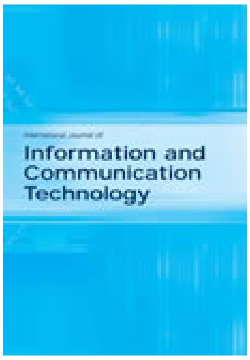


International Journal of Information and Communication Technology

ISSN online: 1741-8070 - ISSN print: 1466-6642
https://www.inderscience.com/ijict

Simulation and evaluation of green power consumption policies driven by spatio-temporal graph convolutional networks

Jie Jiao, Jiyuan Zhang, Wenshi Ren

DOI: 10.1504/IJICT.2025.10074817

Article History:

Received:	13 August 2025
Last revised:	26 September 2025
Accepted:	27 September 2025
Published online:	12 December 2025

Simulation and evaluation of green power consumption policies driven by spatio-temporal graph convolutional networks

Jie Jiao*, Jiyuan Zhang and Wenshi Ren

Economic Research Institute,
State Grid Sichuan,
Chengdu, 610041, China
Email: jiaojie_jyy@sina.com
Email: zhangjiyuan_jyy@126.com
Email: renwenshi_jyy@126.com
*Corresponding author

Abstract: Green power consumption has become a key challenge in the energy transition. Existing research struggles to capture complex relationships between the spatio-temporal dynamics of the power system and policy interventions. To this end, this paper first designs a power load forecasting model based on spatio-temporal graph convolutional networks. The model dynamically adjusts the graph structure according to users' electricity consumption patterns and introduces a weighted skip connection mechanism, assigning different weights to connections at different time steps. Then, a mathematical model for optimal combinations of power consumption policies is established. Through deep reinforcement learning algorithms interacting with the environment, it solves for the optimal combination of power consumption policies that minimise economic and carbon emission costs. Experimental outcome demonstrates that the proposed method achieves a green power consumption rate of 97.16%, outperforming comparison methods, thus helping to promote efficient green power consumption.

Keywords: green power; consumption policy; spatio-temporal graph convolutional network; deep reinforcement learning algorithms; skip connections.

Reference to this paper should be made as follows: Jiao, J., Zhang, J. and Ren, W. (2025) 'Simulation and evaluation of green power consumption policies driven by spatio-temporal graph convolutional networks', *Int. J. Information and Communication Technology*, Vol. 26, No. 45, pp.34–51.

Biographical notes: Jie Jiao obtained his Master's degree from North China Electric Power University in 2018. Currently, he serves as a Senior Economist at the State Grid Sichuan Economic Research Institute. His research interests include the electricity market and electricity price mechanisms.

Jiyuan Zhang obtained her Master's degree from Chongqing University in 2013. She is currently a Senior Economist at the State Grid Sichuan Economic Research Institute. Her research interests lie in the technical and economic aspects of power grids.

Wenshi Ren obtained her Master's degree from Sichuan University in 2020. She is currently an Economist at the State Grid Sichuan Economic Research Institute. Her research interests lie in the technical and economic aspects of power grids.

1 Introduction

Green power consumption has become a core issue in achieving the dual-carbon goals and promoting energy sustainability (Khalil and Sheikh, 2024). The intermittent and fluctuating characteristics of renewable energy sources such as wind and solar power are increasingly conflicting with the requirements for safe and stable operation of power systems, leading to frequent curtailment of wind and solar power, which seriously hinders the large-scale application of green energy (Rashid, 2024). Under this context, scientifically formulating and optimising green power consumption policies has become a key aspect in balancing energy structure adjustment with the economy and reliability of the power system (Constantinou et al., 2022). Traditional policy evaluation methods often rely on experience-based models or static data analysis, making it difficult to accurately capture the spatio-temporal dynamic characteristics in power systems (Press and Arnould, 2009). For example, the time-delay effects of regional power transmission, the spatio-temporal differences in load demand distribution, and the complex interactions of random fluctuations in renewable energy output. These factors make it challenging for policymakers to predict the implementation effects of different policy combinations, potentially leading to resource misallocation or failure to fully realise policy benefits (Lu and Park, 2022). How to achieve dynamic simulation of key indicators of green power under different policy scenarios, ensuring the safe, efficient, and low-carbon operation of power systems, is a topic of significant practical value (Vakili et al., 2021).

In the field of simulation and evaluation of green power consumption policies, most studies first predict user electricity loads, and then establish an economic dispatch optimisation model of the power system based on the prediction results to obtain the optimal policy combination, thereby achieving the optimal system operation. Traditional methods are based on statistical research approaches. Chodakowska et al. (2021) established an autoregressive integrated moving average (ARIMA) model (Nepal et al., 2020), analysed the advantages and disadvantages of the ARIMA model in power load forecasting tasks, and used the forecasting results as constraints for low-carbon operation of the power system, obtaining the optimal policy combination results through genetic algorithm (GA). Parvin et al. (2023) used non-parametric regression techniques for electricity load forecasting, achieving probabilistic interval prediction objectives, and used electricity loads, carbon emissions, and so on as constraints, solving the objective function of the green power system through particle swarm optimisation algorithms. However, load curves are nonlinear, so traditional methods cannot perform the prediction tasks well. Machine learning methods excavate deep information from massive data, break through the limitations of existing physical knowledge, treat the power system as a ‘black box’ to fit the input-output relationship, and thus achieve forecasting. Wen et al. (2019) established an electricity load forecasting model based on decision trees and proposed an economic dispatch model that incorporates different policy combinations based on robust optimisation techniques, solving it through an improved GA algorithm. Uwimana et al. (2023) used support vector machines to forecast the load of the power .system and constructed an economic dispatch model to address the uncertainty in net load caused by the randomness of wind power by introducing a flexible generation reallocation mechanism.

Deep learning-based load forecasting methods automatically extract deep features through neural network models, thereby improving prediction efficiency. Traditional methods for feature extraction often require manual design and selection of features based on domain knowledge and experience. This process is not only time-consuming and labour-intensive but also relies on the subjective judgement of experts, potentially failing to comprehensively capture key information within the data. Deep learning methods can automatically learn and extract deep-level features from raw data without the need for manual feature design, thereby improving evaluation efficiency. Wen et al. (2019) selected electrical quantities, performed feature extraction of the power system through auto-encoding, and then input them into a convolutional neural network for load forecasting. By combining the forecasting results, a multi-objective optimisation function for power system scheduling was established, thereby increasing policy accommodation rates. Tang et al. (2022) used convolutional neural network (CNN) to extract characteristics from input variables and then employed a bidirectional gated recurrent unit (BiGRU) to further extract spatial information from power load data. The attention mechanism was used to adjust the weights of the feature data, thereby enhancing the model's generalisation ability and robustness. Zhou et al. (2023) used a recurrent neural network to forecast the load conditions in the power system and employed a reinforcement learning (RL) algorithm to obtain optimal policy combinations for power system scheduling. The model inputs were Euclidean structured data, without considering the impact of system topology. To address more complex non-Euclidean structured data, graph neural network (GNN) focus on the relationships between nodes, incorporating the topological structure information of graphs into the study. In the power system, generators and substation buses are considered as nodes, and transmission lines are considered as connections between nodes, which well align with the characteristics of graphs. Chen et al. (2023) used a graph convolutional network (GCN) to extract and learn the regional adjacency dependencies of power load data, while using a GRU to extract temporal features of the power load data, and modelled policies and power scheduling optimisation with an RL algorithm, thus reducing system operation costs. Solheim et al. (2024) used graph attention neural networks (GAT) for evaluation. Although temporal correlations of the data were considered, the system optimisation efficiency was not high since the adopted model lacked spatiotemporal characteristics. Chen et al. (2024) combined GCN with long short-term memory (LSTM) and used the Spearman rank correlation coefficient (Ali Abd, 2022) to analyse the correlation between load and meteorological factors, taking into account the extraction of spatial characteristics of power data, and solved the optimal policy combinations for power scheduling using a genetic algorithm, thereby improving system operational efficiency.

From the analysis of existing studies, the current methods have not adequately considered the impacts between users in different geographical areas, making it difficult to effectively extract the spatiotemporal correlations in power data. Therefore, this paper proposes a green power consumption policy simulation and evaluation method driven by spatiotemporal GCN. First, a power load forecasting model based on spatiotemporal GCN is designed. The graph structure is dynamically adjusted according to user power consumption patterns to more accurately reflect the spatial correlations between users in different geographical areas. Second, a weighted skip connection mechanism is designed. Connections at different time steps are assigned different weights based on the importance of different time scales in the time series, which helps alleviate the issues of gradient disappearance and gradient explosion in power load forecasting. On this basis, a mathematical model for the optimisation of power consumption policy combinations is established, with constraints such as power load and power generation output. The target function and constraints in the optimisation model are converted into a reward function in deep reinforcement learning (DRL). The environment, its state space, and action space are designed. The interaction with the environment is used to solve for the optimal power consumption policy combinations that minimise economic costs and carbon emissions. Experimental outcome indicates that the green power consumption rate of the suggested method is 97.16%, an improvement of at least 4.34% compared to baseline methods. This is beneficial for promoting a green and low-carbon energy transition and achieving sustainable growth in the power system and socio-economy.

2 Relevant technologies

2.1 Graph convolutional network

GCN is a deep learning model commonly used for processing irregularly structured data. This type of data is typically non-Euclidean space data, where nodes can be connected in any way through edges, forming complex spatial relationships. Through graph convolution operations, GCN can effectively process the relationships between each node and its neighbours, thereby better extracting the structural features of non-regular spatial data (Wang et al., 2021). GCN and GAT both belong to GNN models. For large-scale graph data, such as power grids, or tasks with high real-time requirements, the lightweight nature of GCN makes its training and inference speeds significantly better than those of GAT. GCN implements graph convolution through eigen decomposition of the Laplacian matrix or Chebyshev polynomial approximation. Its computational complexity mainly depends on matrix multiplication, and it does not require additional parameters for learning adjacency weights. GAT, on the other hand, must calculate attention coefficients for each edge, resulting in computational complexity that increases quadratically with the number of nodes and requiring maintenance of an attention weight matrix, thus occupying more memory.

GCN uses the graph's topology to influence feature propagation. In traditional CNNs, convolution operations are performed only within fixed-size local neighbourhoods, whereas the number of neighbours of nodes in graph data may vary (Bhatti et al., 2023). Therefore, GCNs first normalise the adjacency matrix and then perform convolution operations to extract the connection relationships between nodes and their neighbours, as shown in equation (1).

$$H' = \delta\left(\hat{D}^{-\frac{1}{2}}\hat{A}\hat{D}^{\frac{1}{2}}HW\right) \tag{1}$$

where H is the input feature matrix, H' is the output feature matrix, $\hat{A}$ is the symmetrically normalised adjacency matrix, $\hat{D}$ is the diagonal matrix of the degree matrix, W is the weight matrix, and $\delta(\cdot)$is the activation function.

2.2 Deep reinforcement learning algorithm

RL aims to learn an optimal decision-making policy through interaction with the environment to maximise future cumulative rewards (Zhai et al., 2023). Traditional RL uses tables of value functions to store action-value information but shows significant limitations when dealing with complex Markov decision processes (MDP) (Ororbia and

Warn, 2022). DRL approximates value functions and optimises the decision-making process using deep neural networks, aiming to significantly enhance task-processing capabilities in complex environments. Traditional RL algorithms typically employ lookup tables or simple linear function approximators to store and compute state-action value functions. For linear function approximators, manually designed features are required to represent states, and the quality of these features significantly impacts algorithm performance. Moreover, linear function approximators can only capture linear relationships between state features, making it difficult to handle complex nonlinear relationships. DRL algorithms utilise deep neural networks as function approximators. Deep neural networks possess powerful automatic feature extraction capabilities, enabling them to autonomously learn meaningful feature representations from raw high-dimensional data. Their nonlinear activation functions allow approximation of complex nonlinear functions, thereby better representing state-action value functions. This enables deep reinforcement learning algorithms to handle more complex environments and tasks, performing exceptionally well on problems with large or continuous state and action spaces.

The training process of DRL follows a state-action interaction mechanism. At each time step, the agent selects an action a based on the current observed state s, forming the mapping relationship from state to action. This transition mechanism constitutes the policy π, as shown in equation (2), with the corresponding probability distribution as in equation (3). Here, A_t is the full action sequence at time t, S_t is the full state sequence at time t, and ρ is the probability distribution function.

$$a = \pi(s) \tag{2}$$

$$\pi(a|s) = \rho\left[A_t = a \middle| S_t = s\right] \tag{3}$$

Throughout the process, the agent combines its previous exploration experience and keeps trying actions to explore, with the ultimate goal of finding an optimal policy to obtain as much cumulative reward as possible. The computation equation is as follows:

$$G_t = R_{t+1} + \lambda R_{t+2} + \ldots = \sum_{k=0}^{\infty} \lambda^k R_{t+k+1} \tag{4}$$

where G_t is the cumulative reward over the entire training episode, R_t is the reward obtained at state t, λ is the discount factor, a positive number less than 1, reflecting the decaying effect of long-term rewards.

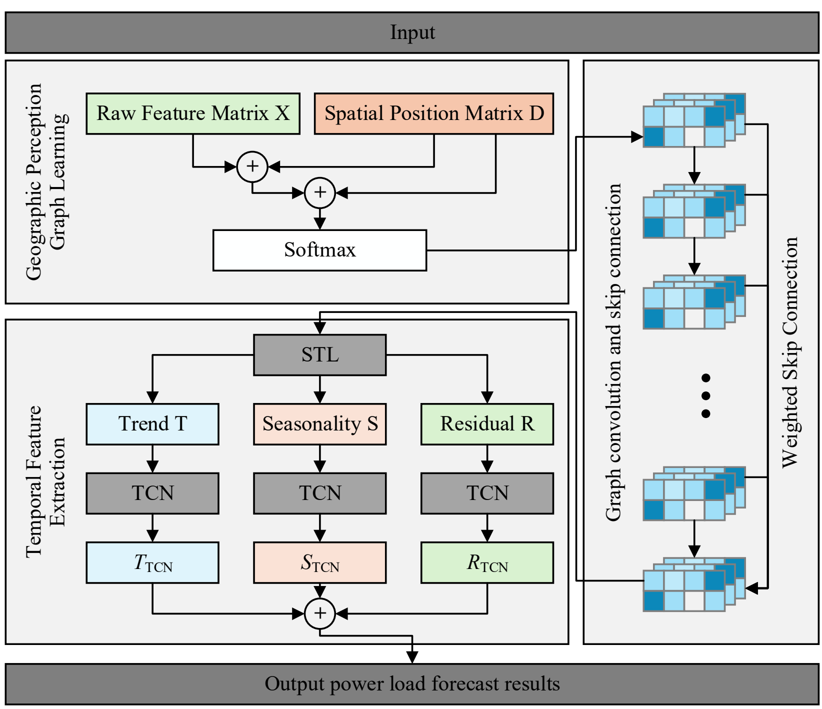
The theoretical core of DRL lies in constructing a nonlinear mapping relationship between high-dimensional state spaces and action policies through deep neural networks, driving the autonomous evolution of intelligent policies (Demirel et al., 2018). Its essence is to build a closed-loop learning system through dynamic interaction between the agent and the environment: by coordinating the optimisation mechanism of value function approximation and policy gradients, it improves the accuracy of value estimation through policy iteration, while using feedback from the value network to correct the direction of policy updates, ultimately converging to a globally optimal decision-making paradigm.

3 Power load prediction based on spatio-temporal graph convolutional network

3.1 Model overview

It can be inferred from the above analysis that power load forecasting is the basis for the optimal scheduling of a combined green power consumption policy. However, existing forecasting models have failed to fully consider the interactions between users in different geographic regions, making it difficult to effectively extract the spatial correlations in power data. In addition, existing models frequently encounter gradient vanishing or gradient exploding problems over longer time spans, thereby reducing the sensitivity of the model to features at different time scales. Therefore, this article puts forward a power load forecasting model in light of spatiotemporal GCNs to improve prediction accuracy for power load forecasting. The model structure is shown in Figure 1. The model constructs an adjacency matrix to represent the interdependency relationships between user nodes in different geographic regions, which helps to enhance the model's accuracy and generalisation ability. At the same time, through a weighted skip connection mechanism, the model can effectively transmit gradients over a longer time span, thus alleviating the gradient vanishing or gradient exploding problems.

Figure 1 The model structure of the suggested power load forecasting model (see online version for colours)



1 Geospatial-aware graph learning module: this module learns the graph structure in a data-driven manner while increasing the weight of geographic region information during the learning process, using geographic region information and electricity consumption patterns to learn the graph's adjacency matrix. The module first initialises an original feature matrix used to represent the latent features of different users. Then, using the spatial location matrix and electricity consumption pattern data, it transforms into a weight matrix via a Gaussian kernel function to learn an adaptive adjacency matrix that reflects the interdependency relationships between user nodes in different geographic regions. By introducing learnable parameters, the adjacency matrix can adaptively adjust weights to more accurately reflect the relationships between user nodes.

2 Graph convolution and skip connection module: this module aggregates information from each user node and its neighbouring nodes through graph convolution operations, using the learned adjacency matrix to guide the flow of feature information. At the same time, a weighted skip connection mechanism is introduced, enabling information to traverse multiple layers in the network to capture long-range dependencies, thereby enhancing the model's ability to integrate spatial and temporal information.

3 Temporal feature extraction module: this module is adopted to capture the time series characteristics of electricity data. Using time series decomposition methods, the time series is decomposed into trend, seasonal, and residual components to help the model better understand and interpret the components of the time series. The sliding window technique is utilised to divide the time series data into multiple time windows, followed by the analysis of features at different time scales.

3.2 Geospatial perception map learning

First, initialise an original characteristic matrix X, where N is the number of nodes and F is the number of features per node. Simultaneously, establish a spatial location matrix D, used to represent the positional relationship between user nodes in different geographic areas, where D_{ij} is the geographic distance between node i and node $_j$, and the spatial location matrix is transformed into a weight matrix using a Gaussian kernel function. An adjacency matrix is learned through the original feature matrix and spatial location matrix to represent the relationship between user nodes in different geographic areas. At the same time, by introducing learnable parameters, the adjacency matrix can adaptively adjust weights to more accurately reflect the relationship between user nodes, as shown in equations (5)–(8).

$$W_{ij} = \exp\left(-\frac{D_{ij}^2}{2\sigma^2}\right) \tag{5}$$

$$Z = [X, \alpha W] \tag{6}$$

$$A' = softmax\left(ZZ^{\mathrm{T}} + \beta W\right) \tag{7}$$

$$softmax(z_i) = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}} \tag{8}$$

where W is the spatial position weight matrix between user nodes in different geographic areas, W_{ij} is the spatial position weight between node i and node j, σ is the hyperparameter controlling the degree of distance influence, Z is the node feature matrix obtained by combining X and W, used to initially integrate geographic spatial features, α and β are hyperparameters, A' is the adjacency matrix obtained through graph learning, and *softmax* is the linear activation function.

3.3 Graph convolution and skip connections

To effectively alleviate the gradient vanishing or gradient exploding problem during the propagation process, this paper designs weighted skip connections for information propagation, combining the output of each layer with the features of all previous layers in a weighted manner, thereby effectively alleviating the gradient vanishing problem during the propagation process, as shown in equation (9).

$$H_{(l+1)} = \sigma\left(\sum_{k=0}^{K} \tilde{A}^k H_{(l)} W_{(l,k)} + \sum_{j=0}^{l} \alpha_{(l,j)} H_{(j)}\right) \tag{9}$$

where K is the propagation depth, $H_{(l+1)}$ is the characteristic matrix of the $(l + 1)^{\text{th}}$ layer node, $\tilde{A}$ is the normalised adjacency matrix, $W_{(l,k)}$ is the trainable weight matrix of the l^{th} layer and k^{th} order, $H_{(j)}$ is the feature matrix of the j^{th} layer nodes, $\alpha_{(l,j)}$ is the trainable weight parameter of the l^{th} layer for the j^{th} layer feature, and σ is the activation function.

3.4 Time series feature extraction and prediction result output

First, sliding window technology is used to divide the time series data into multiple windows. Assuming the window size is w and the step size is S. For each window, the output of the graph convolution $H_{(l+1)}$ is a three-dimensional tensor, where w is the window size and F is the number of features. Subsequently, time series decomposition is performed on the graph convolution output of each window to obtain the trend component, seasonal component, and residual component, as shown in equation (10), where $S_{i(l+1)}$ is the trend component of the i^{th} window, $S_{i(l+1)}$ is the seasonal component of the i^{th} window, and $R_{i(l+1)}$ is the residual component of the i^{th} window.

$$H_{i(l+1)} = T_{i(l+1)} + S_{i(l+1)} + R_{i(l+1)} \tag{10}$$

For each decomposed component, a temporal convolutional network (TCN) (Wan et al., 2019) is used to extract temporal features, as shown in equations (11)–(13).

$$T_{TCN(l+1)} = TCN\left(T_{(l+1)}\right) \tag{11}$$

$$S_{TCN(l+1)} = TCN\left(S_{(l+1)}\right) \tag{12}$$

$$R_{TCN(l+1)} = TCN\left(R_{(l+1)}\right) \tag{13}$$

The trend component, seasonal component, and residual component after TCN processing are weighted and merged to gain the final feature matrix, as shown in equation (14), where γ_{trend}, $\gamma_{seasonal}$ and $\gamma_{residual}$ are trainable weight parameters used to adjust the impact weights of the trend component, seasonal component, and residual component.

$$H_{(l+2)} = \gamma_{trend} T_{TCN_{(l+1)}} + \gamma_{seasonal} S_{TCN_{(l+1)}} + \gamma_{residual} R_{TCN_{(l+1)}} \tag{14}$$

Finally, the temporal features containing spatial features are input into a fully connected network to obtain the final load forecasting results.

4 Optimisation of green power consumption policies based on deep reinforcement learning algorithms

4.1 Optimisation model for green power consumption policy combination

Electric load forecasting is a crucial foundation for the operation and planning of power systems. Accurate load forecasting can offer strong support for power scheduling and resource allocation, thereby influencing the consumption of green power. Green power consumption policies consist of a series of policy measures formulated to promote the effective utilisation of green power, including feed-in tariff policies, subsidy policies, quota-based policies, and green certificate trading policies.

The green power consumption policy portfolio optimisation problem is generally modelled as a non-convex optimisation problem with highly nonlinear constraints. The typical optimisation objectives are lowest generation cost and lowest carbon emissions. This article's optimisation objective is to determine the green power consumption policy portfolio that minimises the total generation cost and carbon emissions. Assuming that based on the electric load forecasting results, the grid G is equipped with N buses, M branches, and K generating units, its optimisation objective function is shown in equation (15).

$$\min \alpha \left(\sum_{k=1}^{K} \left(a_k \left(P_k^G \right)^2 + b_k P_k^G + c_k \right) + \beta \sum_{j=1}^{J} k_j \left(\overline{P}_{j,t} - P_{j,t} \right) \right) \tag{15}$$

where P_k^G is the output of generator k, a_k, b_k and c_k are the operating cost coefficients of this generating unit, $\overline{P}_{j,t}$ is the maximum generation capacity prediction value, $P_{j,t}$ is the actual consumption power, and k_j is the cost corresponding to unit carbon emissions.

The constraints for green power consumption policy portfolio optimisation can be divided into four categories: load constraints, nodal voltage constraints, line transmission capacity constraints, and generator output constraints. Load constraints mainly refer to load spatiotemporal characteristics constraints, that is, the objective function should satisfy $y_t \le \hat{y}_t$, where y_t is the current electric load, and $\hat{y}_t$ is the forecasted electric load.

Nodal voltage constraints are shown in equation (16), ensuring the node voltage stays within the voltage control standard range, where V_i is the voltage magnitude of node i, and $V_i^{\min}$ and $V_i^{\max}$ are its allowable upper and lower voltage magnitude limits, respectively.

$$V_i^{\min} \le V_i \le V_i^{\max} \tag{16}$$

Line transmission power constraints are shown in equation (17), ensuring the power on the line remains within the rated power, where P_j^B is the active power transmitted on line *j*, and $P_j^{B,\max}$ is the line rating power.

$$P_j^B \le P_j^{B,\max} \tag{17}$$

Generator output constraints ensure the active and reactive power output of the generator remains within allowable ranges. As shown in equation (18), it is divided into active power output upper and lower bounds constraints and reactive power output upper and lower bounds constraints, where P_k^G and Q_k^G are the active and reactive power output of generator *k*, $P_k^{G,\min}$ and $Q_k^{G,\min}$ are the minimum active and reactive power output of generator *k*, and $P_k^{G,\max}$ and $Q_k^{G,\max}$ are its maximum active and reactive power outputs, respectively.

$$\begin{cases} P_k^{G,\min} \le P_k^G \le P_k^{G,\max} \\ Q_k^{G,\min} \le Q_k^G \le Q_k^{G,\max} \end{cases} \tag{18}$$

4.2 Model solving based on deep reinforcement learning algorithms

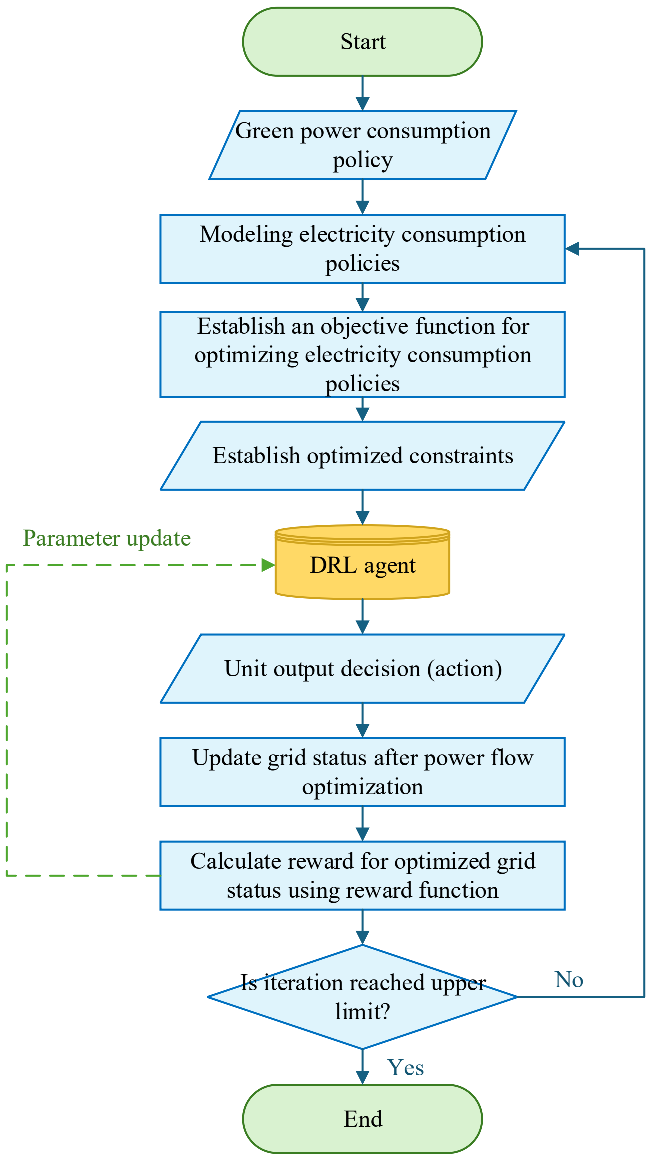
After establishing the mathematical model, this paper solves for the optimal combination of green power consumption policies by interacting with the environment using DRL algorithms. Since this paper defines the optimal combination of green power consumption policies as a single-step MDP, there is only one step of calculation per episode in the DRL training. Figure 2 is the flowchart of the simulation-based offline training environment for the optimisation problem proposed in this paper, divided mainly into three parts.

1 Initialisation of random optimisation case: randomising the topology structure of a source scenario using the $N-1$ method, i.e., randomly removing one branch. Based on this, taking the active and reactive power loads and line impedance in the source scenario as benchmarks, allowing them to vary randomly within a certain range to generate the policy combination optimisation case with uncertain topology structure and load, i.e., generate random state s^0.

2 Agent interacting with the environment: extract the node feature matrix *H*, edge feature matrix *E*, and node adjacency matrix *A* from the generated green power consumption policy combination optimisation cases, input them to the DRL agent to map to the action output a, and transform it into power generation unit output decision instructions. Update the green power consumption policy combination based on the decision instructions, i.e., update the state to s^1.

3 Calculate rewards and update the agent: calculate the updated optimisation policy. If the calculation diverges, return a large negative value as the non-convergence penalty r^D. If the power flow converges, calculate the degree of violation of all system indicators under convergence, and provide the reward under convergence based on the reward function r^C.

Figure 2 Flowchart of the proposed optimisation problem simulation offline training environment (see online version for colours)



DRL evaluates the performance of the model through a reward function and uses it as a basis for optimising parameters in the agent. However, unlike the constraints in optimisation problems that are typically enforced rigidly, the reward function in DRL is generally used to impose constraints on the strategies proposed by the agent either by imposing high penalties or by transforming some constraints into rules within the simulation environment. The total reward calculation is shown in equation (19).

$$r = S \times r^C + (1 - S) \times r^D \tag{19}$$

In this binary, variable S represents whether the model converges. If it converges, $S = 1$; otherwise, $S = 0$. r is the total reward returned by the environment, r^D is the reward in the non-convergent case, where this paper directly provides a large negative value as a penalty, i.e., a negative reward, r^C is the reward in the model convergence case, and the specific calculation method is as follows.

For the objective function of the green power consumption policy combination optimisation, the reward function is shown in equation (20), where R^C is the total reward for generator power generation cost, which is scaled to be between (–1, 0), P_k^G is the power output of generator k, $P_k^{G,\max}$ is its active power output upper limit, a_k, b_k and c_k are its operation cost coefficients.

$$R^C = -\frac{\sum_{k=1}^{K} a_k \left(P_k^G\right)^2 + b_k P_k^G + c_k}{\sum_{k=1}^{K} a_k \left(P_k^{G,\max}\right)^2 + b_k P_k^{G,\max} + c_k} \tag{20}$$

Regarding the node voltage constraints, the voltage at each node is also produced by the power flow calculation, thus there is also a risk of voltage magnitude exceeding the limit. Therefore, this paper converts the node voltage magnitude upper and lower limit constraints into the reward functions shown in equations (21) and (22).

$$r_i^V = \begin{cases} V_i - V_i^{\max}, & V_i > V_i^{\max} \\ 0, & V_i^{\min} \le V_i \le V_i^{\max} \\ V_i^{\min} - V_i, & V_i \le V_i^{\min} \end{cases} \tag{21}$$

$$R^V = \exp\left(-\frac{\sum_{i=1}^{N} r_i^V}{N}\right) - 1 \tag{22}$$

where R^V is the total reward for node voltage violation, which is scaled to be between (–1, 0), r_i^V is the voltage violation constraint reward for node i, N is the total amount of nodes in the power flow optimisation case, V_i is the voltage magnitude of node i, and $V_i^{\max}$ and $V_i^{\min}$ are its voltage magnitude upper and lower bounds, respectively.

Regarding the line power transmission constraints, the power flow of each branch is assigned by the power flow calculation, so there is also a risk of line power transmission exceeding the limit. Therefore, this paper converts the line transmission capacity constraint in equation (17) into the reward functions shown in equations (23) and (24). Among them, R^B is the total reward for line power transmission limit violation, which is scaled to be between (–1, 0); r_j^B is the line power constraint reward obtained by line j,

P_j^B is the active power transmitted on line j, $P_j^{B,\max}$ is its rated transmission power, and M is the number of policy combinations in the power flow optimisation method.

$$r_j^B = \begin{cases} P_j^B - P_j^{B,\max}, & P_j^B > P_j^{B,\max} \\ 0, & P_j^B \le P_j^{B,\max} \end{cases} \tag{23}$$

$$R^B = \exp\left(-\frac{\sum_{j=1}^{M} r_j^B}{M} \right) - 1 \tag{24}$$

Regarding the generator power output constraints, since this paper calculates the non-swing generator power outputs through equation (18), it can be considered as forcibly satisfying its active power upper and lower bound constraints. However, the swing generator output is determined by the power flow calculation, and its value depends on the power imbalance in the system, so there is a risk of boundary violation. Therefore, the active power upper and lower bounds constraint in equation (18) is converted into the reward function shown in equation (25).

$$R^P = \begin{cases} \exp\left(P^{G_B,\max} - P^{G_B}\right) - 1, & P^{G_B} > P^{G_B,\max} \\ 0, & P^{G_B,\min} < P^{G_B,\max} \\ \exp\left(P^{G_B} - P^{G_B,\min}\right) - 1, & P^{G_B} < P^{G_B,\min} \end{cases} \tag{25}$$

where R^P is the reward for the swing generator power output limit violation, which is scaled to be between (–1, 0), P^{G_B} is the swing generator output, and $P^{G_B,\max}$ and $P^{G_B,\min}$ are its output upper and lower bounds, respectively.

After calculating each reward, the total reward under the power convergence is calculated through equation (26).

$$r^C = \rho^V R^V + \rho^B R^B + \rho^P R^P + \rho^Q R^Q + \rho^C R^C \tag{26}$$

In this context, ρ^V, ρ^B, ρ^P, ρ^Q and ρ^C are the penalty coefficients for the node voltage violation reward, the line transmission capacity violation reward, the balance unit output violation reward, the unit output violation reward, and the generation cost reward, respectively.

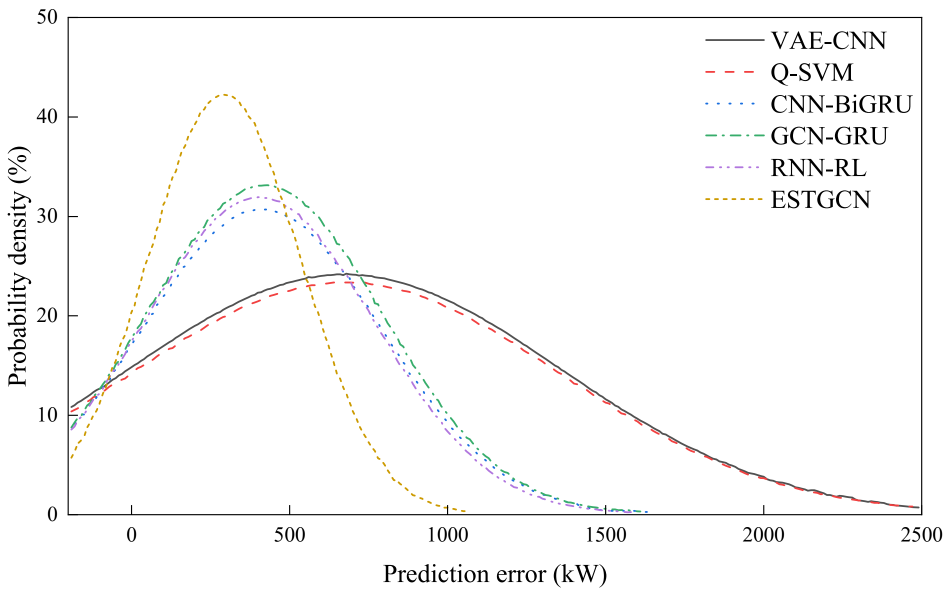
Finally, the parameters in the DRL agent are updated according to the reward to optimise its decision-making strategy. Through interaction with the environment, the optimal green electricity consumption policy combination is iteratively solved to achieve economic and low-carbon operation of the power system.

5 Experiment and performance analysis

5.1 Power load forecasting accuracy analysis

The experiment selected 2,158 pieces of power system operation data and 84 pieces of policy and market mechanism data collected from reference (Wang and Li, 2022) as the dataset. The power system operation data included the installed capacity, output curves, ramp rates, and start-up and shutdown costs of thermal power, wind power, photovoltaic, and other power generation units. Policy and market mechanism data include renewable energy consumption ratios, grid-connected electricity prices for photovoltaic/wind power, energy storage subsidies, etc. In the experiment, the GCN model uses the ReLU activation function, a learning rate of 0.03, and the Adam optimiser. The system is equipped with a GPU server featuring a TITAN Xp graphics card and 32 GB of memory. The operating system is Ubuntu 18.04, the deep learning framework is PyTorch, and the programming language is Python 3.8.

Figure 3 Distribution of prediction errors for different models of power load (see online version for colours)



This paper selects Q-SVM (Uwimana et al., 2023), VAE-CNN (Wen et al., 2019), CNN-BiGRU (Tang et al., 2022), RNN-RL (Zhou et al., 2023), and GCN-GRU (Chen et al., 2024) as comparative models, and the proposed model is denoted as ESTGCN. An analysis of the prediction error distribution of different models for power load is conducted, as shown in Figure 3. The predicted errors of the ESTGCN model are all within a lower error range, with significantly more concentrated distribution, which indicates that the ESTGCN prediction results are closer to the actual condition and have less fluctuation in prediction results, with better predictive stability and reliability. From the above statistical data, it can also be found that the ESTGCN improves prediction accuracy by introducing an adaptive graph construction method and a weighted skip connection mechanism, which can more accurately extract the correlation of electricity consumption among users in different geographic regions. Meanwhile, it can effectively improve the prediction accuracy for power load. In contrast, CNN-BiGRU, RNN-RL, and GCN-GRU mainly focus on capturing the temporal association of data and relatively lack the ability to extract energy consumption patterns and coupling characteristics of the system. Additionally, Q-SVM and VAE-CNN not only neglect the spatial correlation of power load characteristics but also ignore temporal features, leading to larger prediction errors compared to the other four models.

5.2 Performance analysis of green power consumption policy combinations

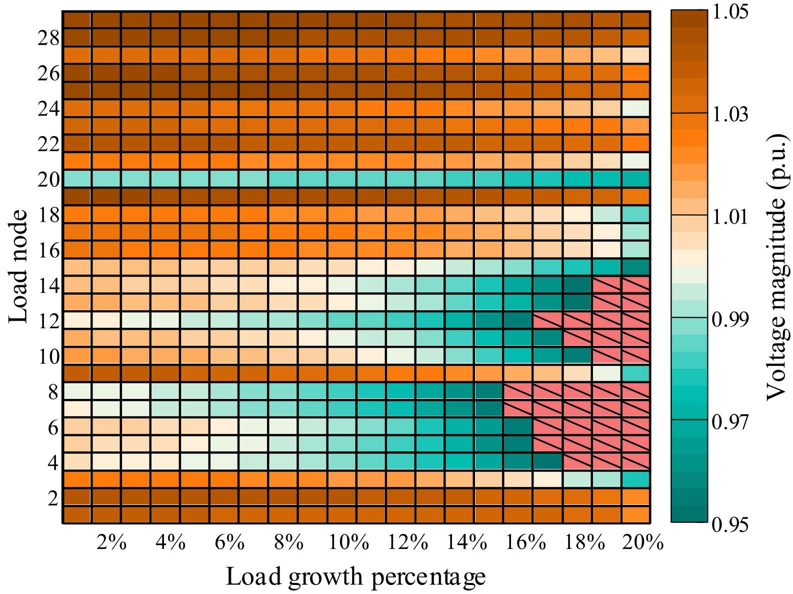
To further verify the performance of the green power consumption policy combination optimisation based on the DRL algorithm, this paper first analyses the optimisation results of DRL. In a single round without any regulating actions, and with active power load increasing by 1% per step, the system's supply shortage will lead to voltage drops. As shown in the heatmap of Figure 4, all load nodes begin to experience voltage violations below the lower limit (0.95 p.u.) when active power increases by 16%, and significant voltage violations and rapid system collapse occur around an 18% increase. Through repeated experiments, nodes with small voltage fluctuations are ignored, and ten easily over-limit load nodes bus-4, bus-5, bus-6, bus-7, bus-8, bus-10, bus-11, bus-12, bus-13 and bus-14 are selected as voltage amplitude observation points. Compared to conventional optimisation algorithms, the number of voltage-violation nodes is reduced from 10 to zero, the voltage violation rate is decreased by 34.5%, and voltage collapse risks are avoided, demonstrating the effectiveness of the optimisation strategy.

This paper selects the green power consumption rate (GPCR), green power penetration rate (GPPR), energy consumption growth rate (ECGR), and computation time T to conduct a comparative analysis of the performance of different models in combination optimisation of green power consumption policies, as shown in Table 1. ESTGCN achieves a GPCR and GPPR of 97.16% and 85.37%, respectively, which represent improvements of 18.52% and 15.18% over Q-SVM, 15.57% and 9.73% over VAE-CNN, 10.25% and 6.45% over CNN-BiGRU, 8.53% and 4.83% over RNN-RL, and 4.34% and 1.76% over GCN-GRU. Further comparison of the ECGR indicator shows that ESTGCN improves by 2.24–13.14% over the other five models. In terms of computation time for the optimal combination of green power consumption policies, ESTGCN takes 1.24 s, which is at least 3.72 s less than the baseline methods. Through analysis of the above performance indicators, it is evident that ESTGCN significantly accelerates the solution speed of the optimisation problem while ensuring optimisation effectiveness.

Table 1 Optimisation of green power consumption policy combinations

Indicator	*GPCR (%)*	*GPPR (%)*	*ECGR (%)*	*T (s)*
Q-SVM	78.64	70.19	3.92	13.69
VAE-CNN	81.59	75.64	7.81	10.92
CNN-BiGRU	86.91	78.92	10.69	8.03
RNN-RL	88.63	80.54	13.51	5.81
GCN-GRU	92.82	83.61	14.82	4.96
ESTGCN	97.16	85.37	17.06	1.24

Figure 4 DRL optimisation solves heat maps (see online version for colours)



6 Conclusions

With the acceleration of global energy transition and the advancement of dual carbon goals, green electricity consumption has become a key issue for the sustainable development of energy systems, and its efficiency improvement is highly dependent on scientific policy guidance. To address the problem that current research struggles to effectively extract spatiotemporal correlations from power data, this paper proposes a spatiotemporal graph convolutional network-driven simulation and evaluation method for green electricity consumption policies. First, a power load forecasting model based on spatiotemporal graph neural networks is designed, which dynamically adjusts the graph structure according to users' electricity consumption patterns to more accurately reflect the spatial correlations of users in different geographic areas. Second, a weighted skip connection mechanism is designed, which can assign different weights to connections at different time steps according to the importance of different time scales in the time series, to alleviate the problem of gradient explosion in prediction. On this basis, a mathematical model for the combination optimisation of power consumption policies is established, with power load, power generation unit outputs, etc., as constraints, and by interacting with the environment through deep reinforcement learning algorithms to solve for the optimal combination of power consumption policies that minimise both economic and carbon emission costs. Experimental outcome implies that the suggested approach achieves a green electricity consumption rate of 97.16%, with an optimisation computation time of 1.24 s, significantly accelerating the solution speed while ensuring the optimisation effect of the green electricity consumption policy combination.

Declarations

All authors declare that they have no conflicts of interest.

References

Ali Abd, K. (2022) 'Spearman's correlation coefficient in statistical analysis', *International Journal of Nonlinear Analysis and Applications*, Vol. 13, No. 1, pp.3249–3255.

Bhatti, U.A., Tang, H., Wu, G., Marjan, S. and Hussain, A. (2023) 'Deep learning with graph convolutional networks: an overview and latest applications in computational intelligence', *International Journal of Intelligent Systems*, Vol. 10, No. 1, pp.83–94.

Chen, G., Hui, W., Huan, Y., Bingchen, L. and Xingxing, Z. (2024) 'Research on optimization of distribution network connection mode based on graph neural network and genetic algorithm', *Distributed Generation & Alternative Energy Journal*, Vol. 39, No. 6, pp.1179–1208.

Chen, J., Yu, T., Pan, Z., Zhang, M. and Deng, B. (2023) 'A scalable graph reinforcement learning algorithm based stochastic dynamic dispatch of power system under high penetration of renewable energy', *International Journal of Electrical Power & Energy Systems*, Vol. 152, pp.32–45.

Chodakowska, E., Nazarko, J. and Nazarko, Ł. (2021) 'ARIMA models in electrical load forecasting and their robustness to noise', *Energies*, Vol. 14, No. 23, pp.52–64.

Constantinou, S., Konstantinidis, A. and Zeinalipour-Yazti, D. (2022) 'Green planning systems for self-consumption of renewable energy', *IEEE Internet Computing*, Vol. 27, No. 1, pp.34–42.

Demirel, B., Ramaswamy, A., Quevedo, D.E. and Karl, H. (2018) 'DeepCAS: a deep reinforcement learning algorithm for control-aware scheduling', *IEEE Control Systems Letters*, Vol. 2, No. 4, pp.737–742.

Khalil, M. and Sheikh, S.A. (2024) 'Advancing green energy integration in power systems for enhanced sustainability: a review', *IEEE Access*, Vol. 12, pp.2169–3536.

Lu, Y. and Park, S.D. (2022) 'Time series analysis of policy discourse on green consumption in China: text mining and network analysis', *Sustainability*, Vol. 14, No. 22, pp.14–27.

Nepal, B., Yamaha, M., Yokoe, A. and Yamaji, T. (2020) 'Electricity load forecasting using clustering and ARIMA model for energy management in buildings', *Japan Architectural Review*, Vol. 3, No. 1, pp.62–76.

Ororbia, M.E. and Warn, G.P. (2022) 'Design synthesis through a Markov decision process and reinforcement learning framework', *Journal of Computing and Information Science in Engineering*, Vol. 22, No. 2, pp.21–29.

Parvin, M., Yousefi, H. and Noorollahi, Y. (2023) 'Techno-economic optimization of a renewable micro grid using multi-objective particle swarm optimization algorithm', *Energy Conversion and Management*, Vol. 277, pp.39–54.

Press, M. and Arnould, E.J. (2009) 'Constraints on sustainable energy consumption: market system and public policy challenges and opportunities', *Journal of Public Policy & Marketing*, Vol. 28, No. 1, pp.102–113.

Rashid, S.M. (2024) 'Employing advanced control, energy storage, and renewable technologies to enhance power system stability', *Energy Reports*, Vol. 11, pp.3202–3223.

Solheim, Ø.R., Høverstad, B.A. and Korpås, M. (2024) 'Using graph neural networks in reinforcement learning with application to Monte Carlo simulations in power system reliability analysis', *IEEE Access*, Vol. 12, pp.160175–160189.

Tang, H., Lv, K., Bak-Jensen, B., Pillai, J.R. and Wang, Z. (2022) 'Deep neural network-based hierarchical learning method for dispatch control of multi-regional power grid', *Neural Computing and Applications*, Vol. 34, No. 7, pp.5063–5079.

Uwimana, E., Zhou, Y. and Zhang, M. (2023) 'Long-term electrical load forecasting in Rwanda based on support vector machine enhanced with Q-SVM optimization kernel function', *Journal of Power and Energy Engineering*, Vol. 11, No. 8, pp.32–54.

Vakili, G., Mishra, S.B., Mousa, H. and Ziaiefar, H. (2021) 'Simulation and optimization of hybrid green energy systems for desalination purposes', *Environmental Progress & Sustainable Energy*, Vol. 40, No. 2, pp.15–23.

Wan, R., Mei, S., Wang, J., Liu, M. and Yang, F. (2019) 'Multivariate temporal convolutional network: a deep neural networks approach for multivariate time series forecasting', *Electronics*, Vol. 8, No. 8, pp.87–96.

Wang, L. and Li, K. (2022) 'Research on renewable energy consumption and emission reduction in power market based on bi-level decision making in China', *Energy*, Vol. 260, pp.12–26.

Wang, Z., Xia, M., Lu, M., Pan, L. and Liu, J. (2021) 'Parameter identification in power transmission systems based on graph convolution network', *IEEE Transactions on Power Delivery*, Vol. 37, No. 4, pp.3155–3163.

Wen, L., Zhou, K., Yang, S. and Lu, X. (2019) 'Optimal load dispatch of community microgrid with deep learning based solar power and load forecasting', *Energy*, Vol. 171, pp.1053–1065.

Zhai, S., Li, W., Qiu, Z., Zhang, X. and Hou, S. (2023) 'An improved deep reinforcement learning method for dispatch optimization strategy of modern power systems', *Entropy*, Vol. 25, No. 3, pp.54–66.

Zhou, X., Wang, J., Wang, X. and Chen, S. (2023) 'Optimal dispatch of integrated energy system based on deep reinforcement learning', *Energy Reports*, Vol. 9, pp.373–378.