuan, B., Yuan, Y., Xu, J., Zhao, C., Huang, Y., Zhao, X., et al. (2016). Small signal stability analysis of dual infeed VSCs connected to weak AC system and design of virtual impedance control strategy. *J. North China Electr. Power Univ.* 43 (06), 38–46. doi:10.3969/j.ISSN.1007-2691.2016.06.07

Zhang, H., Harnefors, L., Wang, X., Hasler, J. P., and Nee, H. P. (2019). SISO transfer functions for stability analysis of grid-connected voltage-source converters. *IEEE Trans. Industry Appl.* 55 (3), 2931–3941. doi:10.23919/IPEC.2018.8507452

Zhang, H., Harnefors, L., Wang, X., Hong, G., and Hasler, J. P. (2018). Stability analysis of grid-connected voltage-source converters using SISO modeling. *IEEE Trans. Power Electron.* 34 (8), 8104–8117. doi:10.1109/TPEL.2018.2878930

Zhang, X., Pei, W., Deng, W., Yu, T., Fan, S., and Huang, R. (2017). Stability analysis of AC/DC hybrid distribution system with constant power loads. *Proc. CSEE* 37 (19), 5572–5582+5834. doi:10.13334/j.0258-8013.pcsee.161579

Zhao, Z., Su, X., Yang, J., Zhou, P., Chen, H., and Lu, G. (2022). Droop-based active voltage regulation control and robust stability analysis for VSCs in the large-scale RES-integrated power system. *Front. Energy Res.* 09 (23). doi:10.3389/fenrg.2022.1005593

frontiers | Frontiers in Energy Research

TYPE Original Research
PUBLISHED 17 March 2023
DOI 10.3389/fenrg.2023.1164198

OPEN ACCESS

EDITED BY
Tao Chen,
Southeast University, China

REVIEWED BY
Haixiang Zang,
Hohai University, China
Wei Yang,
Southwest Petroleum University, China

*CORRESPONDENCE
Yue Xiang,
xiang@scu.edu.cn

SPECIALTY SECTION
This article was submitted to Smart Grids, a section of the journal Frontiers in Energy Research

RECEIVED 12 February 2023
ACCEPTED 28 February 2023
PUBLISHED 17 March 2023

CITATION
Wang S, Xiang Y, Chen L, Sun Z, Hu S, Guo M, Hu H, Tong J, Huang Y, Chen L and Liu J (2023), Electric vehicle navigation and cluster dispatch for reliable low-carbon traffic–Power systems.
Front. Energy Res. 11:1164198.
doi: 10.3389/fenrg.2023.1164198

Electric vehicle navigation and cluster dispatch for reliable low-carbon traffic–Power systems

Shunqi Wang[1], Yue Xiang[1]*, Lin Chen[1], Zhou Sun[2,3], Shangjian Hu[2], Mingyang Guo[3], Haifeng Hu[4], Jun Tong[4], Yifan Huang[4], Liang Chen[4] and Junyong Liu[1]

[1]College of Electrical Engineering, Sichuan University, Chengdu, China, [2]Sichuan Road & Bridge Group Co, Ltd., Chengdu, China, [3]Sichuan Shudao Clean Energy Group Co, Ltd., Chengdu, China, [4]Chengdu TELD New Energy Co, Ltd., Chengdu, China

Large-scale integration of electric vehicles (EVs) into the city system for charging will affect the operation of both traffic and distribution networks. An electric vehicle navigation and cluster dispatch model is proposed for improving the overall charging efficiency of EVs on the transportation network and increasing the voltage level of the distribution network. First, a simplified model of vehicles and traffic road network is established, and a cell transmission model (CTM) is used to simulate the real traffic network. The traffic system takes into account charging EVs, discharging EVs, and other vehicles, and traffic congestion is considered. Then, a coupled model of the traffic–power system is built for the orderly charging of electric vehicles upon arrival at a charging station. The model considers the coupling of the two systems on a time scale, and the charging/discharging power at each charging station node is controlled. The validity of the model is verified in a coupled system of 357 cell traffic network and modified IEEE33 nodes. The results show that the proposed model can ensure good guarantee of the distribution network voltage reliability and reveal the scheduling process of the traffic network. The proposed model also provides a reference for planning of charging stations in the distribution network.

KEYWORDS

electric vehicle, navigation, cluster dispatch, traffic–power system, cell transmission model

1 Introduction

In recent years, air pollution, global warming, and other environmental issues have become prominent and urgently need to be addressed. By implementing low-carbon policies, electric vehicles (EVs) are widely used in transportation. Different from traditional fuel cars, electric vehicles have a large difference with them in terms of range capacity. At the same time, the high current involved in the charging process of electric vehicles has a significant impact on the reliability of the distribution network. Therefore, scholars have conducted a series of studies on the problem of electric vehicles (Xiang et al., 2021) (Li et al., 2020).

Large-scale access of electric vehicles to the network will have a greater impact on the way the distribution network operates. The online evolution mechanism of the EV–grid system was proposed in Dong et al. (2021) for penetration of intermittent renewable energy and the

time variance problem of the system model. Li et al. (2023) proposed a two-stage EV management scheme to promote sustainable transportation. In the first stage, a toll scheduling model based on fuzzy multi-criteria decision-making was proposed, and in the second stage, a non-cooperative game model that incentivizes electric vehicles to participate in supplementary frequency regulation was proposed. Another two-stage scheme was proposed in Yan et al. (2023) to solve the problem of privacy of individual electric owners, the interest competition of different charging stations, and the restriction of coupled distribution networks. In the first stage, the total electric vehicle power flexibility region was derived by solving the optimization problem. In the second stage, a distributed coordination mechanism with a clear physical interpretation was established considering the network constraints based on AC power flow. To simulate a real customer EV distribution scenario, Shi et al. (2022) established a dynamic EV dispatching model considering the multi-source data association of customers, vehicles, charging stations, and service stations. Tao et al. (2022) proposed an improved generative adversarial network (GAN)-based coordinated scheduling strategy for electric vehicles with thermostatically controlled loads (TCLs). Machine learning was integrated into a two-layer optimization problem to determine the steady-state power dispatch and energy storage control of the energy storage systems. Feizi et al. (2022) presented a framework for determining feasible dispatch limits for solar photovoltaic (PV) generation in an unbalanced distribution network considering EV interconnection and associated uncertainties. To ensure frequency stability while reducing load shedding, an optimization strategy based on a two-layer confidence interval for EV participation in ultra-high-speed transmission systems was proposed in Liu et al. (2022a). Due to problems such as computational complexity, these studies focused on electric vehicles as grid participants but ignored the characteristics of transportation participants.

On the other hand, with the development of electric vehicles in power electronics and mechanics, related technologies such as vehicle-to-grid (V2G) (Yue et al., 2023) (Shang et al., 2022), vehicle-to-building (V2B) (Liu et al., 2022b), and battery swapping (Wu, 2022) (Jia et al., 2022) are also applied in network dispatching. Aiming at the randomness of current EV users' participation in V2G, an EV cluster evaluation index model was established in Yue et al. (2023) by analyzing the impact of various pieces of information declared by EV users on the scheduling plan of aggregators and using the declared scheduling power, user credit, battery loss, and user participation as evaluation indexes. Due to the high computational complexity of large-scale EVs and impracticality of future power data collection, Shang et al. (2022) proposed a distributed edge computing framework to ensure the efficiency of scheduling and flexible availability of raw data sets. To minimize cost and maximize satisfaction, Liu et al. (2022b) combined the building energy demand and the safety/willingness of EVs to find and dispatch the optimal vehicle for auxiliary or support actions. Wu (2022) reviewed the state-of-the-art literature and business models for battery swapping stations (BSSs), which provide rechargeable batteries for upcoming electric vehicles in low-battery states. Yang et al. (2019) proposed a shared battery station (SBS) model, which is a multifunctional facility with charging, discharging, dormancy, and switching functions. Argiolas et al. (2022) evaluated the potential business case for battery storage systems for PV-assisted electric vehicle fast charging stations with integrated market applications and services. In Jia et al. (2022), a hybrid model of continuous time-domain battery exchange charging and logistics scheduling was proposed to optimize the joint operation of battery charging and exchange systems. In general, leveraging these new technologies can help solve the charging problem for EVs.

On the other hand, EVs, as transportation participants, are also subject to the rules of the transportation network. Diaz-Cachinero et al. (2020) proposed an enhanced operational planning model for EV routing and charging that considers technical and economic realistic constraints such as battery degradation, speed-dependent power consumption, and penalties for non-performance. Liu et al. (2019) formulated a traffic network charging navigation considering stochasticity based on dynamic programming. A simplified charging control algorithm was proposed to address the computational complexity of the navigation system. To investigate the vehicle–vehicle interaction of multiple vehicles in path planning and charging, an EV charging path optimization method based on an event-driven pricing strategy was proposed in Xiang et al. (2022). To address the charging time-consuming energy efficiency problem and the negative energy cost of regenerative braking in the road network, Wang et al. (2013) proposed an energy-driven and environment-aware EV route planning framework. A charging facility planning model coupled with the transportation and distribution system that considers traffic congestion was proposed in Zhang et al. (2019). Morlock et al. (2020) proposed a strategy for computing time-optimal routes for EVs. A simplified road network was first obtained from a leading routing service. Then, a detailed fuel consumption model was applied to solve the multi-objective shortest path problem using the improved Moore–Bellman–Ford algorithm. Zhang et al. (2020) used a two-loop structure to find the route that leads to the lowest energy consumption. In the outer loop, the path with the lowest energy consumption was obtained by minimizing the difference between the value function of the current round and the optimal values of the previous rounds. In the inner loop, the energy consumption metric of the plug-in hybrid vehicle power management under each feasible path was trained using reinforcement learning (RL). Most of the studies on EV distribution and path planning problems have focused on refined modeling in the direction of energy consumption and charging costs. However, path planning for large-scale vehicles is lacking due to the computational complexity requirements.

In summary, existing studies focus on scheduling or navigation unilaterally. Based on the computational complexity, traditional methods are not convenient for carrying out path planning for collaborative navigation of large-scale vehicles. A few studies have dealt with the coupling of transportation–energy systems, but they are also limited by the aforementioned scale problems and lack a realistic portrayal of the two networks.

This paper studies the navigation and dispatching problems for cluster EVs in order to create a highly reliable low-carbon traffic–power system. The main contributions of this paper are as follows:

- A charging and discharging co-navigation model is proposed. There are charging and discharging vehicles and other vehicles

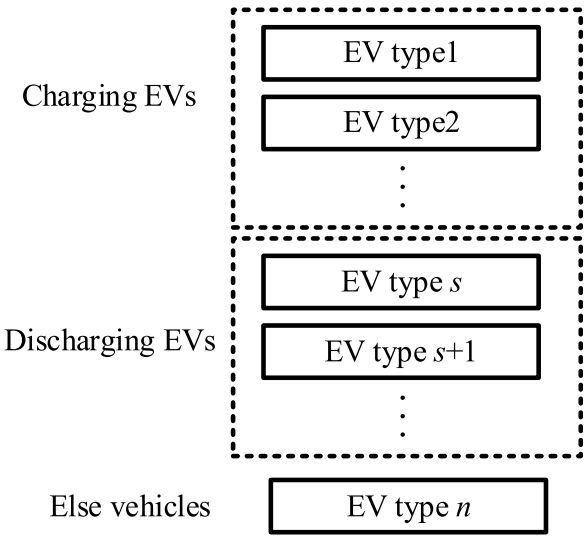


FIGURE 1
Classification of vehicles.

in the traffic network. A cell transmission model (CTM) is used to simulate the real road network. The navigation contains the path planning for charging vehicles and discharging vehicles. The charging vehicles reflect the power demand, and the discharging vehicles support the operation of the distribution network through V2G.

- A clustered electric vehicle dispatching model is proposed for a reliable low-carbon coupled traffic–power system. A cluster EV dispatching model is proposed for improving the overall charging efficiency of EVs on the transportation network and increasing the voltage level of the distribution network. After the electric vehicles arrive at the charging station, the distribution network acts as the control body for orderly charging of the EVs. The charging/discharging power is controlled by an optimization model.

The rest of the paper is organized as follows. The charging and discharging co-navigation model for improving the overall charging efficiency of EVs on the transportation network and increasing the voltage level of the distribution network is presented in Section 2. Section 3 presents the EV cluster dispatching model. The case study is demonstrated in Section 4. Finally, conclusions are drawn in Section 5.

2 Charging and discharging co-navigation models

2.1 Establishment of the traffic system

A low-carbon traffic system refers to the low energy consumption and low emission transportation mode with the goal of reducing greenhouse gas emissions from transportation behavior and is a way to realize low-carbon economy in the transportation field. Among them, the popularity of electric vehicles is the most crucial part.

In this section, there are three participating subjects in the transportation network: electric vehicles with charging demand, electric vehicles that can participate in V2G discharge, and other vehicles. To simplify the vehicle models of different types, we have

$$\Lambda_s = \{p_s, e_s, c(s)\} \tag{1}$$

$$\Gamma = \{\Lambda_1, \Lambda_2, \ldots \Lambda_s, \ldots \Lambda_n\} \tag{2}$$

where Λ_s denotes the properties of vehicles of type s, p_s denotes the maximum charge/discharge power of vehicle type s, e_s denotes the energy margin of vehicle type s, $c(s)$ is the set of terminal cells that EVs of type s finally reach, and Γ is the set of all vehicle types.

It is worth noting that Λ_s represents only one property of the vehicle type s. As shown in Figure 1, if the vehicles have the same power demand value and driving endpoint decision, they are considered to belong to the same type. The existence of multiple vehicles of the same type in the traffic network is considered, implying all these vehicles satisfy the same constraints. Assuming that the power loss during the navigation process is ignored, the energy margin of charging vehicles is less than 0 and finally should arrive at the charging station. The discharged vehicle is used to support the operation of the distribution network through V2G, so its value is positive and eventually arrives at the charging station as well. Other vehicles are not part of the dispatching object of this study, so there is no charging demand. To simplify the modeling, it is considered that the vehicles are eventually allowed to reach any node. Then, three vehicle subjects satisfy the following relations:

$$p_s > 0, \quad e_s < 0, \quad c(s) \subseteq \psi^{\mathrm{CS}}, \quad \forall s \in V^{\mathrm{char}} \tag{3}$$

$$p_s > 0, \quad e_s > 0, \quad c(s) \subseteq \psi^{\mathrm{CS}}, \quad \forall s \in V^{\mathrm{dis}} \tag{4}$$

$$p_s = 0, \quad e_s = 0, \quad \forall c(s), \quad \forall s \in V^{\mathrm{else}} \tag{5}$$

where ψ^{CS} is the set of all cells of charging stations. V^{char} is the set of all charging vehicles. V^{dis} is the set of all discharging vehicles. V^{else} is the set of all other vehicles. Figure 2 illustrates how the traffic network can be converted into a cell structure.

Figure 2A shows the model transforming a section of a one-way road into cells. Assuming that the road section has a travel time of t without considering traffic congestion, it can be divided into multiple cell strings with each travel period τ, satisfying the sum of travel times equaling t.

Figure 2B shows the transformation of a simple road network structure into a cell model. In this case, there are crossed cells that accept path inputs from multiple cells, or outputs that point to multiple cells, which are called parent cells and child cells, respectively.

2.2 Navigation operating strategy

Based on the establishment of the traffic network, the navigation of the charging and discharging vehicles satisfies the CTM constraints described in Wang et al. (2020). We have

$$x^{\mathrm{cell}}_{a,t,s} = x^{\mathrm{cell}}_{a,t-1,s} + \sum_{c \in \alpha(a)} y^{\mathrm{cell}}_{ca,t-1,s} - \sum_{b \in \beta(a)} y^{\mathrm{cell}}_{ab,t-1,s}, \forall a, \quad \forall t, \quad \forall s \tag{6}$$

$$\sum_{b \in \beta(a)} y^{\mathrm{cell}}_{ab,t,s} \le x^{\mathrm{cell}}_{a,t,s}, \quad \forall a, \quad \forall t, \forall s \tag{7}$$

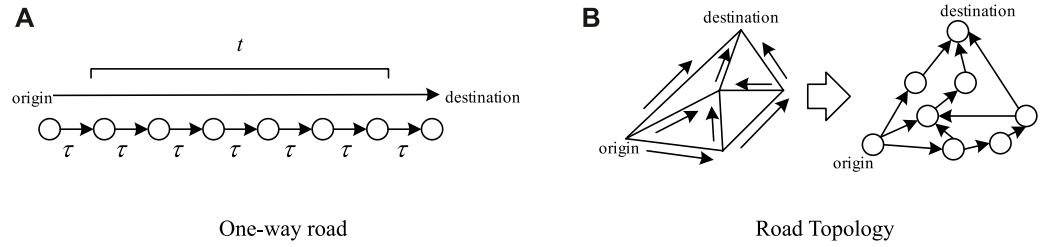


FIGURE 2
Conversion to cell structure.

$$\sum_{\forall s}\sum_{b\in\beta(a)} y_{ab,t,s}^{\text{cell}} \le Q_a^{\max}, \quad \forall a, \forall t \tag{8}$$

$$\sum_{\forall s}\sum_{c\in\alpha(a)} y_{ca,t,s}^{\text{cell}} \le Q_a^{\max}, \quad \forall a, \forall t \tag{9}$$

$$\sum_{\forall s}\sum_{c\in\alpha(a)} y_{ca,t,s}^{\text{cell}} \le k_a\left(N_a^{\max} - \sum_{s} x_{a,t,s}^{\text{cell}}\right), \quad \forall a, \forall t \tag{10}$$

$$y_{ab,t,s}^{\text{cell}} \ge 0, \forall a, \quad \forall b \in \beta(a), \quad \forall t, \quad \forall s \tag{11}$$

where $x_{a,t,s}^{\text{cell}}$ denotes the number of type s vehicles in cell a at time t. $y_{ab,t,s}^{\text{cell}}$ denotes the number of type s vehicles moving from cell a to child cell b at time t. $\alpha(a)$ is the set of all parent cells of cell a. $\beta(a)$ is the set of all child cells of cell a. $Q_a^{\max}$ is the maximum traffic flow of cell a. $N_a^{\max}$ is the maximum traffic amount that cell a can store. k_a is the congestion factor. Constraint (6) represents the traffic flow balance. Constraints (7)–(11) represent the traffic capacity limit.

On the other hand, it is also necessary to consider the changes in the energy of the transportation network. We have

$$e_{a,t,s}^{\text{cell}} = e_{a,t-1,s}^{\text{cell}} + e_s \sum_{c\in\alpha(a)} y_{ca,t,s}^{\text{cell}} - e_s \sum_{b\in\beta(a)} y_{ab,t,s}^{cell} + P_{a,t}^{\text{cs}}, \quad \forall a, \forall t, \quad \forall s \in V^{\text{char}} \tag{12}$$

$$e_{a,t,s}^{\text{cell}} \le \varepsilon, \forall a, \quad \forall t, \quad \forall s \in V^{\text{char}} \tag{13}$$

$$e_{a,end,s}^{\text{cell}} \ge 0, \quad \forall a, \quad \forall s \in V^{\text{char}} \tag{14}$$

where $e_{a,t,s}^{\text{cell}}$ denotes the energy of type s vehicles that cell a could supply at time t. e_s denotes the energy supply of each type s vehicle. $P_{a,t}^{cs}$ denotes the inject active power of cell a at time t. ε is an error factor. Constraint (12) represents the energy flow balance. Constraint (13) avoids the overcharging of vehicles. Constraint (14) ensures that at the end of the entire time scale, all charging actions must be completed. Moreover, for the initial moment, all conditions should be given. We have

$$x_{a,1,s}^{\text{cell}} = X_{a,s}, \quad \forall a, \forall s \tag{15}$$

$$e_{i,1,s}^{\text{cell}} = e_s X_{a,s}, \quad \forall a, \forall s \tag{16}$$

where $X_{a,s}$ denotes the number of type s vehicles in cell a at the initial moment.

Finally, charging and discharging vehicles need to satisfy the following path planning constraints:

$$\sum_{a\in c(s)} x_{a,end,s}^{\text{cell}} = N_s^V, \quad \forall s \in V^{\text{char}} or V^{\text{dis}} \tag{17}$$

where N_s^V denotes the number of type s vehicles in the traffic system.

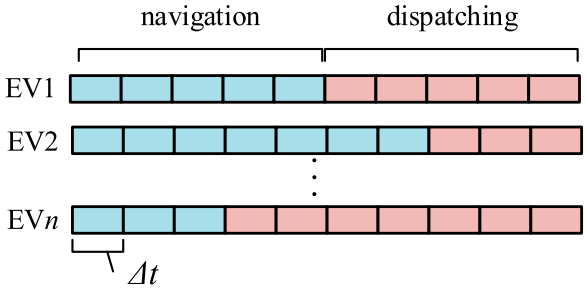


FIGURE 3
Coupling of navigation and dispatching on the time scale.

3 EV cluster dispatching model

3.1 Coupling for the traffic–grid system

When the electric vehicle arrives at the charging station node, the distribution network acts as the control master and unifies the charging power control. Figure 3 illustrates the coupling of the navigation process and the dispatching process on the time scale.

Sequential charging satisfies the following constraints:

$$-\sum_{m\in V^{\text{dis}}} p_m x_{a,t,m}^{\text{cell}} \le P_{a,t}^{\text{cs}} \le \sum_{n\in V^{\text{char}}} p_n x_{a,t,n}^{\text{cell}}, \quad \forall a, \forall t \tag{18}$$

$$-\sum_{m\in V^{\text{dis}}} q_m x_{a,t,m}^{\text{cell}} \le Q_{a,t}^{\text{cs}} \le \sum_{n\in V^{\text{char}}} q_n x_{a,t,n}^{\text{cell}}, \quad \forall a, \forall t \tag{19}$$

where $Q_{a,t}^{\text{cs}}$ denotes the inject reactive power of cell a at time t. q_s denotes the maximum reactive power for each type s vehicle. Constraints (18) and (19) ensure that the lower limit of the node injection power is the sum of the maximum power of the V2G and the upper limit is the sum of the maximum charging power of all charging vehicles.

In addition, the distribution network model constraints should be satisfied:

$$\sum_{j\in\lambda(i)} P_{ij,t} - \sum_{k\in\pi(i)} P_{ki,t} = P_{i,t}^{\text{in}} - P_i^{\text{load}}, \quad \forall i, \quad \forall t \tag{20}$$

$$\sum_{j\in\lambda(i)} Q_{ij,t} - \sum_{k\in\pi(i)} Q_{ki,t} = Q_{i,t}^{\text{in}} - Q_i^{\text{load}}, \quad \forall i, \quad \forall t \tag{21}$$

$$P_{i,t}^{\text{in}} = P_{i,t}^{\text{gen}} - P_{a,t}^{\text{cs}}, \quad \forall (i,a) \in \mathbb{C}, \quad \forall t \tag{22}$$

$$Q_{i,t}^{\text{in}} = Q_{i,t}^{\text{gen}} - Q_{a,t}^{\text{cs}}, \quad \forall (i,a) \in \mathbb{C}, \quad \forall t \tag{23}$$

$$U_{i,t} - U_{j,t} = P_{ij,t}R_{ij} + Q_{ij,t}X_{ij}, \quad \forall (i,j) \in \Omega^{\mathrm{L}} \tag{24}$$

$$\sqrt{{P_{ij,t}}^2 + {Q_{ij,t}}^2} \le {S_{ij}}^{\max}, \quad \forall (i,j) \in \Omega^{\mathrm{L}} \tag{25}$$

$${U_i}^{\min} \le U_{i,t} \le {U_i}^{\max} \tag{26}$$

where $P_{ij,t}$ and $Q_{ij,t}$ denote the active/reactive power flowing from node i to node j at time t. $P_{i,t}^{\mathrm{in}}$ and $Q_{i,t}^{\mathrm{in}}$ denote the injected power of node i at time t. $P_{i,t}^{\mathrm{load}}$ and $Q_{i,t}^{\mathrm{load}}$ denote the active/reactive load power at node i at time t. $P_{i,t}^{\mathrm{gen}}$ and $Q_{i,t}^{\mathrm{gen}}$ denote the active/reactive power from the generator of node i at time t. $\mathbb{C}$ is the set of all the coupling pairs of all nodes and cells, where $(i,a) \in \mathbb{C}$ indicates that there is a coupling between node i and cell a. $U_{i,t}$ denotes the amplitude voltage of node i at time t. R_{ij} and X_{ij} denote the resistance/reactance of branch (i,j). ${S_{ij}}^{\max}$ denotes the maximum capacity of branch (i,j). ${U_i}^{\max}$ and ${U_i}^{\min}$ denote the upper and lower limits of voltage at node i. $\lambda(i)$ is the set of all the children nodes of node i. $\pi(i)$ is the set of all parent nodes of node i. Ω^{L} denotes the set of all branches. Constraints (20)–(23) indicate the power balance of the distribution network, where the injected power of the cell corresponds to the outflow power of the distribution network node. Constraint (24) indicates the voltage drop balance. Constraint (25) indicates the line capacity limit. Constraint (26) indicates the voltage amplitude limit.

3.2 Optimization models

For the traffic system, the shortest navigation time of total charging EVs is guaranteed to reduce the impact of charging vehicles on the road conditions of the traffic network. That is, all charging electric vehicles are present at the charging station for the longest time in the whole time scale of dispatch. Thus, we have

$$\max F_1 = \sum_{s \in V^{\mathrm{char}}} \sum_{i \in \psi^{CS}} \sum_{\forall t} x_{i,t,s}^{\mathrm{cell}} \tag{27}$$

For the power system, the guaranteed voltage level is more likely to be close to the rated value, which is used to measure the reliability of the power system. Thus, we have

$$\min F_2 = \sum_{\forall t} \sum_{i \in \Omega^{\mathrm{N}}} \left|1 - U_{i,t}\right| \tag{28}$$

where Ω^{N} is the set of all nodes in the distribution network. Since constraint (28) is a non-linear equation containing an absolute value function, a linearization transformation is applied. Let auxiliary variables $z_{i,t}$ replace $|1 - U_{i,t}|$. We have

$$\min F_2 = \sum_{\forall t} \sum_{i \in \Omega^{N}} z_{i,t} \tag{29}$$

$$z_{i,t} \ge 1 - U_{i,t} \tag{30}$$

$$z_{i,t} \ge U_{i,t} - 1 \tag{31}$$

In summary, the optimization model is established as follows:

$$\mathrm{obj} = \omega_1 \sum_{\forall t} \sum_{i \in \Omega^{N}} z_{i,t} - \omega_2 \sum_{s \in V^{char}} \sum_{i \in \psi^{CS}} \sum_{\forall t} x_{i,t,s}^{\mathrm{cell}} \tag{32}$$

The constraints include

$$(6) - (26), (30) - (31) \tag{33}$$

Constraint (25) is a second-order cone-convex function, and all the remaining constraints are linear. Therefore, the optimization model can be solved by directly calling the common solvers.

4 Case study

4.1 Case description

The simulation is set up with a modified road network structure with 357 cells, coupled with modified IEEE33 nodes. The three charging stations in the traffic network are located in cells corresponding to nodes 7, 22, and 24 of the IEEE33 node power network. In this case, 70 EVs with charging demand and 10 discharged EVs with V2G willingness are set up in the traffic network. A certain number of other vehicles are generated by a random model in each cell. Set the time scale of the whole scheduling process to 40. Each time scale corresponds to a 3-min time interval. The maximum traffic flow $Q_i^{\max}$ is set at a uniform rate of 3. The congestion factor k_a is 1. The maximum number of vehicles per cell $N_i^{\max}$ is set to 10. In particular, for the traffic network end cells or charging station cells, let $N_i^{\max} = 30$. Generators of the distribution network are set at nodes 1 and 15. Assuming that all charging vehicles or discharging vehicles have the same charging demand and charging station arrival end condition, three types of vehicles were established. The maximum power of charging/discharging is set to 40 kW, and the charging demand is 10 kWh. The dispatch interval for each time scale is 3 min, for a total of 40 time scales. The discharging vehicle can provide a maximum power of 10 kWh. Let $\boldsymbol{\omega_1 = 1, \omega_2 = 10}$. A schematic diagram of the coupling is shown in Figure 4. The road conditions and topology of the traffic network are shown in Figure 5. The simulation platform is MATLAB 2019b and Gurobi 9.5.2.

4.2 Navigation simulation results

Figure 6 shows the partial time-scale results for the charging co-navigation of 70 EVs. The red nodes represent electric vehicles, and the blue nodes represent charging stations. There is only one type of charging vehicle in this example, and it is stipulated that all charging vehicles can eventually reach any charging station node. For visual differentiation, the random generation method of vehicles specifies that each vehicle is at a different cell location. Moment 0 shows the locations of all vehicles with charging demand at the initial moment. As collaborative navigation proceeds, the number of vehicles traveling in the topology keeps decreasing. When 105 min is reached, all vehicles in this round arrive at the charging station. It can be seen that vehicles generally tend to lead to the charging station with the shortest single time. This is the result given by the objective function based on the shortest dispatch time at the traffic level.

Figure 7 shows the navigation results for all V2G vehicles. It can clearly be seen that nine EVs lead to charging Station 2, while only one vehicle leads to charging Station 1, and no vehicle leads to charging Station 3. In contrast to the path planning of charging EVs, there is no clear tendency for V2G vehicles to travel on the shortest path. Further analysis shows that it is the reliability of the distribution network that acts on the traffic dispatch and has an

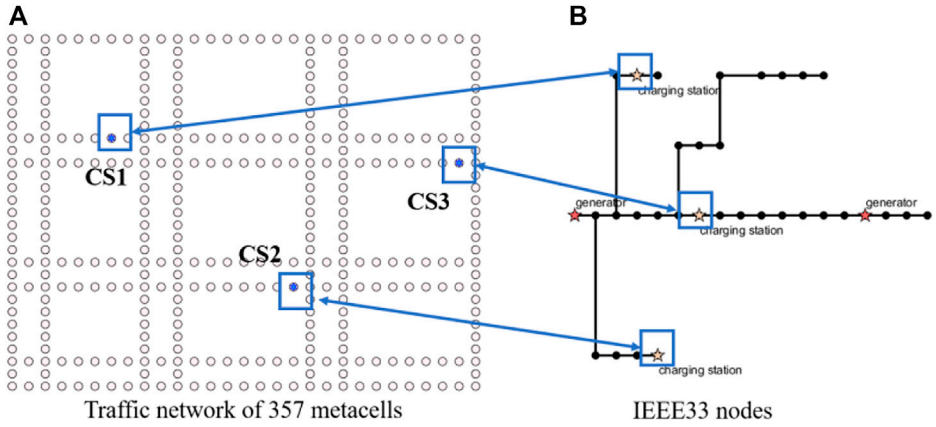


FIGURE 4
Coupling of the traffic–power system.

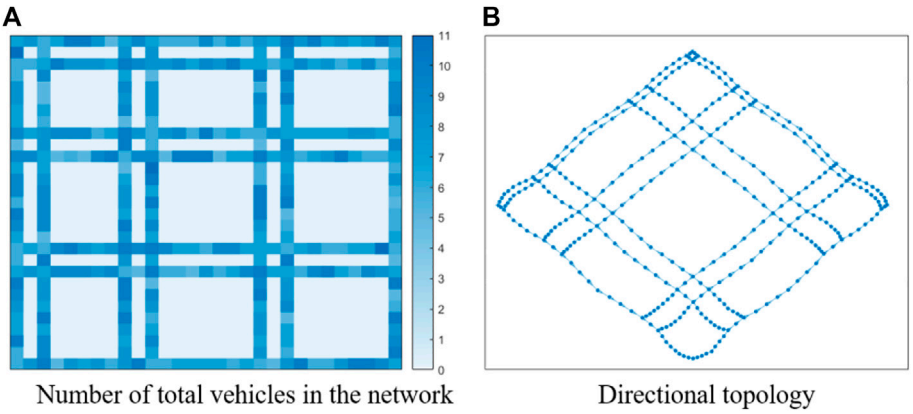


FIGURE 5
Specific parameters of the traffic network.

impact on path planning. Figure 8 shows the variation in the number of vehicles at the charging station nodes. It is worth noting that since the arithmetic example is a one-way topological road network structure, the number of vehicles at the charging station nodes can reflect the variation of the congestion level of the whole road where they are located. As the navigation process proceeds, it can be seen that the charging station nodes gradually become congested. Therefore, cluster scheduling will increase the traffic burden around the charging stations.

In fact, by observing the distribution of vehicles at all time scales, all charging vehicles almost always pass at the optimal time, and at the intuitively optimal distance. Congestion hardly affects vehicle path planning, but can exacerbate traffic network congestion.

4.3 Cluster dispatching results

Figure 9 illustrates the power variation curve of charging stations. The maximum actual power of charging Station 1 reaches 1120 kW and that of charging Station 2 reaches 680 kW, while the maximum actual charging power of charging Station 3 is only 257 kW. Since this scheduling process incorporates an objective function indicating voltage reliability, it is obvious that the charging power of charging Station 3 is very flat. In contrast, charging Station 1 and Station 2 tend to proceed at maximum power for fast charging. Further analysis shows that the staggered charging moments of charging Station 1 and Station 2 with maximum power are carried out. Therefore, it can be concluded that the distribution network can accommodate more capacity at charging Station 1 and Station 2 nodes than at Station 3 nodes, when the main factor affecting the voltage level of the distribution network is the topology of the distribution network's own power distribution rather than the charging power distribution, provided there is a capacity margin. Excessive power will aggravate the line voltage drop at charging Station 3, which is more likely to cause a lower voltage reliability level. As observed in Figures 9A, B, the reason for such a sharp increase in charging power is that the node power has not yet reached the maximum capacity of the distribution network. The

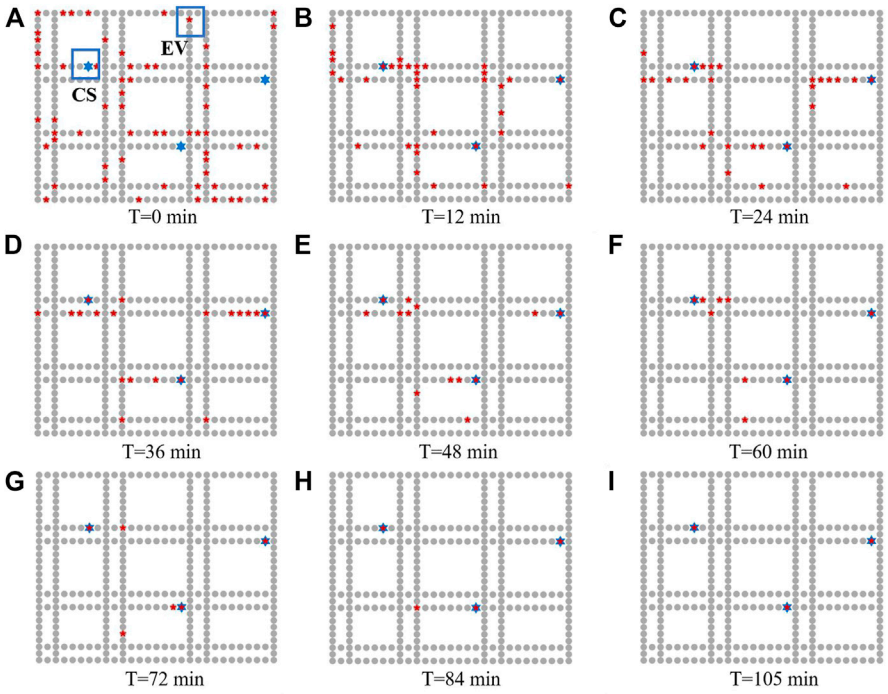


FIGURE 6
Co-navigation process for 70 charging electric vehicles.

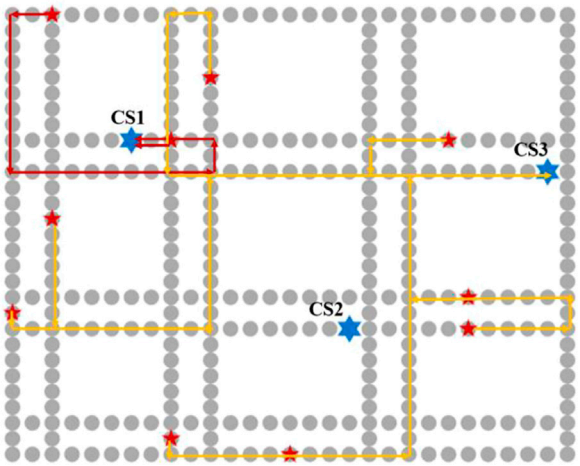


FIGURE 7
Co-navigation process for 10 discharging EVs.

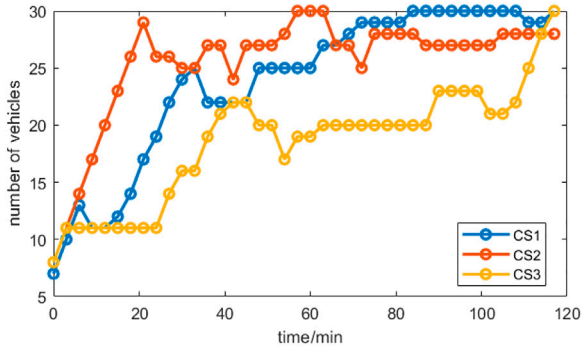


FIGURE 8
Changes in the number of vehicles at the nodes of charging stations.

theory of convex optimization suggests that if there is an optimal solution to a mathematical plan, it must be on the boundary conditions of the feasible domain.

Comparative experiments were carried out with or without considering voltage levels. The control experiments remove the objective function term characterizing the voltage level in Eq. 28. Figure 10 shows the comparison of the dispatch with/without considering voltage reliability. The voltage levels of all nodes are selected at 3 min, 57 min, and 114 min for the entire 2-h scheduling process. It is clear that in the case of scheduling without considering voltage reliability, the voltage easily reaches the upper bound of 1.1 and the lower bound of 0.9, while the latter voltage level is closer to the nominal value and never reaches the upper or lower bound. The highest voltage is only around 1.05, while only node 16 reaches a minimum voltage of 0.9. On the other hand, without considering voltage reliability, the voltage level varies significantly throughout the dispatch process, and we consider this to be a less reliable low-carbon traffic–power system. As shown in Figure 10A, the voltage level is relatively normal in the early stage at 3 min and 57 min, but deteriorates in the later stage, which is because the EVs are still in the

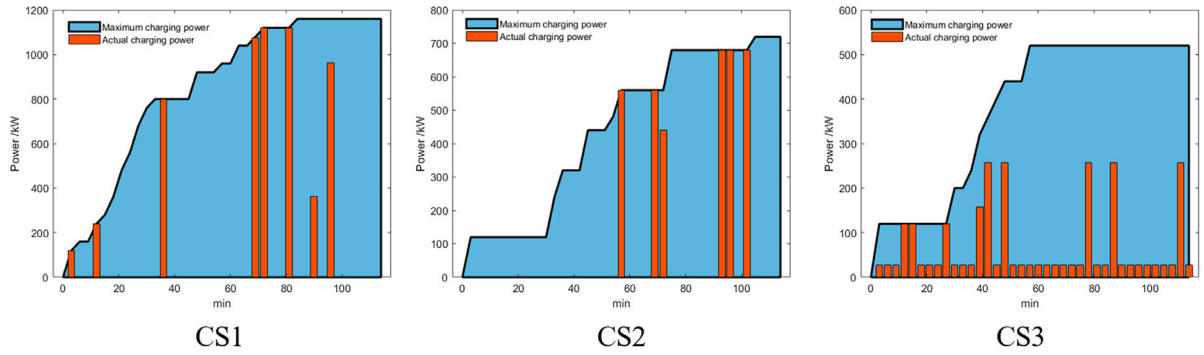


FIGURE 9
Charging power variation of charging stations.

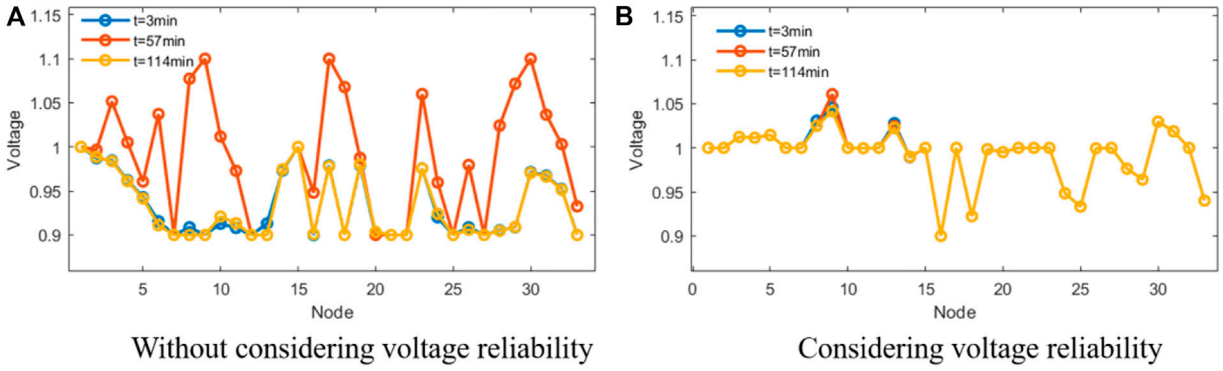


FIGURE 10
Comparison of voltage reliability.

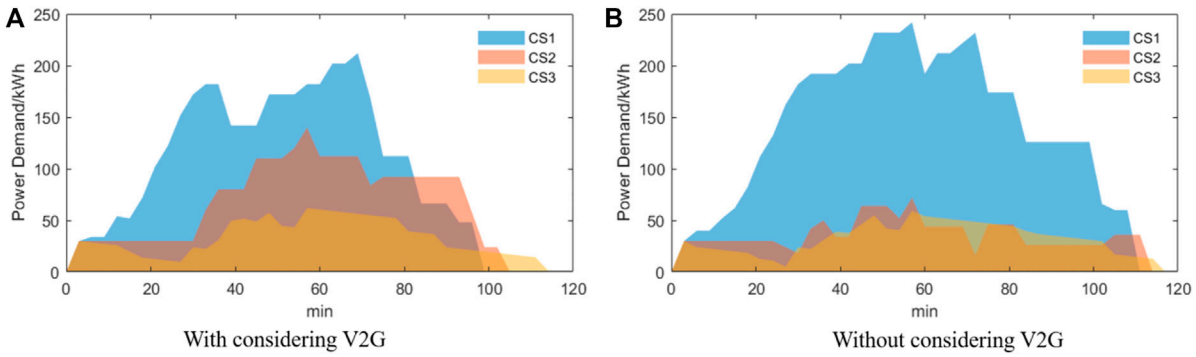


FIGURE 11
Changes in power demand during dispatching.

navigation stage in the early stage and the charging station node load level is low. Also, when they arrive at the charging station, they are charged in a disorderly manner, which significantly increases the random load of the charging station.

Comparative experiments were carried out with and without V2G vehicles. The control experiment removes the dispatch of 10 V2G vehicles from the traffic network. Figure 11 shows the changes in power demand at the three charging station nodes throughout the dispatch process. The charging Station 2 node is generally a regular process where the power demand increases first and then decreases, which is because the charging demand increases as the EVs keep reaching this node; while on the other hand, the

TABLE 1 Computational time analysis by different cases.

	Case A	Case B	Case C	Case D
Optimization time (seconds)	106.977669	62.152826	113.649501	47.258919

charging demand decreases at this point because of the unified charging and discharging scheduling management after reaching this node. It is worth noting that charging demand does not absolutely increase first and then decrease, which is because navigation and scheduling is a coupled process that takes place simultaneously. The power demand variation curve of charging Station 1 and Station 3 is more irregular, which is due to the addition of V2G vehicles, which also relieves the pressure of power demand.

Comparing Figure 11A with Figure 11B, it is found that the most significant change without the addition of V2G vehicles is a significant increase in the load accommodated at charging Station 1, with the maximum demand increasing from 200 kWh to nearly 250 kWh. There is little change at charging Station 3 and a significant decrease in the charging load at charging Station 2, with the maximum demand decreasing from 150 kWh to nearly 60 kWh. Therefore, it can be concluded that charging Station 1 is more reliable than charging Station 2 at the power system level. Analyzing scheduling together with navigation, it can be found that V2G vehicles almost always flock to charging Station 3, and the charging power is flat at charging Station 3. Therefore, it can be considered that this is the vulnerability of the distribution network. In the subsequent distribution network planning, a new distributed power supply or increased line capacity near the charging station at this location can be considered to ensure the reliability of the large-scale access of EVs. Ultimately, a rough conclusion can be drawn that the reliability of charging Station 1 is greater than that of charging Station 2, and the reliability of charging Station 2 is greater than that of charging Station 3.

These analyses also provide a reference for charging station planning. The traffic system data can be obtained by extracting real-time traffic flow data and building a cell network. The charging stations at different locations are further simulated to get the optimized ordered charging and discharging power. Based on the reliability of the traffic network as part of the objective function, the ordered charging curve of the charging stations is obtained, and further analysis of the power demand and charging power can find a balance point of power and traffic, ensuring the charging power and charging navigation of the synergy are reasonable.

4.4 Computational time analysis

Furthermore, we choose four different cases under this model to analyze the computational complexity. Case A is the optimization model proposed in Section 3.2. Case B is the modified model without considering grid reliability in Eq. 28. Case C is the modified model without considering traffic efficiency in Eq. 27. Case D is the time required to find the first feasible solution for the optimization model as given in Section 3.2. The program is set up on a personal computer with a four-core i5-8265 CPU and an 8G RAM to further study the computational performance of the proposed model. The optimal gap is set as 0.01%.

Table 1 shows that not considering grid reliability speeds up the runtime, where most of the runtime is used to find the first feasible solution. This is due to the fact that the model is a macroscopic scale and the CTM is essentially concerned with whether and how many cars each cell has, while path planning is achieved through traffic flow only. Therefore, the search for the feasible domain of preliminary convex optimization is time-consuming. In contrast, when a feasible solution is found, collaborative path planning for large-scale vehicles is much faster according to the boundary conditions. Table 1 shows computational time analysis by different cases. It is worth noting that the operation time of Case C is longer than that of Case A, which is the result of the divergence of the boundary conditions due to the change of the objective function, which increases the difficulty of the boundary convergence.

5 Conclusion

This paper presents an electric vehicle navigation and cluster dispatch model proposed for a reliable traffic–power system. A simplified model of the vehicle and traffic road network is established, and a cell transmission model (CTM) is used to simulate the real traffic network. A coupled model of the traffic–power system is built for the orderly charging of electric vehicles upon arrival at a charging station. The model considers the coupling of the two systems on the time scale, and the charging and discharging power at each charging station node is controlled. The validity of the model is verified in a coupled system of 357 cell traffic system and IEEE33 power system. The results show that it can ensure a good guarantee of distribution network voltage reliability and reveal the scheduling process of the traffic network. The model also provides a reference for the planning of charging stations in the distribution network. Through case study verification, the main conclusions are summarized as follows:

- Different from traditional path planning, path navigation based on the cell transmission model could plan a large number of traffic subjects simultaneously, and at the same time, the model can realistically simulate the traffic network and imitate traffic congestion, which has high applicability and practicality.
- The model performs unified charging scheduling for electric vehicles, while considering the scheduling of V2G. Through time-scale coupling, the transportation system is well-coupled with the electric power system, solving the problem that the existing research cannot take into account both systems at the same time.
- The model can also reveal the weaknesses of the distribution network. The comparison of the dispatch endpoints and charging and discharging power of V2G vehicles can reveal the vulnerable nodes corresponding to charging stations in the system. This provides a reference for the subsequent distribution network expansion planning including charging

station planning, line capacity increase, and line new construction.

At the same time, there are shortcomings in this study. Considering the computational complexity, this study does not consider the fuel consumption of the traffic on the way, which cannot be ignored in a large-scale traffic network. The model does not consider the path planning constraints of other vehicles, which can have an impact on the congestion level of the traffic network. Further research will be conducted later to address these issues.

In the future, as electric vehicle penetration rate increases year-by-year, the integration of traffic–power networks will gradually deepen. Joint planning of the traffic–power system needs to be urgently carried out, which can also combine distributed data, such as distributed photovoltaics, to realize the integration of new energy and electrified transportation. It is also necessary to carry out joint planning to combine distributed generator data to realize integrated planning of new energy and electrified transportation to help realize the dual carbon strategy.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material; further inquiries can be directed to the corresponding author.

Author contributions

SW and YX: conceptualization and methodology. SW: writing—original draft preparation. YX, LnC, ZS, SH, MG, HH, JT, YH, and LaC: writing—reviewing and editing. JL: supervision.

Conflict of interest

Authors ZS and SH were employed by Sichuan Road & Bridge Group Co., Ltd. Authors ZS and MG were employed by Sichuan Shudao Clean Energy Group Co., Ltd. Authors JT, YH, and LaC were employed by Chengdu TELD New Energy Co., Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Argiolas, L., Stecca, M., Ramirez-Elizondo, L. M., Soeiro, T. B., and Bauer, P. (2022). Optimal battery energy storage dispatch in energy and frequency regulation markets while peak shaving an EV fast charging station. *IEEE Open Access J. Power Energy* 9, 374–385. doi:10.1109/oajpe.2022.3198553

Diaz-Cachinero, P., Muñoz-Hernandez, J. I., Contreras, J., and Muñoz-Delgado, G. (2020). An enhanced delivery route operational planning model for electric vehicles. *IEEE Access* 8, 141762–141776. doi:10.1109/access.2020.3013144

Dong, C., Chu, R., Morstyn, T., McCulloch, M. D., and Jia, H. (2021). Online rolling evolutionary decoder-dispatch framework for the secondary frequency regulation of time-varying electrical-grid-electric-vehicle system. *IEEE Trans. Smart Grid* 12 (1), 871–884. doi:10.1109/tsg.2020.3020983

Feizi, M. R., Khodayar, M. E., and Chen, B. (2022). Feasible dispatch limits of PV generation with uncertain interconnection of EVs in the unbalanced distribution network. *IEEE Trans. Veh. Technol.* 71 (3), 2267–2280. doi:10.1109/tvt.2021.3096459

Jia, W., Ding, T., Bai, J., Bai, L., Yang, Y., and Blaabjerg, F. (2022). Hybrid swapped battery charging and logistics dispatch model in continuous time domain. *IEEE Trans. Veh. Technol.* 71 (3), 2448–2458. doi:10.1109/tvt.2022.3143370

Li, X., Li, C., Luo, F., Chen, G., Dong, Z. Y., and Huang, T. (2023). Electric vehicles charging dispatch and optimal bidding for frequency regulation based on intuitionistic fuzzy decision making. *IEEE Trans. Fuzzy Syst.* 31 (2), 596–608. doi:10.1109/tfuzz.2022.3220964

Li, X., Xiang, Y., Lyu, L., Ji, C., Zhang, Q., Teng, F., et al. (2020). Price incentive-based charging navigation strategy for electric vehicles. *IEEE Trans. Industry Appl.* 56 (5), 5762–5774. doi:10.1109/tia.2020.2981275

Liu, C., Zhou, M., Wu, J., Long, C., and Wang, Y. (2019). Electric vehicles en-route charging navigation systems: Joint charging and routing optimization. *IEEE Trans. Control Syst. Technol.* 27 (2), 906–914. doi:10.1109/tcst.2017.2773520

Liu, H., Pan, H., Wang, N., Yousaf, M. Z., Goh, H. H., and Rahman, S. (2022). Robust under-frequency load shedding with electric vehicles under wind power and commute uncertainties. *IEEE Trans. Smart Grid* 13 (5), 3676–3687. doi:10.1109/tsg.2022.3172726

Liu, Y., Zhou, P., Yang, L., Wu, Y., Xu, Z., Liu, K., et al. (2022). Privacy-preserving context-based electric vehicle dispatching for energy scheduling in microgrids: An online learning approach. *IEEE Trans. Emerg. Top. Comput. Intell.* 6 (3), 462–478. doi:10.1109/tetci.2021.3085964

Morlock, F., Rolle, B., Bauer, M., and Sawodny, O. (2020). Time optimal routing of electric vehicles under consideration of available charging infrastructure and a detailed consumption model. *IEEE Trans. Intelligent Transp. Syst.* 21 (12), 5123–5135. doi:10.1109/tits.2019.2949053

Shang, Y., Shang, Y., Yu, H., Shao, Z., and Jian, L. (2022). Achieving efficient and adaptable dispatching for vehicle-to-grid using distributed edge computing and attention-based LSTM. *IEEE Trans. Industrial Inf.* 18 (10), 6915–6926. doi:10.1109/tii.2021.3139361

Shi, L., Zhan, Z. -H., Liang, D., and Zhang, J. (2022). Memory-based ant colony system approach for multi-source data associated dynamic electric vehicle dispatch optimization. *IEEE Trans. Intelligent Transp. Syst.* 23 (10), 17491–17505. doi:10.1109/tits.2022.3150471

Tao, Y., Qiu, J., and Lai, S. (2022). A data-driven management strategy of electric vehicles and thermostatically controlled loads based on modified generative adversarial network. *IEEE Trans. Transp. Electrification* 8 (1), 1430–1444. doi:10.1109/tte.2021.3109671

Wang, Y., Jiang, J., and Mu, T. (2013). Context-aware and energy-driven route optimization for fully electric vehicles via crowdsourcing. *IEEE Trans. Intelligent Transp. Syst.* 14 (3), 1331–1345. doi:10.1109/tits.2013.2261064

Wang, Y., Yu, X., Li, J., and He, J. (2020). Dynamic load restoration considering the interdependencies between power distribution systems and urban transportation systems. *CSEE J. Power Energy Syst.* 6 (4), 772–781. doi:10.17775/CSEEJPES.2020.02250

Wu, H. (2022). A survey of battery swapping stations for electric vehicles: Operation modes and decision scenarios. *IEEE Trans. Intelligent Transp. Syst.* 23 (8), 10163–10185. doi:10.1109/tits.2021.3125861

Xiang, Y., Wang, Y., Xia, S., and Teng, F. (2021). Charging load pattern extraction for residential electric vehicles: A training-free nonintrusive method. *IEEE Trans. Industrial Inf.* 17 (10), 7028–7039. doi:10.1109/tii.2021.3060450

Xiang, Y., Yang, J., Li, X., Gu, C., and Zhang, S. (2022). Routing optimization of electric vehicles for charging with event-driven pricing strategy. *IEEE Trans. Automation Sci. Eng.* 19 (1), 7–20. doi:10.1109/tase.2021.3102997

Yan, D., Ma, C., and Chen, Y. (2023). Distributed coordination of charging stations considering aggregate EV power flexibility. *IEEE Trans. Sustain. Energy* 14 (1), 356–370. doi:10.1109/tste.2022.3213173

Yang, J., Wang, W., Ma, K., and Yang, B. (2019). Optimal dispatching strategy for shared battery station of electric vehicle by divisional battery control. *IEEE Access* 7, 38224–38235. doi:10.1109/access.2019.2906488

Yue, H., Zhang, Q., Zeng, X., Huang, W., Zhang, L., and Wang, J. (2023). Optimal scheduling strategy of electric vehicle cluster based on index evaluation system. *IEEE Trans. Industry Appl.* 59 (1), 1212–1221. doi:10.1109/tia.2022.3213639

Zhang, Q., Wu, K., and Shi, Y. (2020). Route planning and power management for PHEVs with reinforcement learning. *IEEE Trans. Veh. Technol.* 69 (5), 4751–4762. doi:10.1109/tvt.2020.2979623

Zhang, X., Li, P., Hu, J., Liu, M., Wang, G., Qiu, J., et al. (2019). Yen's algorithm-based charging facility planning considering congestion in coupled transportation and power systems. *IEEE Trans. Transp. Electrification* 5 (4), 1134–1144. doi:10.1109/tte.2019.2959716

Nomenclature

Index and sets

ψ^{CS} Set of all cells of charging stations.

V^{char} Set of all charging vehicles.

V^{dis} Set of all discharging vehicles.

V^{else} Set of all other vehicles.

Ω^{L} Set of all branches.

$\beta(a)/\alpha(a)$ Set of all children/parent cells of cell a.

$\mathbb{C}$ Set of all the coupling pair of nodes and cells.

$\mathbb{\lambda}(i)/\pi(i)$ Set of all the children/parent nodes of node i.

Ω^{N} Set of all nodes in the distribution network.

Constants

$Q_a{}^{max}$ Maximum traffic flow of cell a.

$N_a{}^{max}$ Maximum traffic amount that cell a can store.

k_a Congestion factor.

e_s Energy margin of vehicle type s.

ε Error factor.

$X_{a,s}$ Number of type s vehicles at cell a at the initial moment.

N_s^V Number of type s vehicles in the traffic system.

p_s/q_s Maximum active/reactive power for each type s vehicle charging/discharging.

Variables

$x_{a,t,s}^{cell}$ Number of type s vehicles in cell a at time t.

$y_{ab,t,s}^{cell}$ Number of type s vehicles moving from cell a to child cell b at time t.

$e_{a,t,s}^{cell}$ Energy of all type s vehicles in cell a that could supply at time t.

$P_{a,t}^{cs}/Q_{a,t}^{cs}$ Inject active/reactive power of cell a at time t.

$P_{i,t}^{in}/Q_{i,t}^{in}$ Inject power of node i at time t.

$P_{i,t}^{gen}/Q_{i,t}^{gen}$ Active/reactive power from the generator of node i at time t.

$P_{ij,t}/Q_{ij,t}$ Active/reactive power flowing from node i to node j at time t

$U_{i,t}$ Amplitude voltage of node i at time t.

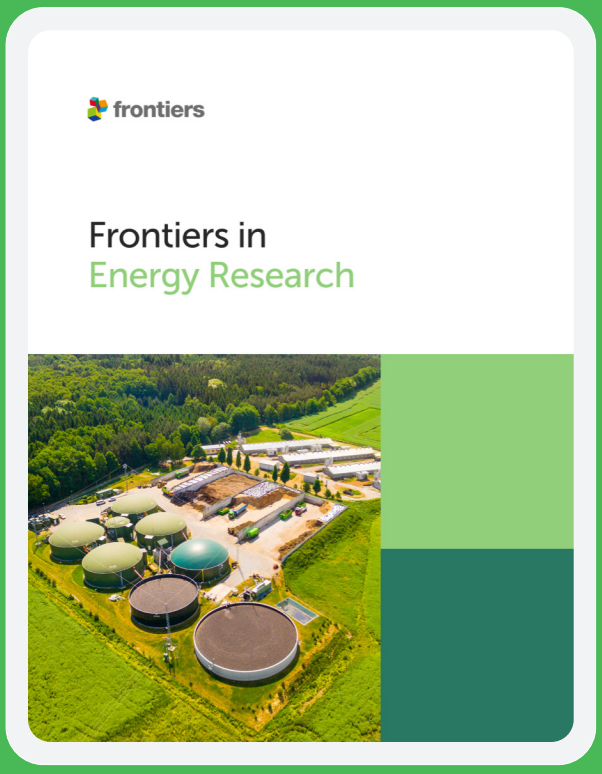
Frontiers in
Energy Research

Advances and innovation in sustainable, reliable and affordable energy

Explores sustainable and environmental developments in energy. It focuses on technological advances supporting Sustainable Development Goal 7: access to affordable, reliable, sustainable and modern energy for all.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact