IIETA International Information and Engineering Technology Association
Advancing the World of Information and Engineering

International Journal of Heat and Technology
Vol. 43, No. 2, April, 2025, pp. 502-512
Journal homepage: http://iieta.org/journals/ijht

Spatiotemporal Graph Neural Network-Based Optimization Strategies for Thermal Consumption in Dynamic Heat Conduction Networks within Smart Manufacturing

Lu Gao[1], Lei Bai[2*]

[1] College of Physical Science and Technology, Bohai University, Jinzhou 121000, China
[2] College of Computer Science and Technology, North China Institute of Science & Technology, Langfang 065201, China

Corresponding Author Email: leib_cn@126.com

https://doi.org/10.18280/ijht.430212

Received: 3 September 2024
Revised: 20 February 2025
Accepted: 9 March 2025
Available online: 30 April 2025

Keywords:
smart manufacturing, dynamic heat conduction network, spatiotemporal graph neural network, thermal consumption prediction, thermal optimization

ABSTRACT

Amid the global shift of smart manufacturing towards greener and more intelligent paradigms, the spatiotemporal coupling characteristics of dynamic heat conduction networks pose significant challenges for optimizing thermal consumption. Traditional mathematical models often lack adaptability to complex operational conditions, while conventional machine learning methods struggle to capture deep spatiotemporal dependencies in data. Moreover, static graph neural networks fail to account for the evolving relationships among nodes in heat conduction systems, limiting the accuracy of current prediction and optimization approaches. To address these issues, this study proposes a thermal consumption prediction algorithm based on spatiotemporal graph neural networks tailored for smart manufacturing. The proposed model integrates spatial graph structures with temporal sequence features: it employs graph convolution operations to capture the dynamic evolution of spatial heat conduction relationships among nodes, and leverages temporal analysis techniques to model time-dependent patterns. This unified approach enables comprehensive modeling of spatiotemporal dependencies in dynamic heat conduction networks. The study details the model architecture, spatial and temporal dependency modeling, and prediction methodology, and validates its effectiveness using real-world production data. The findings offer a precise prediction framework and strategic support for optimizing thermal energy use in smart manufacturing, overcoming the limitations of traditional methods in capturing dynamic spatiotemporal characteristics. This contributes both theoretical insight and practical value to enhancing energy efficiency and advancing green manufacturing initiatives.

1. INTRODUCTION

Under the global trend of intelligent and green transformation of the manufacturing industry, smart manufacturing, as the core direction for industrial upgrading [1, 2], faces the dual challenge of improving production efficiency and reducing energy consumption. The heat conduction process widely exists in many aspects of smart manufacturing [3-5], such as processing technology, equipment operation, and material handling. Its dynamic characteristics directly affect the stability of the production process, product quality, and energy utilization efficiency. Dynamic heat conduction networks have complex spatiotemporal coupling characteristics [6], and the heat conduction relationships between nodes dynamically evolve with time and production conditions [7, 8]. Traditional thermal consumption management methods are difficult to accurately capture this dynamic nature, resulting in poor optimization performance. With the development of technologies such as the Industrial Internet and the Internet of Things [9], large amounts of multi-source heterogeneous data related to heat conduction have been accumulated in smart manufacturing systems. How to effectively utilize these data to build accurate dynamic heat conduction network models has become a key issue in achieving thermal consumption optimization.

Modeling dynamic heat conduction networks and exploring thermal consumption optimization strategies have important theoretical and practical value in the field of smart manufacturing. Dynamic heat conduction networks involve complex spatiotemporal dependencies [10]. Integrating theories and methods such as graph theory and neural networks to model them [11, 12] can enrich and expand the theoretical system of complex system modeling in smart manufacturing and provide new ideas and methods for solving similar spatiotemporal dynamic system problems. Moreover, accurate thermal consumption optimization strategies can effectively reduce energy consumption in smart manufacturing processes, lower production costs, and improve the economic benefits of enterprises [13]. At the same time, they help improve the stability of production processes and product quality, enhancing the market competitiveness of enterprises. In addition, this aligns with the global concept of green manufacturing and sustainable development and plays a positive role in promoting the realization of the "dual carbon" goals.

At present, for the research on heat conduction networks in smart manufacturing, scholars have adopted a variety of methods. For example, some studies are based on traditional mathematical models, such as partial differential equations, to model the heat conduction process. However, such methods have poor adaptability when facing complex and variable production conditions and are difficult to accurately describe the dynamic characteristics of heat conduction [14, 15]. In data-driven methods, some studies use machine learning algorithms, such as support vector machines and random forests, for thermal consumption prediction. However, these methods are insufficient in mining the complex spatiotemporal dependencies hidden in the data, and their prediction accuracy needs improvement [16, 17]. In recent years, graph neural networks have shown advantages in processing graph-structured data. Some scholars have applied static graph neural networks to heat conduction network modeling [18]. However, static graph neural networks cannot effectively capture the dynamic changes of node relationships in heat conduction networks and ignore the dependencies in the time dimension, resulting in limited model performance in dynamic scenarios.

This paper mainly explores a thermal consumption prediction algorithm for smart manufacturing based on spatiotemporal graph neural networks, aiming to construct a model that can simultaneously capture the spatiotemporal dependencies of dynamic heat conduction networks. The specific content includes: first, a detailed explanation of the model structure, designing a spatiotemporal graph neural network architecture suitable for dynamic heat conduction networks, integrating modules for processing spatial graph structures and time series data; second, an in-depth study of spatial dependency modeling, using graph convolution operations to capture the spatial heat conduction relationships between nodes and their dynamic changes; then, modeling temporal dependencies, using time series analysis methods such as recurrent neural networks and temporal convolutional networks to explore the evolution patterns of the heat conduction process in the time dimension; finally, based on the constructed spatiotemporal graph neural network model, carrying out research on thermal consumption prediction in smart manufacturing, and verifying the effectiveness and superiority of the model using real production data.

2. RESEARCH PROBLEM DESCRIPTION AND MODEL FRAMEWORK

2.1 Research problem

The dynamic heat conduction network in intelligent manufacturing is defined as a directed graph $H=(N, R)$, where the node set $N=\{n_1, ..., n_V\}$ represents physical entities with heat conduction characteristics in the manufacturing system, such as processing equipment, material units, cooling devices, etc. Each node n_u is associated with a set of static and dynamic attributes: The static attributes include physical parameters such as thermal conductivity of materials, geometric dimensions, rated power, etc., while the dynamic attributes cover time-varying state data such as real-time temperature, operating load, energy consumption rate, etc. The directed edge set R represents the heat conduction relationship between nodes. The edge direction $n_u \rightarrow n_k$ reflects the direction of heat flow from entity n_u to n_k, and the edge weight q_{uk} quantifies the conduction intensity, determined by contact area, thermal resistance coefficient, operating condition parameters, etc. The time dimension $S=\{s_1, ..., s_l\}$ characterizes the dynamic evolution features of the network, making the network present differentiated connection weights and node attributes at different moments s_j, forming a time-dependent graph sequence $\{H_{s1}, ..., H_{sl}\}$. Figure 1 shows the schematic diagram of the spatiotemporal data flow of dynamic heat conduction in intelligent manufacturing.

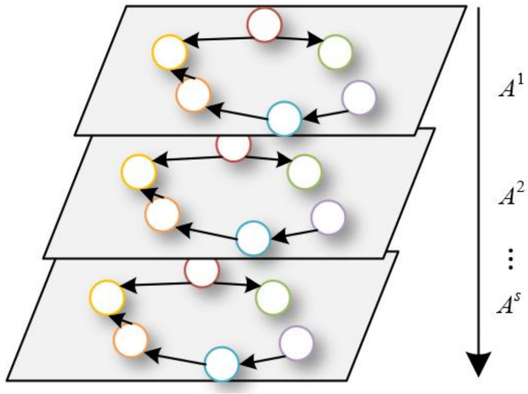


Figure 1. Spatiotemporal data flow of dynamic heat conduction in intelligent manufacturing

The core data attribute characteristics of the directed graph are reflected in spatiotemporal coupling and multi-source nature: in the spatial dimension, the directed edge structure between nodes explicitly expresses the physical path of heat flow conduction, while the node attributes not only include the thermodynamic parameters of the device itself, but also integrate real-time operating condition data collected by sensors, such as motor temperature, surface heat flux density of workpieces, providing the basis for graph neural networks to capture nonlinear dependencies of spatial heat conduction. In the time dimension, the node attributes and edge weights dynamically change along with the production process, such as heat source fluctuations caused by the switching on and off of processing equipment, and changes in thermal resistance induced by cooling system adjustments, forming a time-series-driven graph structure evolution. In addition, the data attributes integrate multi-source heterogeneous information, also including material property parameters in the process design stage and historical energy consumption records. After time-stamp alignment, these data form a spatiotemporal graph structure containing the node feature matrix A_s and the adjacency matrix X_s, providing rich input features for graph neural network modeling of the dynamic heat conduction process.

Based on this model, the core objective of heat consumption optimization is transformed into: designing strategies that can dynamically adapt to changes in production conditions by capturing the spatiotemporal coupled heat conduction patterns in the graph structure, to minimize the total system heat consumption under the premise of ensuring production quality. Based on the above directed graph model, the specific research problem of this paper can be summarized as: how to construct a graph neural network architecture that integrates spatiotemporal dependencies, to accurately model the spatiotemporal evolution patterns of node attributes and edge weights in the dynamic heat conduction network, and generate optimal heat consumption control strategies accordingly. This problem includes three core sub-problems: 1) Spatiotemporal feature modeling problem – Existing static graph models are difficult to handle the dynamic changes of edge weights and node attributes over time, requiring the design of suitable spatiotemporal graph convolution operators to jointly model the neighborhood dependency in spatial dimension and the sequential dependency in temporal dimension of the heat conduction network; 2) Multi-source data fusion problem – The heat conduction-related data in manufacturing systems have multi-source and heterogeneous characteristics, requiring research on how to convert these data into a unified representation processable by graph neural networks, avoiding information loss or feature conflict; 3) Optimization strategy mapping problem – Heat consumption optimization needs to balance production efficiency and energy cost, requiring the construction of a decision model with the prediction results of the graph neural network as input, combining constraints and objective functions to generate control strategies that can directly guide the actuators. Solving the above problems can break through the dependence of traditional heat consumption management methods on fixed conduction models, and provide a data-driven accurate solution for heat consumption optimization in complex dynamic scenarios of intelligent manufacturing. Figure 2 shows the overall architecture of the heat consumption optimization platform constructed in this paper for intelligent manufacturing.

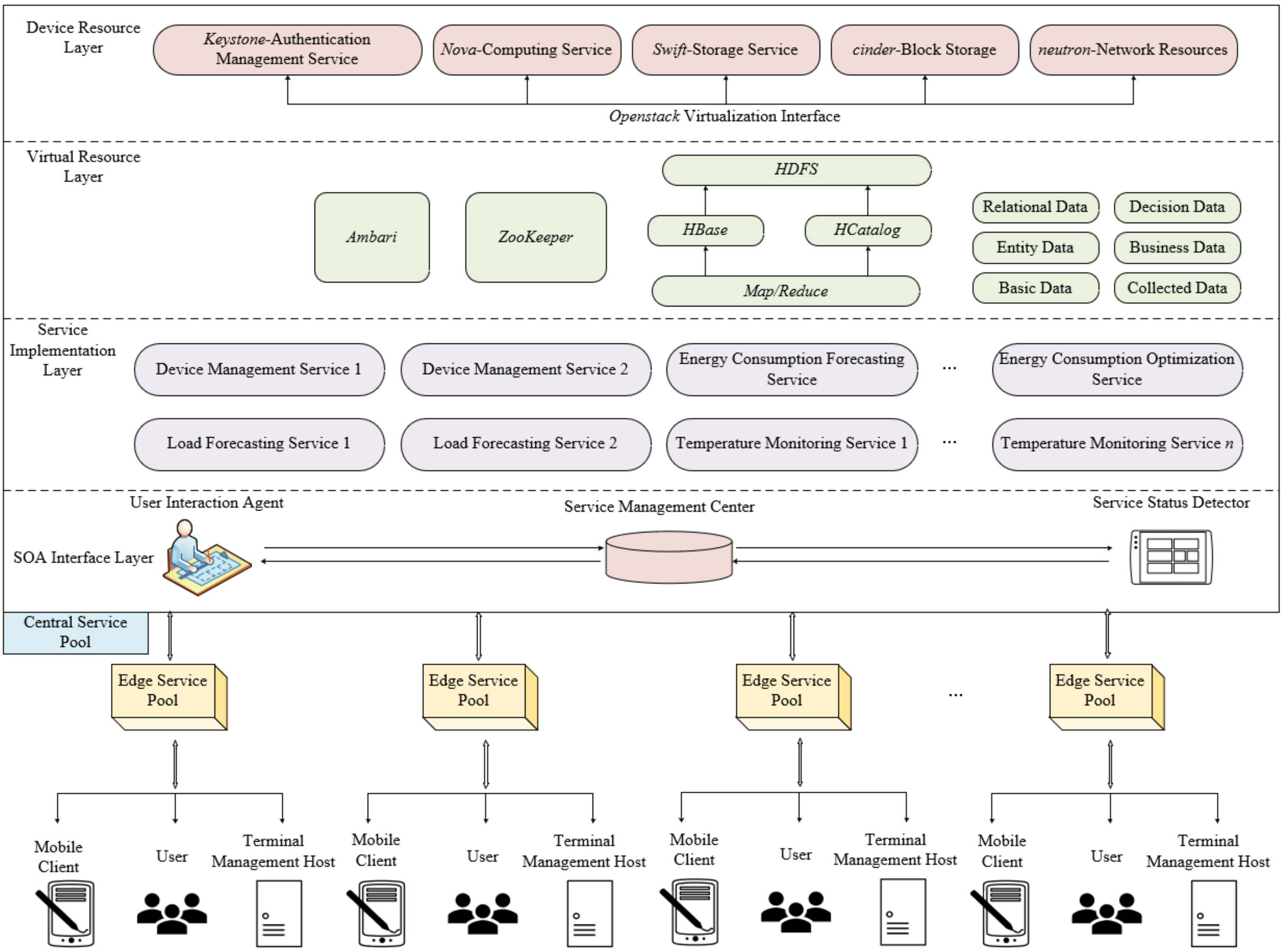


Figure 2. Overall architecture of heat consumption optimization platform in intelligent manufacturing

2.2 Model composition

For the directed graph structure of the dynamic heat conduction network in intelligent manufacturing, the model first captures the spatial heat conduction relationships between nodes through the diffusion graph convolution module. This module is based on node position representations, converting the heat flow conduction process among physical entities into signal diffusion on the graph structure. By designing diffusion operators regularized by the Laplacian matrix, the module realizes hierarchical aggregation of spatial features in node neighborhoods. Specifically, the inputs of the module are the node attribute matrix A_s and the adjacency matrix X_s. Through diffusion convolution operations, each node's local neighborhood information is fused with its own attributes, generating node embeddings with spatial dependencies. This module not only retains the directional edge structure information of the heat conduction network, but also distinguishes the spatial position characteristics of different physical entities through a position encoding mechanism. Its output is directly used as the input features of the subsequent dynamic Gated Recurrent Unit (GRU) module, providing basic spatial representations for spatiotemporal joint modeling.

The model achieves hierarchical modeling of temporal dependencies through the dynamic GRU module and Informer module. The D-GCGRU module embeds graph convolution operations into the gating mechanism of the traditional GRU, taking the spatial feature sequence generated by the diffusion graph convolution as input. It dynamically fuses the spatial neighborhood information at the current time and the historical hidden states through the gating units, thereby capturing local temporal patterns in time series data such as node temperature changes and heat flow rates. Complementarily, the Informer module addresses issues of gradient vanishing and insufficient global dependency capture in long-term heat consumption prediction. Through a self-attention mechanism and Fourier transform, it converts time series into frequency domain features, efficiently capturing global dependences at different time scales. Specifically, the Informer reduces computational complexity through probabilistic self-attention mechanisms, and extracts multi-resolution temporal features using hierarchical temporal aggregation strategies. Finally, it fuses with the output features of the D-GCGRU to generate prediction sequences containing spatiotemporal coupling information. The entire model architecture forms a three-level processing flow of "spatial feature extraction — local temporal modeling — global temporal dependency enhancement," which can not only accurately depict the real-time thermal coupling effects among devices in the heat conduction network, but also capture the long-term evolution trends of heat consumption in the production process, providing multi-dimensional prediction support for generating heat consumption optimization strategies. Figure 3 shows the heat consumption prediction model architecture in intelligent manufacturing.

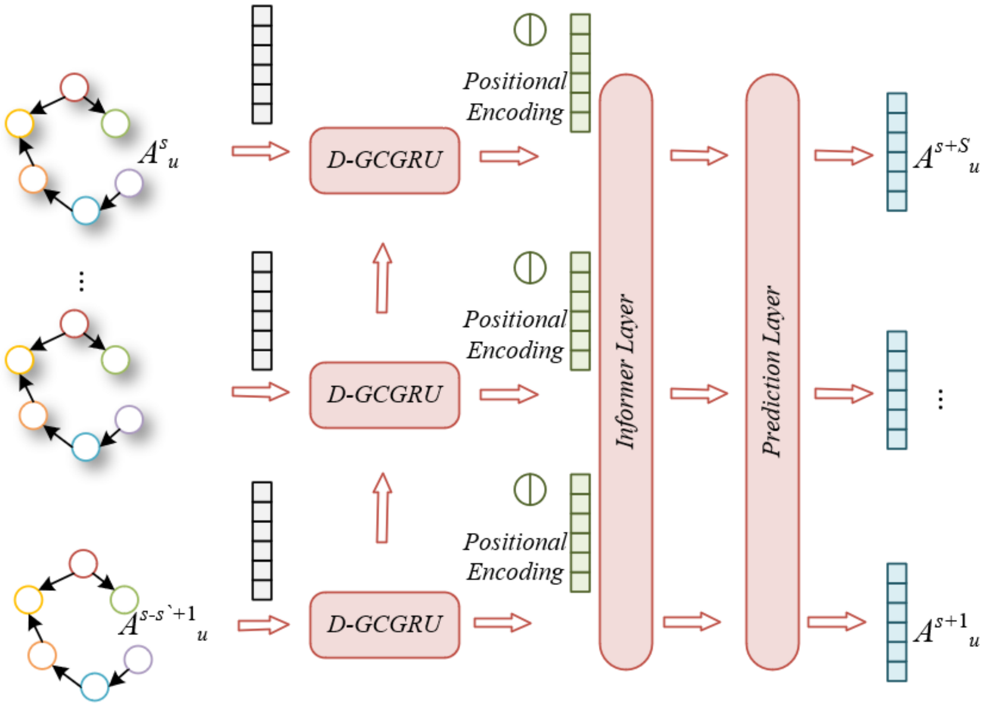


Figure 3. Heat Consumption prediction model architecture in intelligent manufacturing

3. SPATIAL HEAT MODELING IN SMART MANUFACTURING

In the dynamic heat conduction network of intelligent manufacturing, spatial dependency essentially refers to the spatial correlation characteristics formed by heat flow conduction among physical entities in the manufacturing system, manifested in the directionality of heat transfer between nodes, conduction intensity, and the influence of spatial layout on heat diffusion paths. In the dynamic heat conduction network of intelligent manufacturing, heat flow transfer has a clear physical directionality, and the directed graph modeling characteristics of diffusion graph convolution precisely match this requirement. To effectively capture the spatial dependencies of the dynamic heat conduction network, the model designs a diffusion graph convolution module that includes position representation. This module uses a diffusion operator regularized by the Laplacian matrix to hierarchically aggregate each node's local neighborhood information with its own attributes, realizing nonlinear modeling of spatial heat conduction patterns. Specifically, the module inputs the node attribute matrix A_s and the adjacency matrix X_s, and through diffusion convolution operations, aggregates the features of the neighborhood nodes of node n_u weighted by conduction intensity, generating node embeddings containing spatial dependencies. This forward modeling not only retains the physical causal relationships of heat conduction, but also captures long-range heat diffusion effects of higher-order neighborhoods through multi-layer convolution operations, enabling the model to accurately characterize the cascading propagation patterns of heat flow in the spatial dimension, and providing dependency information of forward conduction paths for heat consumption prediction.

In the design of heat consumption optimization strategies, in addition to forward prediction of heat flow influence, it is also necessary to determine key control nodes through backward analysis, and the mathematical properties of diffusion graph convolution provide support for this kind of backward modeling. From the perspective of backward derivation of spatial dependencies, although the adjacency matrix X_s of diffusion graph convolution is constructed based on the physical conduction direction, its transposed matrix X_s^T can be used to characterize the reverse dependency relationships between nodes during the optimization process, such as the sensitivity of the temperature change of node n_k to the heat consumption regulation of node n_u. For example, when the goal is to reduce the total heat consumption of the system, through backward propagation gradient analysis, diffusion graph convolution can identify upstream heat source nodes that have the most significant impact on the thermal states of downstream nodes, thereby guiding the optimization strategy to prioritize adjustment of the operating parameters of these nodes. In addition, the position encoding mechanism of diffusion graph convolution can quantify the impact of spatial layout on heat consumption in backward modeling. For example, high-heat-generating devices located at the center of the production line have a comprehensive impact on surrounding devices through multi-directional conduction paths. This can be captured by the model through gradient aggregation in backward propagation, and thus prioritized for regulation during optimization. This combination of forward modeling and backward derivation enables diffusion graph convolution to not only describe the physical propagation process of heat flow, but also provide key node identification and conduction path weight analysis in the spatial dimension for optimization targets, forming a "prediction–analysis–regulation" closed loop, and ultimately improving the relevance and effectiveness of heat consumption optimization strategies. Specifically, suppose the out-degree matrix is represented by $F_P=\Sigma_k X$, the in-degree matrix is represented by $F_U=\Sigma_k X^T$, the Sigmoid activation function is represented by δ, and the learnable parameters are represented by ϕ. The diffusion graph convolution operation is defined as:

$$A_{OUT}=\delta\left(\sum_{j=0}^{1}\left(\left(F_P^{-1}X_s\right)^j+\left(F_U^{-1}X_s^{\ T}\right)^j\right)A\varphi\right) \tag{1}$$

In the dynamic heat conduction network of intelligent manufacturing, the static weights of traditional predefined adjacency matrices are difficult to adapt to complex and variable production conditions. The fundamental flaw lies in ignoring the dynamic and multi-source nature of heat conduction relationships. The intensity of heat flow conduction is not only affected by spatial position but also closely related to dynamic factors such as real-time operating status and process parameters. For example, when the spindle of a machine tool adjusts its speed due to changes in the hardness of the processed material, the contact thermal resistance between it and the workpiece changes in real time, resulting in dynamic changes in conduction weight; similarly, when multiple devices work collaboratively, the cluster thermal effect may cause indirect thermal coupling between non-adjacent nodes, and this implicit correlation cannot be captured by predefined static weights. In addition, the rich information contained in multi-source data in manufacturing systems needs to be fused through learnable weights, while static weights can only express a single spatial proximity relationship and cannot realize the joint modeling of "spatial position + operating status + process constraints." Therefore, the adjacency matrix weights must be learned dynamically as model parameters along with the data to accurately portray the dynamic dependencies between nodes in the heat conduction network, and to avoid spatial feature extraction bias caused by fixed weights.

To achieve adaptive learning of adjacency matrix weights, the model adopts a dual-layer mechanism of "initial matrix initialization + data-driven optimization": First, an initial adjacency matrix is constructed based on geographical proximity and physical connection relationships from design drawings, serving as prior knowledge for weight learning to ensure the physical interpretability of model convergence; second, by introducing learnable parameters, the static matrix is upgraded to a dynamic matrix. Specifically, for each node n_u, a position representation is learned, that is, through linear transformation $o_u=Qa_u$, the pairwise relationship between arbitrary nodes is represented as follows:

$$E_{u,k}=\frac{\exp\left(\theta\left(T\left(o_u,o_k\right)\right)\right)}{\sum_{j=1}^{V}\exp\left(\theta\left(T\left(o_u,o_k\right)\right)\right)} \tag{2}$$

where, $T(o_u,o_k) = o^S_u o_k$. The following operation is then performed:

$$\tilde{E}=\begin{cases}E_{uk}, X_{uk}>0\\ 0, others\end{cases} \tag{3}$$

Combining the ideas of diffusion graph convolution and position representation learning, assuming that the out-degree and in-degree matrices of $\tilde{E}$ are represented by $\tilde{F}_P$ and $\tilde{F}_U$, respectively, and the activation function is denoted by δ, the graph convolution operation can be defined as follows:

$$A_{OUT}=\beta\left(\sum_{j=0}^{1}\left(\left(\tilde{F}_P^{-1}\tilde{E}\right)^j+\left(\tilde{F}_U^{-1}\tilde{E}^S\right)^j\right)A\varphi\right) \tag{4}$$

4. TEMPORAL HEAT MODELING IN MANUFACTURING

In the dynamic heat conduction network of intelligent manufacturing, local time dependency mainly manifests as high-frequency fluctuations and instantaneous associations of node states on a short time scale, such as transient heat flow changes caused by device start/stop or sudden temperature changes caused by cooling system adjustments. As a variant of recurrent neural networks, GRU's gating mechanism can effectively capture such short-term dynamic features: the update gate determines the retention degree of historical state to the current state, and the reset gate controls the forgetting rate of past information, thereby focusing on recent data's local dependencies when processing time series. Specifically in the heat consumption optimization scenario, GRU can dynamically integrate node historical temperature, power load and other dynamic attributes, such as the impact of the spindle temperature change of a machine tool at the previous moment on the current heat consumption. It also incorporates the heat conduction states of spatial neighbors into gate computation through graph convolution operations, achieving local time modeling of "node's own temporal features + spatial coupling effect of neighborhood." This mechanism is suitable for capturing short-term heat consumption fluctuations at the minute or second level, providing real-time feedback for heat flow regulation, and avoiding the bias of local dependency characterization caused by the lack of spatial information fusion in traditional temporal models.

Global time dependency focuses on long-term evolution patterns of the heat conduction network on time scales of hours, shifts, or even days, such as periodic changes of device accumulated heat loss during continuous processing, or the effect of diurnal temperature variation on the overall heat dissipation efficiency of the factory. Traditional recurrent neural networks have problems like gradient vanishing and high computational complexity when dealing with long sequences, while the Informer model, through probabilistic self-attention mechanisms and Fourier transform, can efficiently capture long-distance time dependencies, becoming the key to solving this issue. Specifically, Informer reduces computational cost through sparse self-attention operations, allowing the model to process longer time series without losing key information, and extracts multi-resolution temporal features using a hierarchical time aggregation strategy. In the heat consumption optimization scenario, Informer can capture heat consumption trends within production planning cycles and long-term influences such as seasonal variations on the energy efficiency of cooling systems, avoiding long-term prediction bias caused by relying solely on local temporal modeling. Complementing GRU's local modeling, Informer's global perspective provides cross-period planning support for heat consumption optimization strategies. The combination of the two achieves full coverage of the time dimension with "short-term dynamic response + long-term trend prediction."

Specifically, this paper replaces the matrix multiplication operation in GRU with the graph convolution operation shown in Eq. (4). Suppose the input and output at time s are represented by A^s and G^s, the reset and update gates at time s are denoted by e^s and i^s, the diffusion graph convolution operation is represented by $*\xi$, and the corresponding convolution kernel parameters are represented by ϕ_e, ϕ_i, and ϕ_Z, then the model expression is as follows:

$$\begin{aligned} e^s &= \delta\left(\varphi_e * \xi\left[A^s, G^{s-1}\right] + y_e\right) \\ i^s &= \beta\left(\varphi_i * \xi\left[A^s, G^{s-1}\right] + y_i\right) \\ Z^s &= \tanh\left(\varphi_Z * \xi\left[A^s, \left(e^s \otimes G^{s-1}\right)\right] + y_Z\right) \\ G^s &= i^s \otimes G^{s-1} + \left(1 - i^s\right) \otimes Z^s \end{aligned} \tag{5}$$

Based on the physical layout and design drawings of intelligent manufacturing equipment, an initial adjacency matrix is constructed with geographic proximity as the core, where nodes represent thermal conduction entities such as machining equipment and cooling devices, and the initial edge weights are determined by static spatial parameters such as equipment distance and contact area. Subsequently, the initial adjacency matrix is incorporated into model training as learnable parameters, allowing its weights to be dynamically updated with real-time thermal conduction data. For example, when the contact state between a machine tool and the workpiece changes due to process adjustments, the corresponding edge weight in the adjacency matrix is updated in real time through backpropagation to reflect the actual thermal conduction intensity under current working conditions.

On the basis of the dynamic adjacency matrix, a recurrent neural network layer D-GCGRU is constructed. This module embeds the diffusion graph convolution operation into the GRU gating mechanism. It first aggregates the spatial neighborhood information of nodes at the current moment through diffusion graph convolution, generating node feature vectors that contain spatial dependencies. Then, the update gate and reset gate of GRU combine the current spatial features with the historical hidden state to dynamically determine the degree of forgetting and retaining past information. Taking machine tool heat consumption optimization as an example, when the spindle experiences a rapid temperature rise due to high-speed processing, D-GCGRU captures the short-term temperature fluctuation sequence of the node and its neighboring cooling devices through the gating mechanism, and generates a hidden state representation of local thermal flow variation in real time, providing instant decision-making support for second-level or minute-level heat regulation, effectively coping with high-frequency dynamic events such as equipment switching and load mutation.

After being processed by the D-GCGRU layer, the output local temporal features are input to the Informer module to capture long-range temporal dependencies. First, positional encoding is performed on the local temporal features to generate an encoding vector containing relative temporal information for each time point, solving the self-attention mechanism’s insensitivity to temporal order. For example, when processing equipment heat consumption data recorded by shift, positional encoding can distinguish cyclical temporal features such as morning and night shifts. Then, the encoded sequence is input into the multi-head probabilistic sparse self-attention layer of the Informer. Through the sparsification operation, computational complexity is reduced, allowing the model to efficiently process long-sequence data at hour or day scale. Specifically, the sequence of node n_u over a certain time period $G_u[:]=[G^1{}_u, G^2{}_u, ..., G^s{}_u]$ is computed as follows:

$$X\left(W, J, N\right) = \text{softmax}\left(\frac{\bar{W}J^S}{\sqrt{f}}\right)N \tag{6}$$

To enhance the model’s expressive power for complex temporal patterns, the Informer adopts a multi-head attention mechanism, mapping the input sequence into multiple independent subspaces for parallel processing. Each head learns a different attention distribution, capturing dependencies in heat consumption sequences at different temporal scales. For instance, the first head focuses on minute-level fluctuations due to equipment switching, the second head attends to hour-level shift cycle patterns, and the third head captures day-level differences between weekdays and weekends. Taking the heat consumption of stamping dies in automobile manufacturing as an example, multi-head attention can simultaneously identify the instantaneous heat flow peaks from single stamping operations, the accumulated thermal loss after 100 continuous stampings, and the baseline temperature after daily mold cooling. The outputs of each head are fused through concatenation to form a comprehensive representation containing multi-dimensional temporal information, enabling the model to cope with high-frequency noise, mid-term trends, and long-term cycles in heat consumption sequences, providing richer temporal feature inputs for optimization strategies. Assuming the output of the u-th attention head is $HEAD_u$, and the output linear transformation matrix is Q^P, it can be expressed as:

$$MU\text{-}HEAD\left(G_u\right) = CONCAT\left(HE_1, ..., HE_T\right)Q^P \tag{7}$$

To ensure the Informer layer understands the relative positions of G^s, the output of the D-GCGRU is positionally encoded, that is, the time segment sequence $G_u[:]=[G^1{}_u, G^2{}_u, ..., G^s{}_u]$ is combined with the positional encoding r^s to generate a new representation. The formula is:

$$G_u^{s'} = G_u^s + r^s \tag{8}$$

where,

$$r^s = \begin{cases} SIN\left(s/10000^{2u/f_{MODrm}}\right), if \ s = 0, 2, 4... \\ COS\left(s/10000^{2u/f_{MODrm}}\right), others \end{cases} \tag{9}$$

After the multi-head attention layer, a self-attention distillation layer is introduced to optimize the computational results and address the complexity issue in long-sequence processing. The distillation layer adopts a hierarchical aggregation strategy to progressively compress the attention weights of long sequences, retaining only the most critical global dependencies for heat consumption prediction. For example, when processing 30-day heat consumption data of a production line, the distillation layer can automatically identify the 5 most influential historical time points for the current heat consumption, ignoring irrelevant noise data, thereby effectively reducing computational complexity while maintaining prediction accuracy. Specifically, for the output $G^{sv}{}_u$ of the multi-head attention, the computation of the self-attention distillation layer is:

$$G_u^{s^l} = \mathrm{maxpool}\left(\delta\left(\left(CONV1D\left(\left[G_u^{s^v}\right]\right)\right)\right)\right) \tag{10}$$

Finally, the global temporal features output by the Informer module are fused with the local spatiotemporal features from the D-GCGRU layer and input into the prediction layer to generate future heat consumption sequences for each node. During the fusion process, the model dynamically adjusts the weights of local and global features through a fully connected layer. For instance, in scenarios of sudden heat anomalies due to equipment failure, the weight of local features from D-GCGRU is automatically increased for rapid response; in stable production phases, the model relies on the global trend features from Informer for optimization. The final heat consumption prediction results can be directly used to guide actuators: short-term predictions drive real-time regulation of cooling valves, heating elements, etc.; long-term predictions assist production planning systems in adjusting process parameters. Figure 4 shows the constructed Informer module architecture.

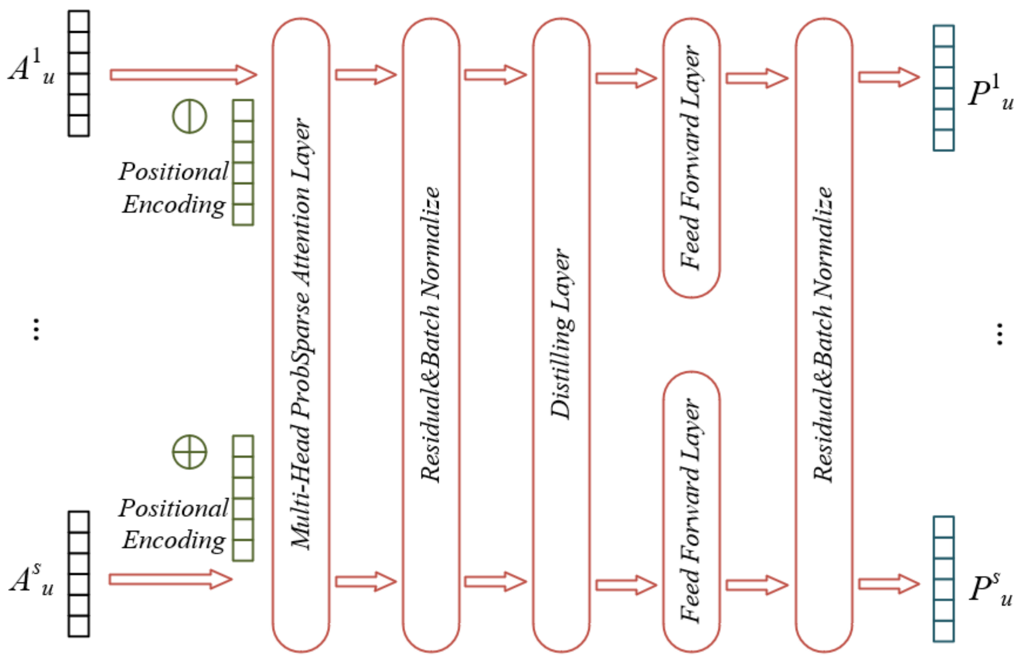


Figure 4. Informer module architecture

5. HEAT PREDICTION VIA DYNAMIC NETWORKS

The core input of the prediction layer is the global temporal features output by the Informer module, which have integrated the long-term temporal dependencies and multi-scale time patterns of each node in the dynamic heat transfer network. To adapt to the heat consumption prediction task, the node feature matrix output by the Informer is first dimensionally integrated, concatenating each node's global temporal representation with the local spatiotemporal features extracted by the D-GCGRU module to form a composite feature vector containing "node attributes–spatial correlation–temporal evolution". Then, a multi-layer feedforward network processes this feature through hierarchical nonlinear transformations: the first layer of neurons captures the nonlinear mapping relationship between heat consumption and device operating parameters and process conditions through the ReLU activation function; the middle layer utilizes Dropout to avoid overfitting; the final layer outputs the heat consumption sequence $[A^{s+1}, ..., A^{s+S}]$ for the next S time points according to the prediction length S. Taking machine tool heat consumption prediction as an example, this network can convert features such as the spindle's historical temperature curve, the real-time heat dissipation efficiency of adjacent cooling devices, and the seasonal variation of workshop ambient temperature into hourly heat consumption predictions for the next 24 hours, providing data support for strategies such as equipment preheating and cooling system power regulation.

6. EXPERIMENTAL RESULTS AND ANALYSIS

As shown in Table 1, regarding the R^2 metric, the proposed method achieves 0.9752, significantly higher than ST-GNN (0.9652), GraphSAGE (0.9613), LSTM (0.9546), and ARIMA (0.9485), indicating that the proposed method has better goodness of fit for the heat consumption data and can explain more variation of the dependent variable. In terms of error metrics, the proposed method's RMSE is 0.945, lower than ST-GNN (1.1236), GraphSAGE (1.2358), etc., and the MAPE and MAE are 2.8954 and 0.4752 respectively, also outperforming other models. This demonstrates that the proposed method controls both absolute and relative errors at a lower level when predicting heat consumption, and the deviation between the predicted and actual values is smaller. Combining with the content of this study, the proposed method, by designing a spatiotemporal graph neural network architecture adapted to the dynamic heat transfer network, effectively integrates spatial graph convolution operations and time series analysis, and can accurately capture the spatiotemporal dependencies in the heat transfer process. In contrast, although ST-GNN, GraphSAGE and other models involve graph structure or sequence processing, they are insufficient in capturing dynamics or spatiotemporal collaborative modeling, resulting in poorer fitting and error control compared to the proposed method. Traditional sequence models such as LSTM and ARIMA lack the characterization of spatial heat transfer relationships, leading to even more significant performance gaps.

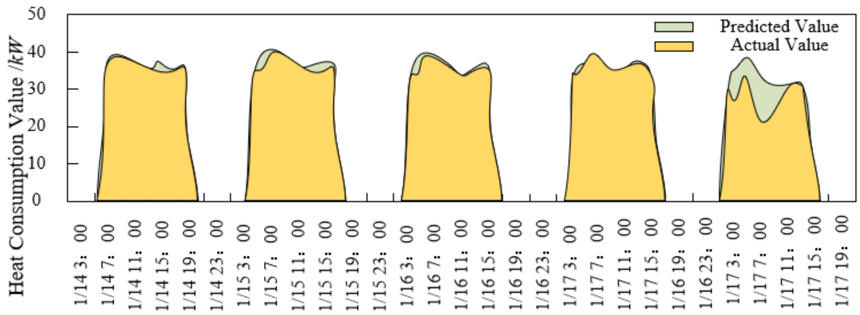
Table 2 shows the comparison of error between simulated and actual values of intelligent manufacturing heat consumption under different environments and layouts. For the Normalized Mean Bias Error (NMBE) metric, the open layout is 7.65%, and the dense layout is 6.4%; for the Coefficient of Variation of Root Mean Square Error (CVRMSE) metric, the open layout is 14.23%, and the dense layout is 13.26%. These two metrics both reflect the error level between simulated and actual values, with smaller values indicating smaller errors and more accurate predictions. The model captures the dynamic variation of spatial heat transfer relationships between nodes through graph convolution operations, demonstrating effective modeling capability of spatial dependencies under different layouts. In dense layouts, nodes have tighter heat transfer connections, and the model captures the dynamic features of heat transfer in such scenarios more precisely, thus resulting in relatively smaller errors; in open layouts, heat transfer is more complex due to environmental influences, but the model still controls errors within a reasonable range. Overall, the model can accurately predict heat consumption under different environmental and layout scenarios, effectively reflecting the spatiotemporal dependencies of the dynamic heat transfer network, further verifying the effectiveness and adaptability of the spatiotemporal graph neural network-based heat consumption prediction model in intelligent manufacturing, and providing a reliable predictive basis for subsequent heat consumption optimization strategies.

Table 1. Performance comparison of different intelligent manufacturing heat consumption prediction models

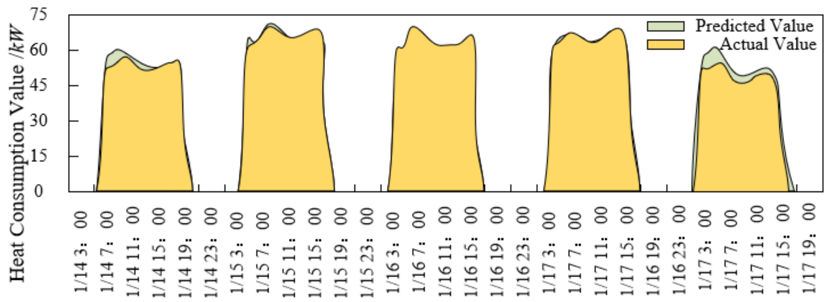
	R^2	*RMSE*	*MAPE*	*MAE*
Proposed Method	0.9752	0.945	2.8954	0.4752
ST-GNN	0.9652	1.1236	4.1235	0.6895
GraphSAGE	0.9613	1.2358	5.1236	0.8452
LSTM	0.9546	1.3265	5.5632	0.9123
ARIMA	0.9485	1.5689	6.3245	1.1256

Table 2. Comparison of error between simulated and actual values of intelligent manufacturing heat consumption under different environments and layouts

Metric	Open Layout	Dense Layout
NMBE	7.65%	6.4%
CVRMSE	14.23%	13.26%



1) Open-layout



2) Dense-layout

Figure 5. Comparison of predicted and actual heat consumption values under different environments and layouts in intelligent manufacturing

Figure 5 shows the comparison between predicted and actual heat consumption values in intelligent manufacturing under different environmental and layout conditions. In the open-layout scenario, the fluctuation trend of the predicted values (light green) and the actual values (yellow) is highly consistent, and the variation amplitude of heat consumption at each time point is similar. For example, during the period from 14:13 to 14:17, both show a significant upward trend in heat consumption, with peak values being close. Under the dense-layout condition, the predicted values also closely follow the changes of the actual values. For instance, during 14:13 to 14:17, the heat consumption increase amplitude and rhythm of the two are almost identical, indicating that the predicted values can reflect the dynamic change of actual heat consumption well under different layouts. Combined with the content of this paper, the model effectively captures the spatiotemporal dependencies of the dynamic heat conduction network by integrating spatial graph structure and time series data processing modules. The graph convolution operation accurately characterizes the dynamic changes of spatial heat conduction relationships between nodes, and the time series analysis method extracts the evolution pattern of heat conduction on the temporal dimension. As shown in the figure, whether it is open-layout or dense-layout, the predicted values are closely aligned with the actual values. Even though the spatial heat conduction characteristics of the two layouts are significantly different, the model can still adapt accurately. This fully verifies the effectiveness and robustness of the heat consumption prediction model based on spatiotemporal graph neural networks under different environmental and layout scenarios, which can provide accurate and reliable prediction support for heat consumption optimization strategies in intelligent manufacturing.

The heat consumption prediction results present the heat consumption values and variation trends of equipment, process, and other links in future time periods. Through in-depth analysis of the prediction data, the key factors affecting heat consumption are identified. For example, the prediction shows that the heat consumption of a certain machine tool increases significantly during a specific machining phase. By combining equipment operation parameters and the spatial dependencies of the heat conduction network, it is determined that the increase in heat generation is caused by excessive cutting speed or significant thermal aggregation effect due to proximity to adjacent equipment, thereby identifying key parameters to be adjusted and directions for optimization. According to the identified key influencing factors, corresponding heat consumption optimization strategies are matched. If the prediction shows that the heat consumption is too high due to unreasonable operation parameters of the equipment, then a dynamic adjustment strategy of equipment operation parameters is adopted. If it is found that the heat consumption of different production tasks is concentrated in time distribution, affecting overall heat consumption efficiency, a production task scheduling optimization scheme is selected. For example, if multiple high-heat-generating devices are predicted to operate at the same time period, resulting in excessive local thermal load in the workshop, a production task scheduling optimization strategy is matched to reschedule some tasks to other time periods.

Using the spatial dependencies of the dynamic heat conduction network, the initially matched strategy is refined. Taking the equipment layout and spatial optimization scheme as an example, according to the heat consumption prediction results and the heat conduction relationships and edge weights between nodes in the directed graph, it is determined which devices have strong thermal coupling effects and need to be rearranged. For example, if the prediction shows that several devices in a certain area have high overall heat consumption due to close proximity and mutual heat conduction, a specific equipment relocation plan is formulated based on the spatial dependencies, including node position encoding and adjacency matrix information. This plan clarifies which devices should be moved to where, and how to adjust the layout of cooling devices to reduce heat consumption. A quantitative analysis is conducted on the refined plan to evaluate its feasibility and optimization effect. By simulating the changes in heat consumption after the implementation of different plans and comparing the prediction results, the optimal plan is selected. For example, for the strategy of dynamic adjustment of equipment operation parameters, the degree of heat consumption reduction under different parameter adjustment ranges is simulated, and the best parameter adjustment value is determined by combining factors such as production efficiency and machining accuracy, ultimately forming a complete and executable heat consumption optimization plan.

Table 3. Total heat consumption and energy-saving rate of intelligent manufacturing in different test cases

Optimization Strategy	Test Period							
	Case 1		Case 2		Case 3		Case 4	
Dynamic Adjustment of Equipment Operation Parameters	456.32	7.6%	512.23	6.2%	489.36	6.3%	465.32	6.6%
Optimization of Production Task Scheduling	458.32	7.4%	514.36	6.3%	489.32	6.1%	468.62	6.1%
Heat Recovery and Reuse	459.31	7.3%	512.39	5.7%	487.21	6.1%	469.21	6.8%
Intelligent Temperature Control System Coordination	512.36	7.4%	528.69	6.3%	512.36	6.1%	512.58	6.6%

Table 3 presents the total heat consumption and energy-saving rates of intelligent manufacturing under different optimization strategies across various test cases. The dynamic adjustment of equipment operating parameters strategy achieved total heat consumption values of 456.32, 512.23, 489.36, and 465.32 in Cases 1 to 4, with energy-saving rates of 7.6%, 6.2%, 6.3%, and 6.6%, respectively. The production task scheduling optimization strategy resulted in heat consumption totals of 458.32, 514.36, 489.32, and 468.62, with energy-saving rates of 7.4%, 6.3%, 6.1%, and 6.1%. The heat recovery and reuse strategy yielded heat consumption values of 459.31, 512.39, 487.21, and 469.21, with energy-saving rates of 7.3%, 5.7%, 6.1%, and 6.8%. The intelligent temperature control system linkage strategy achieved heat consumption totals of 512.36, 528.69, 512.36, and 512.58, with energy-saving rates of 7.4%, 6.3%, 6.1%, and 6.6%. Each strategy achieved varying degrees of energy savings, with total heat consumption values fluctuating within a reasonable range. Based on the research presented in this paper, the spatiotemporal graph neural network model accurately captures the spatiotemporal dependencies of dynamic heat conduction networks, enabling the effective implementation of various optimization strategies. Strategies such as dynamic adjustment of equipment operating parameters and production task scheduling optimization leverage the model's analysis of spatial heat conduction relationships and temporal evolution patterns to achieve targeted control of heat consumption. For instance, the heat recovery and reuse strategy enhances waste heat utilization efficiency through the model's modeling of heat conduction paths, achieving an energy-saving rate of 6.8% in Case 4. The intelligent temperature control system linkage strategy utilizes the model's exploration of temporal heat consumption evolution to adjust temperature control parameters in real time, achieving an energy-saving rate of 7.4% in Case 1. These data demonstrate that heat consumption optimization strategies based on this model effectively reduce heat consumption and improve energy-saving rates across different scenarios, validating the effectiveness of the optimization strategies and the reliability of the model's support, providing practical and feasible pathways for achieving energy savings in intelligent manufacturing.

Figure 6 illustrates the variations in heat consumption values at different times of the day for various intelligent manufacturing heat consumption optimization strategies in Case 1. The intelligent temperature control system linkage strategy had a heat consumption value of approximately 48 kW at 8:00, which subsequently fluctuated downward. The heat recovery and reuse strategy started at around 44 kW at 8:00, showing a trend of decreasing, then increasing, and then decreasing again. The production task scheduling optimization strategy began at about 43 kW at 8:00, with more frequent fluctuations thereafter. The dynamic adjustment of equipment operating parameters strategy started at approximately 44 kW at 8:00, maintaining a relatively stable overall trend. At all times, the heat consumption values for each strategy were lower than those in the unoptimized state, and they exhibited fluctuations throughout the day that aligned with production activities, reflecting the effective regulation of heat consumption by the optimization strategies. Experimental results indicate that these optimization strategies rely on the spatiotemporal graph neural network model's precise capture of the spatiotemporal dependencies in dynamic heat conduction networks. The intelligent temperature control system linkage strategy utilizes the model's analysis of temporal heat consumption evolution patterns to adjust temperature control parameters in real time, effectively controlling heat consumption growth during peak production periods. The heat recovery and reuse strategy enhance waste heat utilization efficiency through the model's modeling of spatial heat conduction relationships, reducing overall heat consumption. The production task scheduling optimization and dynamic adjustment of equipment operating parameters strategies also achieve targeted control of heat consumption based on the model's analysis of spatial dependencies and temporal sequence features. The data in the figure demonstrate that each strategy effectively reduces heat consumption at different times of the day, validating the effectiveness and adaptability of the heat consumption optimization strategies based on this model in actual production, and highlighting the feasibility and advantages of achieving heat consumption optimization through dynamic heat conduction network modeling.

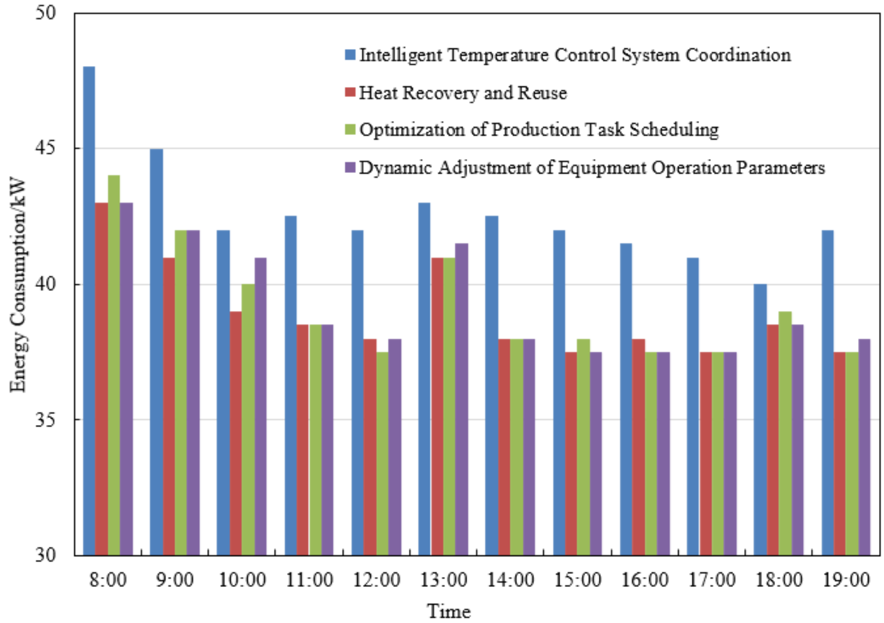


Figure 6. Heat consumption at all optimized time points of intelligent manufacturing in case 1 throughout one day

7. CONCLUSION

This study focused on optimizing heat consumption in intelligent manufacturing by proposing a heat consumption prediction algorithm based on spatiotemporal graph neural networks. The model integrated spatial graph structures with time series processing modules to deeply explore the modeling of spatial and temporal dependencies. Validated by real production data, the model outperformed comparative methods such as ST-GNN, GraphSAGE, and LSTM across various metrics. Additionally, optimization strategies based on the model achieve energy-saving rates exceeding 5% in different test cases, effectively demonstrating the feasibility of precise heat consumption prediction and optimization through dynamic heat conduction network modeling. This provides innovative methods and technical support for energy conservation and consumption reduction in intelligent manufacturing, highlighting the study's significant value in enhancing energy utilization efficiency and promoting green manufacturing.

However, the study has limitations: the model relies on specific scenario data, and its generalizability needs further validation; the complex architecture of the spatiotemporal graph neural network demands high computational resources, limiting its applicability in certain scenarios. Future work may expand data coverage to include heat consumption data from multiple industries and scenarios to enhance model generalization; explore model lightweighting techniques to reduce computational costs and improve deployment convenience; and further integrate emerging technologies such as reinforcement learning and digital twins to deepen the regulation of complex behaviors in dynamic heat conduction networks, advancing the research toward more efficient, intelligent, and universal directions, aiding the manufacturing industry in achieving green and low-carbon transformation, and expanding the application boundaries and depth of the research in cross-industry intelligent manufacturing scenarios.

REFERENCES

[1] Famularo, D., Fortino, G., Pupo, F., Giannini, F., Franzè, G. (2024). An intelligent multi-layer control architecture for logistics operations of autonomous vehicles in manufacturing systems. IEEE Transactions on Automation Science and Engineering, 22: 7296-7311. https://doi.org/10.1109/TASE.2024.3435342

[2] Szántó, N., Monek, G.D., Fischer, S. (2024). Development of an immersive, digital twin-supported smart reconfigurable educational platform for manufacturing training: A proof of concept. Journal of Engineering Management and Systems Engineering, 3(4): 199-209. https://doi.org/10.56578/jemse030402

[3] Wang, K., Peng, Y., Duan, M.W., Zhang, X.Y., Huang, Y., He, S., Chen, Z.W., Guo, S., Li, X.P. (2024). Review on intelligent additive manufacturing of multidimensional heterogeneous large components. Chinese Science Bulletin, 69(17): 2401-2416. https://doi.org/10.1360/TB-2023-0569

[4] Arain, M.B., Hussain, D., Al-Yarimi, F.A., Shaheen, S., Hu, J. (2025). Intelligent predictive networks of swimming of microorganism for tangent hyperbolic fluid flow with heat and mass transfer and Thompson and Troian slip conditions using Levenberg-Marquardt algorithm. Engineering Applications of Artificial Intelligence, 152: 110749. https://doi.org/10.1016/j.engappai.2025.110749

[5] Savaş, A. (2024). Numerical simulation and analysis of residual stress in B91 steel deposition using wire arc additive manufacturing. Precision Mechanics & Digital Fabrication, 1(2): 102-110. https://doi.org/10.56578/pmdf010205

[6] Ferreira, R.P.M., de Freitas Pinto, R.L.U., Mahey, P. (2025). Network optimization model approach to the volume-to-point heat conduction problem. Optimization and Engineering, 26(1): 229-251. https://doi.org/10.1007/s11081-024-09895-3

[7] Manavi, S., Becker, T., Fattahi, E. (2023). Enhanced surrogate modelling of heat conduction problems using physics-informed neural network framework. International Communications in Heat and Mass Transfer, 142: 106662. https://doi.org/10.1016/j.icheatmasstransfer.2023.106662

[8] Lee, S. (2023). Heat conduction simulation of one-dimensional element structure using interaction network. Transactions of the Korean Society of Mechanical Engineers, A, 47(12): 995-1002.

[9] Kumar Sah, D., Vahabi, M., Fotouhi, H. (2025). Federated learning at the edge in Industrial Internet of Things: A review. Sustainable Computing: Informatics and Systems, 46: 101087. https://doi.org/10.1016/j.suscom.2025.101087

[10] Dong, Y., Cao, X., Peng, X., Lv, J., Li, Q. (2025). In-pile heat conduction model of the dispersion nuclear fuel plate with particle agglomeration Part II: Predicting the effective thermal conductivity under the in-pile thermal transfer pattern based on a deep neural network. Journal of Nuclear Materials, 607: 155692. https://doi.org/10.1016/j.jnucmat.2025.155692

[11] Fu, Y., Horton, R., Heitman, J. (2025). Neural network estimation of thermal conductivity across full saturation for various soil types. Computers and Electronics in Agriculture, 235: 110321. https://doi.org/10.1016/j.compag.2025.110321

[12] Adam, A., Fang, H., Li, X. (2024). Effective thermal conductivity estimation using a convolutional neural network and its application in topology optimization. Energy and AI, 15: 100310. https://doi.org/10.1016/j.egyai.2023.100310

[13] Sagawa, D., Tanaka, K. (2023). Machine learning-based estimation of COP and multi-objective optimization of operation strategy for heat source considering electricity cost and on-site consumption of renewable energy. Energies, 16(13): 4893. https://doi.org/10.3390/en16134893

[14] Pourmirzaagha, H., Rouhi, S. (2023). Molecular dynamic simulations of the heat transfer in double-layered graphene/silicene nanosheets. Physica B: Condensed Matter, 666: 415079. https://doi.org/10.1016/j.physb.2023.415079

[15] Abouelregal, A.E., Mohammad-Sedighi, H., Faghidian, S.A., Shirazi, A.H. (2021). Temperature-dependent physical characteristics of the rotating nonlocal nanobeams subject to a varying heat source and a dynamic load. Facta Universitatis, Series: Mechanical Engineering, 19(4): 633-656. https://doi.org/10.22190/FUME201222024A

[16] Jayashankara, M., Sharma, A., Singh, A.K., Chanak, P., Singh, S.K. (2024). A novel intelligent modeling and prediction of heat energy consumption in smart buildings. Energy and Buildings, 310: 114105. https://doi.org/10.1016/j.enbuild.2024.114105

[17] Koleda, P., Čuchor, T., Koleda, P., Rajko, Ľ. (2024). Prediction models for the milling of heat-treated beech wood based on the consumption of energy. Applied Sciences, 14(20): 9539. https://doi.org/10.20944/preprints202409.0200.v1

[18] Peng, J.Z., Aubry, N., Li, Y.B., Chen, Z.H., Mei, M., Hua, Y. (2024). HCP-PIGN: Efficient heat conduction prediction by physics-informed graph convolutional neural network. International Journal of Heat and Fluid Flow, 109: 109552. https://doi.org/10.1016/j.ijheatfluidflow.2024.109552