

Multi-zone indoor temperature prediction based on Graph Attention Network and Gated Recurrent Unit*

Chunxiang Zhou, *Student Member, IEEE*, Zhanbo Xu, *Member, IEEE*, Jiang Wu, *Member, IEEE*,
Kun Liu, *Member, IEEE*, and Xiaohong Guan, *Fellow, IEEE*

Abstract—Indoor temperature have significant influence on load forecasting, comfort control and security monitoring. Achieving accurate temperature prediction can provide key basic data for energy efficiency and building safety and comfort. In the case of multiple zones, the heat transfer process in adjacent zones can have an important impact on the dynamics of indoor temperature. This paper focuses on the influence of heat transfer process in multiple adjacent zones. To describe the interactions of temperature among the multiple zones, we consider the zones as nodes and the connected walls as edges based on actual layouts to construct the graph network. For the non-linearity of the heat transfer process, we propose a novel multi-zone indoor temperature prediction model based on graph attention mechanism and recurrent network to achieve one-step ahead and multi-step ahead temperature predictions. The accuracy of this model was further verified by using data generated from EnergyPlus simulations. The best predicted result had an RMSE value of 0.47, an MAE value of 0.37, and an R^2 value of 0.94.

I. INTRODUCTION

Temperature is a key basic data for thermal comfort [1], energy consumption [2] and equipment safety [3]. An accurate prediction of temperature is necessary in many fields. Indoor temperature is affected by a variety of factors such as equipment, people, outdoor environment and adjacent zone's temperature. This paper focuses on the influence of the heat transfer process between adjacent zones, aiming to improve the accuracy of indoor temperature prediction. By considering heat transfer process, the accurate prediction can better guide pre-cooling and pre-heating to achieve a win-win situation of personnel comfort and energy consumption reduction [4].

In recent years, many efforts have been made to have a accurate temperature prediction. Prior to the data-driven approach, physical models was widely used, such as using simulation software [5], Resistance-Capacitance (RC) model [6], etc. Mazzeo et al. [7] predicted the indoor temperature in the same building based on the physical models built into three simulation softwares, and compared their prediction accuracy. Li et al. [8] using the unscented Kalman filter in combination with resistance-capacitance (RC) models to online estimate the thermal dynamics of a single room.

*This work was supported in part by National Natural Science Foundation of China (61803297, U1766205)

C. Zhou, Z. Xu, and J. Wu, K. Liu, X. Guan are with the Moe Klinns Lab and Systems Engineering Institute of Xi'an Jiaotong University, Xi'an, Shaanxi, 710049, China (e-mail: cxzhou@stu.xjtu.edu.cn, zbxu@sei.xjtu.edu.cn, jwu@sei.xjtu.edu.cn, kliu@sei.xjtu.edu.cn, xhguan@xjtu.edu.cn).

In addition to the physical modeling methods above, data-driven approaches [9] are also commonly used. Kumar et al. [10] analyzed the correlation between sensors and temperature using particle swarm optimization algorithms and statistical tools to improve the accuracy of temperature prediction. Gustin et al. [11] developed a recursive, autoregressive time series model using exogenous inputs to predict temperature during summer. Xu et al. [12] used an improved long short-term memory neural network to predict multi-room indoor temperature and compared the performance with commonly used prediction methods to verify the superior performance of the proposed method. Afroz et al. [2] presented a nonlinear autoregressive network with exogenous inputs-based system identification method for temperature prediction in spaces with multiple zones.

However, although the above research has achieved relatively accurate prediction of indoor temperature, there are still the following challenges: First, heat transfer process is closely related to wall structure and building materials [13], [14], and has complex non-linearity. Therefore, it is difficult to construct a mechanism model to accurately represent the non-linear relationship. Second, it is costly to obtain accurate detailed information in physical models. Without the detailed information, the accuracy of prediction will be poor. It is noted that existing studies of heat transfer process such as EnergyPlus often use complex physical models with empirical parameters for description.

To address the above challenges, we make the following contributions. First, we propose a novel multi-zone indoor temperature prediction model integrating physics-based and data-driven methods. By combining the physical part and the data-driven part, we got more accurate temperature predictions. Second, we introduce graph attention network to represent the non-linear heat transfer process between physical zones, and recurrent neural network to memorize the temporal features of the temperature data.

Compared with previous studies, on the one hand, this paper predicts the temperature for multiple rooms, taking into account the interaction between rooms, which is more valuable than single-room prediction. On the other hand, existing studies either base on physical models or data-driven models for prediction, this paper combines the advantages of physics-based and data-driven methods to achieve higher prediction accuracy. Since neural networks can accurately describe the non-linearity between inputs and outputs, complex physical models and detailed parameters on heat transfer process are no longer needed. Finally, The proposed model

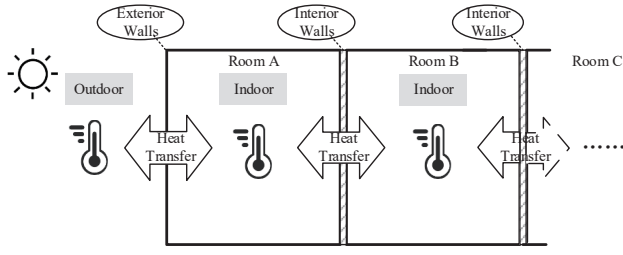


Fig. 1. Heat transfer process in adjacent zones.

is validated using a case study simulated by EnergyPlus.

This paper is organized as follows: Section II describes the problem in general. Section III introduces the modeling methods used in this paper. Section IV conducts numerical tests and analysis for specific cases, and Section V summarizes the whole paper and suggests directions for further research.

II. PROBLEM FORMULATION

In reality, zones have different layouts and complex connection relationships between adjacent zones. Due to the temperature difference between different zones and between zones and outdoor environment, the indoor temperature of each zone is influenced by the heat transfer process through the interior or exterior wall, as shown in Fig.1.

Our model's physical part is built by abstracting the actual zone topology to the graph structure and contains the actual physical connection information. Consider N adjacent zones, the outdoor environment and each zone are considered as nodes and the walls between the nodes are considered as edges, so the directed graph $G=(\mathbf{V},\mathbf{E})$ can be constructed according to the physical topology, where \mathbf{V} is a set of nodes, $\mathbf{V}=\{v_0, v_1, \dots, v_N\}$, and \mathbf{E} is a set of edges, $\mathbf{E} \subseteq \mathbf{V} \times \mathbf{V}$. Note that v_0 represents the outdoor environment node. Node pairs connected with edges mean that they are adjacent to each other physically and there is a heat transfer process between them. On the contrary, node pairs without edges connected mean that they are not adjacent to each other physically and there is no heat transfer process between them. The direction of the edge represents the bidirectional heat transfer process through the wall between the nodes, as shown in Fig.2. The properties of the edge is the intensity of the heat transfer process. Note that, the dashed circles and the dashed lines represent other nodes and edges that may exist, respectively.

The data-driven part of the proposed model consists of properties of each zone and heat transfer process between zones. For zones in Fig.1, each zone may have different usage scenarios. So we treat each zone as a thermal zone, representing that its indoor temperature is uniform and may not be the same as others. And each zone's unique property is its indoor temperature. The heat transfer process between zones can be generally described as shown below:

$$Y_{a,b}^k = f(\mathbf{X}_a^k, \mathbf{X}_b^k, \theta) \quad (1)$$

where $Y_{a,b}^k$ represents the heat transferred from zone a to zone b in time period k , \mathbf{X}_a^k and \mathbf{X}_b^k are the feature vectors of zone

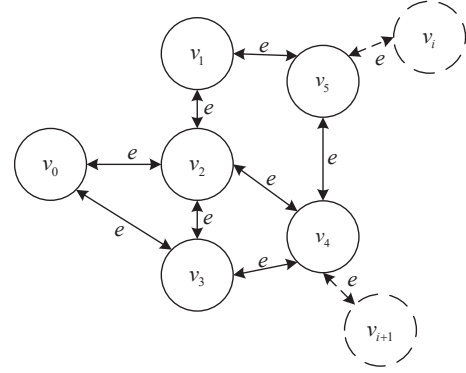


Fig. 2. Directed graph based on adjacent zones.

a and zone b , respectively. θ is an unknown parameter vector of walls, which needs the data-driven part to estimate. f is an function contains the influence of the heat transfer process between zones. Here, zones correspond to the nodes in the physical part. Based on the edge information in the physical part, the influence coefficients between each pair of zones with edges connected will be learned by the data-driven part.

The physical part and the data-driven part complement each other. Without the physical part, physical connection information is difficult to obtain through data-driven means due to the lack of representative data [1]. Similarly, without the data-driven part, the parameters of the nonlinear heat transfer process are difficult to obtain by means of physical modeling due to the non-linearity.

Graph neural network combines the concept of nodes and edges in physical graph structure with the data-driven methods of neural networks. Therefore, it can effectively describe and solve problems mentioned above. By combining the two parts above, our model can take both of the advantages of physics-based and data-driven methods, and we can transform the multi-zone temperature prediction problem into the problem of predicting node properties in graph. In this way, the properties of each node represent the temperature of the corresponding zone, and the heat transfer process is transformed into a flow process along the edges in the graph. Note that, in order to accurately investigate the impact of the heat transfer process on indoor temperature, the heat dissipation of indoor occupants and equipments are not considered in this study.

III. SOLUTION METHODOLOGY

In this paper, the reasons for choosing graph neural network include two: First, the non-linear heat transfer process needs to be described by a neural network that is good at describing non-linearity. Second, each zone has different neighbors, meaning that they are non-Euclidean structured data. Graph neural network incorporates the concepts of nodes and edges in the graph structure and retain the complete topological information of the graph, thus it can handle non-Euclidean structured data.

The concept of graph neural network was first introduced in 2005 by Dr. Franco [15]. Later in 2008, the theoretical basis of graph neural network was defined [16]. It has

been well applied in the fields of traffic forecasting [17], knowledge graph [18] and social network [19], and has broad application prospects for problems that can be described by graph structures.

A. Graph Attention Network

The basic idea of graph network is that the properties of a node in the network are determined by its own properties and the properties of its neighbor nodes. In Graph Attention Network (GAT), the attention mechanism is used to assign different weights among different nodes. In the adjacent zone indoor temperature prediction problem, different neighbors have different temperature differences from the central node. This will lead to different intensity of heat transfer process in different walls and different effects on the central node indoor temperature. Therefore, it is suitable to use an attention mechanism to represent the different impact weights of different neighbors.

Besides, according to [20], the algorithm of GAT has lower complexity, better robustness, and better performance as it does not require the information of the whole graph. The procedure for calculating the attention coefficients between node i and node j in GAT is shown below [20]:

$$\alpha_{ij} = \frac{\exp\left(\text{LeakyReLU}\left(\vec{a}^T \left[\vec{W}\vec{h}_i \parallel \vec{W}\vec{h}_j\right]\right)\right)}{\sum_{k \in N_i} \exp\left(\text{LeakyReLU}\left(\vec{a}^T \left[\vec{W}\vec{h}_i \parallel \vec{W}\vec{h}_k\right]\right)\right)} \quad (2)$$

$$\vec{h}'_i = \sigma\left(\sum_{j \in N_i} \alpha_{ij} \vec{W}\vec{h}_j\right) \quad (3)$$

where LeakyReLU and σ are all activation function, \parallel denotes concatenation, \vec{h}_i is the feature of node i in the input data of the graph attention layer, and \vec{h}'_i is the feature of node i in the output data of the graph attention layer. \vec{a} and \vec{W} are learnable weight vector and learnable weight matrix respectively.

B. Gated Recurrent Unit

In reality, temperature changes gradually, so temporal features are especially important in temperature data. In order to prevent the loss of temporal features of temperature data in GAT, this study introduces Gated Recurrent Unit (GRU) [21] to realize the memory of temperature features at historical moments to help improve the overall prediction accuracy. GRU is a type of recurrent neural network. Similar to long and short term memory (LSTM), GRU is used to solve the long term memory and gradient disappearance problem in temporal data.

The internal structure of the GRU consists of two gates, the reset gate and the update gate, the gating states r^t and z^t can be formulated as follows:

$$r^t = \sigma(\mathbf{W}_r x^t + \mathbf{U}_r h^{t-1}) \quad (4)$$

$$z^t = \sigma(\mathbf{W}_z x^t + \mathbf{U}_z h^{t-1}) \quad (5)$$

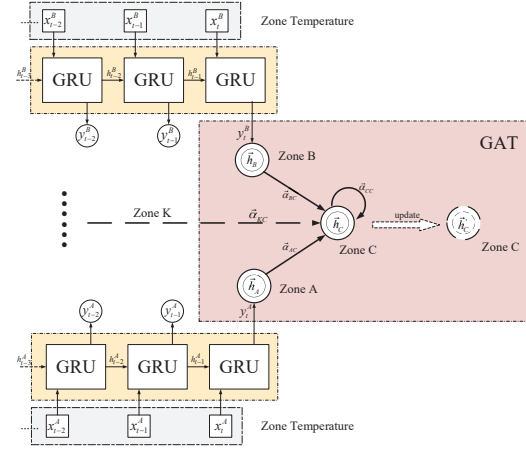


Fig. 3. Overall framework of the model.

where x^t is the input at time t , \mathbf{W} and \mathbf{U} are weight matrixes. h^{t-1} is the hidden state at time $t-1$. After getting the gating states, the hidden state update process is as follows:

$$h^t = (1 - z^t)h^{t-1} + z^t (\tanh(\mathbf{W}x^t + r^t \odot (\mathbf{U}h^{t-1}))) \quad (6)$$

where h^t is the hidden state at time t . $\tanh()$ is the activation function, \odot represents the Hadamard product. Unlike LSTM, GRU has fewer gates and therefore fewer parameters. Considering the computational power and time cost of hardware, GRU is easier to train and can improve the training efficiency while obtaining comparable experimental results [21]. Therefore, GRU replaces the commonly used LSTM prediction method in the following to compare the results.

In this paper, we combine the above two methods as follows. **Step 1**, we use a four-layer GRU network to extract and memorize the temporal features between adjacent moments of indoor temperature. The four-layer structure consists of three hidden layers and one fully connected layer. The input to the GRU network is the historical time-series temperature data for each zone, and the output of the GRU network is the predicted value of each independent zone. **Step 2**, the predicted value is fed into a GAT network with a two-layer structure as the input data. And the final temperature prediction is output after each zone completes the integration of the impact on neighboring zones based on the physical connection information.

In Step 1, the predicted values can capture the past temperature variation patterns, but the effect of heat transfer process in adjacent zones is not considered. In other words, it considers the temporal features of the temperature data without considering its spatial features. In Step 2, the attention mechanism assigns different influence weights to different neighbors to represent different intensity of heat transfer process, implying that the spatial features are also considered to obtain more accurate temperature predictions. The entire model framework is shown in Fig.3.

IV. CASE STUDY

A. Model and Data Generation

Consider the indoor temperature prediction problem for 5 adjacent zones in Xi'an, Shaanxi Province, with the layout

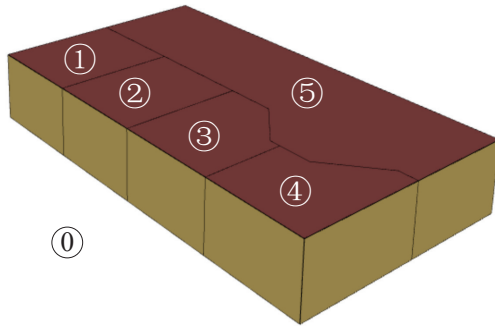


Fig. 4. Layout of the simulated zones.

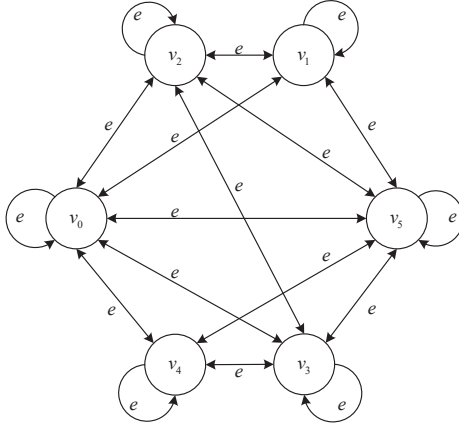


Fig. 5. Directed graph based on simulated zones.

shown in the Fig.4. The above adjacent zones are abstracted into a directed graph structure as the physical part of predict model, which as shown in Fig.5. The input data of the predict model is the time-series temperature data generated by EnergyPlus. And the final output data are the one-step ahead and multi-step ahead predicted temperatures in each zone. Note that in order to study the interaction of heat transfer processes between zones, no air conditioning system is designed in this case, and there are no people or equipment in the zones.

EnergyPlus is a building simulation software that can output time-by-time indoor and outdoor temperature by inputting building envelope information and weather information of geographical location. In this paper, the building envelope information is modified in the Input Data File (IDF) file based on the actual layout, and the weather information is derived from the China Standard Weather Data (CSWD) for a typical meteorological year.

Before generating simulated temperature data, the temporal granularity of the output data needs to be determined. Since heat transfer in walls is a relatively slow process, if the data granularity is too small, it will add a large amount of inefficient data and subsequently increase the model training cost; while if the data granularity is too large, the dynamic evolution of heat transfer in walls cannot be accurately and meticulously expressed, leading to a decrease in prediction accuracy. Therefore, when choosing the temporal granularity, there is a trade-off between accuracy and

complexity. In order to investigate the effect of different temporal granularity of data on the prediction results, we conducted a sensitivity analysis of the temporal granularity of temperature data. In this paper, we choose to record temperature data every 10 minutes, every 30 minutes and every 60 minutes, respectively, then we have 52560, 17520 and 8760 temperature data for the whole year, respectively.

Using the above zone model with weather files, indoor temperature time series data can be generated by EnergyPlus as the proposed model's dataset for off-line training and model testing. The first 80% of the dataset are used for training data and the remaining data are reserved for testing. In order to eliminate the effect of data size on the results, the temperature data obtained from the simulation need to be normalized before using them, and in this paper, the data are all normalized to $[-1, 1]$.

B. Analysis of results

This paper uses Pytorch and Deep Graph Library to build the graph neural network. The graph neural network was trained using Adam optimizer and equal interval adjustment of learning rate. GRU neural network with 64, 256 and 512 hidden layer nodes was also built.

In order to compare the effectiveness of the proposed model, we set up three groups of experiments for comparison: the first group using only the GAT, the second group using only the GRU network, and the third group combining the use of GRU and GAT. The first group predicts temperature based on physical connection topology only, the second group predicts temperature based on temporal features from historical data only, and the third group combines a physics-based approach with a data-driven approach. The three groups of algorithms perform one-step ahead prediction and multi-step ahead prediction based on data with different temporal granularity, respectively. And in the multi-step ahead prediction, we set the number of steps to three. Comparative experimental results are shown in Fig.6.

In Fig.6, the curve GRU represents the prediction results considering the temporal features of temperature data, the curve GAT represents the prediction results considering only the physical connection relationship and heat transfer process, the curve GAT+GRU represents our proposed prediction model integrating the physical model and the data-driven method. It can be seen that the GAT, which only considers the physical connection and heat transfer process and ignores the temporal features of temperature, has relatively poor prediction accuracy in general, and the GRU model, which only considers the temporal features, has relatively good accuracy in one-step prediction, but deteriorates in multi-part prediction. The model combined GRU and GAT performances well compared to the other two models, especially showing the highest accuracy in multi-step predictions. Using the full test data, the performance of the prediction models is further evaluated by RMSE (Root Mean Square Error), MAE (Mean Absolute Error) and R^2 (R Squared), which are defined as:

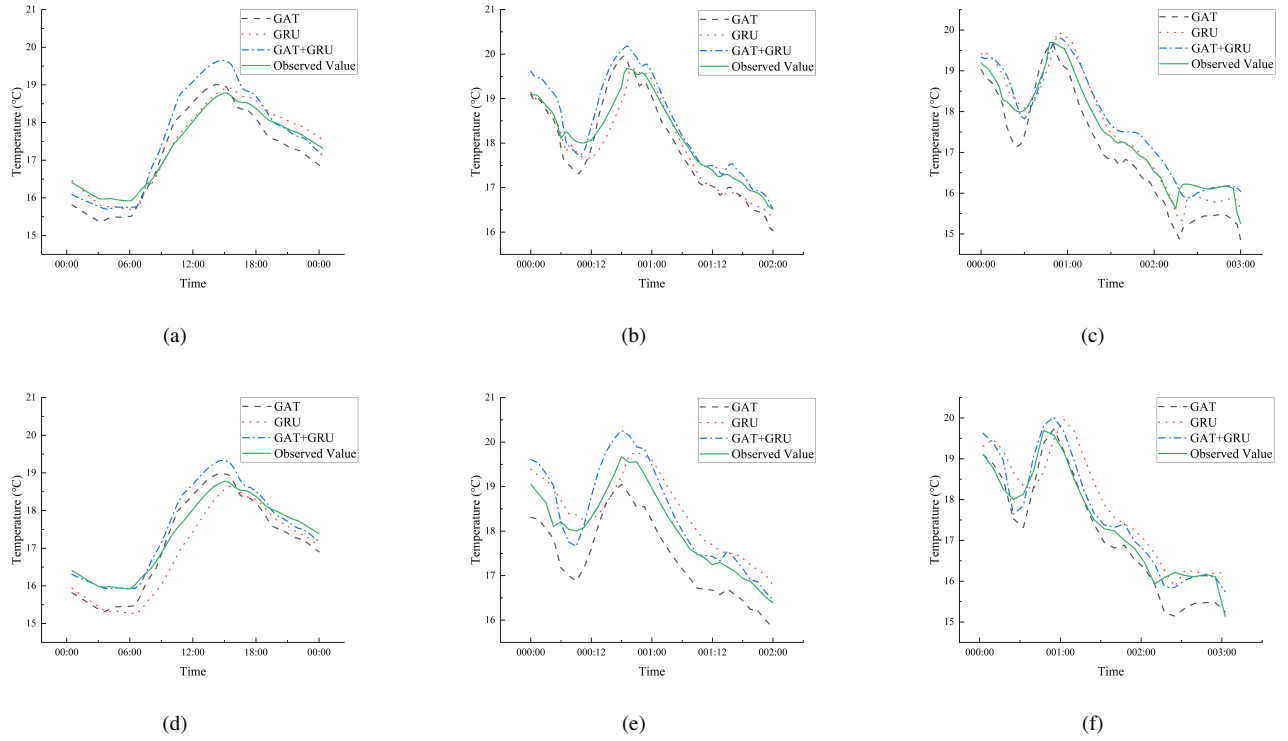


Fig. 6. (a) One-step ahead prediction based on 10-minute granularity data. (b) One-step ahead prediction based on 30-minute granularity data. (c) One-step ahead prediction based on 60-minute granularity data. (d) Multi-step ahead prediction based on 10-minute granularity data. (e) Multi-step ahead prediction based on 30-minute granularity data. (f) Multi-step ahead prediction based on 60-minute granularity data.

TABLE I
ONE-STEP AHEAD PREDICTION

Time interval	Models	RMSE	MAE	R^2
10 minutes	GAT	0.57	0.51	0.74
	GRU	0.49	0.38	0.94
	GRU+GAT	0.47	0.37	0.94
30 minutes	GAT	0.66	0.53	0.87
	GRU	0.63	0.56	0.88
	GRU+GAT	0.62	0.50	0.89
60 minutes	GAT	0.70	0.58	0.86
	GRU	0.66	0.52	0.87
	GRU+GAT	0.64	0.50	0.88

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - y_i^*)^2} \quad (7)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - y_i^*| \quad (8)$$

$$R^2 = 1 - \frac{\sum_i (y_i^* - y_i)^2}{\sum_i (\bar{y} - y_i)^2} \quad (9)$$

where n represents total data volume, y_i represents true value, y_i^* represents predicted value, and \bar{y} represents mean value. RMSE measures the deviation between the predicted value and the true value. MAE calculates the average of the absolute error, which can better reflect the actual situation of the predicted value error, and the smaller the value of the two means the better the model accuracy. R-squared is a measure of the model fit and takes values between [0, 1], the closer to 1 means the better the model accuracy. For one-step ahead prediction, the performance evaluation of the three models is shown in Table I and Table II shows the evaluation results of multi-step ahead prediction.

From Table I, we can have the following conclusions. First, using the data with same granularity, the model using a combination of GRU and GAT is more accurate in most cases. It shows that our prediction model integrates the advantages of both physics-based and data-driven methods to

achieve better results. Second, with same model, the smaller the granularity of the base data, the better the prediction accuracy. It shows that too large temporal granularity does lose the dynamic information of the heat transfer process. Similar conclusions are also available for the multi-step ahead prediction problem shown in Table II. Comparing Table I and Table II, it can be seen that one-step prediction has similar results to multi-step prediction, which is due to the absence of random factors such as personnel and equipment in the numerical tests. Therefore, the algorithm can better capture the pattern of temperature changes over a long period of time. However, one-step ahead prediction under the same conditions usually has smaller RMSE and MAE compared to multi-step prediction, implying a higher prediction accuracy.

In all cases, our combined model predicts the most accu-

TABLE II
MULTI-STEP AHEAD PREDICTION

Time interval	Models	RMSE	MAE	R^2
10 minutes	GAT	0.74	0.64	0.84
	GRU	0.59	0.41	0.90
	GRU+GAT	0.57	0.40	0.90
30 minutes	GAT	0.79	0.68	0.76
	GRU	0.63	0.49	0.88
	GRU+GAT	0.63	0.50	0.89
60 minutes	GAT	0.81	0.69	0.81
	GRU	0.68	0.56	0.69
	GRU+GAT	0.66	0.54	0.71

rately when using 10-minute granularity data, with an RMSE of 0.47, MAE of 0.37, and R^2 of 0.94. In addition, we can see that the model combining the use of GRU with GAT has stable accuracy on data with different temporal granularity, indicating that the combination of physics-based and data-driven methods increases the robustness of prediction to some extent.

V. CONCLUSIONS

Prediction of indoor temperature is important for comfort control, security monitoring, energy consumption optimization, and many other aspects. In this paper, we construct a graph network based on actual layouts to describe the interactions of temperature among the multiple zones. And in order to deal with the non-linearity of the heat transfer process, we use graph attention network and gated recurrent unit to predict multi-zone indoor temperatures. Using EnergyPlus simulation data as the dataset, the validity of the model prediction is verified. We also compare the differences in prediction results between using only the GAT network, only the GRU network, and a combination of the two networks. The results illustrate that our proposed novel multi-zone indoor temperature prediction model has good prediction accuracy and prediction stability. The combined use of the physics-based and data-driven methods outperforms the purely physics-based and purely data-driven methods.

In future work, further consideration will be given to describing the effects of uncertainty in zone occupancy and equipment on zone temperature. Then using the predicted temperature data as the base state for optimal control of equipments, such as air conditions and lights, can further reduce energy consumption and improve energy efficiency while ensuring the comfort of personnel.

ACKNOWLEDGMENT

The authors would like to thank Yuanjun Shen for his valuable advice on EnergyPlus, and Xiaopeng Li for his valuable suggestions.

REFERENCES

- [1] Y. Zhou, Y. Su, Z. Xu, X. Wang, J. Wu, and X. Guan, "A hybrid physics-based/data-driven model for personalized dynamic thermal comfort in ordinary office environment," *Energy and Buildings*, p. 110790, 2021.
- [2] Z. Afroz, T. Urmee, G. Shafullah, and G. Higgins, "Real-time prediction model for indoor temperature in a commercial building," *Applied energy*, vol. 231, pp. 29–53, 2018.
- [3] Z. Chen, R. Xiong, J. Lu, and X. Li, "Temperature rise prediction of lithium-ion battery suffering external short circuit for all-climate electric vehicles application," *Applied energy*, vol. 213, pp. 375–383, 2018.
- [4] C. Ren and S.-J. Cao, "Development and application of linear ventilation and temperature models for indoor environmental prediction and hvac systems control," *Sustainable Cities and Society*, vol. 51, p. 101673, 2019.
- [5] Sullivan, Royer, Stéphane, Thil, Thierry, Talbert, Monique, and Polit, "A procedure for modeling buildings and their thermal zones using co-simulation and system identification - sciencedirect," *Energy and Buildings*, vol. 78, no. 1, pp. 231–237, 2014.
- [6] M. Miguel, W. N. Hien, I. Marcel, H. D. J. Chung, H. Yueer, Y. Zhongqi, D. Ji-Yu, S. V. Raghavan, and N. N. Son, "A physically-based model of interactions between a building and its outdoor conditions at the urban microscale," *Energy and Buildings*, p. 110788, 2021.
- [7] D. Mazzeo, N. Matera, C. Cornaro, G. Oliveti, P. Romagnoni, and L. De Santoli, "Energyplus, ida ice and trnsys predictive simulation accuracy for building thermal behaviour evaluation by using an experimental campaign in solar test boxes with and without a pcm module," *Energy and Buildings*, vol. 212, p. 109812, 2020.
- [8] Y. Li, J. Castiglione, R. Astroza, and Y. Chen, "Real-time thermal dynamic analysis of a house using rc models and joint state-parameter estimation," *Building and Environment*, p. 107184, 2020.
- [9] Q. Xie, J.-Q. Ni, J. Bao, and Z. Su, "A thermal environmental model for indoor air temperature prediction and energy consumption in pig building," *Building and Environment*, vol. 161, p. 106238, 2019.
- [10] S. Kumar, S. K. Pal, and R. Singh, "A novel hybrid model based on particle swarm optimisation and extreme learning machine for short-term temperature prediction using ambient sensors," *Sustainable Cities and Society*, vol. 49, p. 101601, 2019.
- [11] M. Gustin, R. S. Mcleod, and K. J. Lomas, "Forecasting indoor temperatures during heatwaves using time series models," *Building and Environment*, vol. 143, no. OCT., pp. 727–739, 2018.
- [12] C. Xu, H. Chen, J. Wang, Y. Guo, and Y. Yuan, "Improving prediction performance for indoor temperature in public buildings based on a novel deep learning method," *Building and Environment*, vol. 148, pp. 128–135, 2019.
- [13] R. G. Martinez, A. Riverola, and D. Chemisana, "Disaggregation process for dynamic multidimensional heat flux in building simulation," *Energy and Buildings*, vol. 148, pp. 298–310, 2017.
- [14] Y. Wang, Z. Zhao, Y. Liu, D. Wang, C. Ma, and J. Liu, "Comprehensive correction of thermal conductivity of moist porous building materials with static moisture distribution and moisture transfer," *Energy*, vol. 176, pp. 103–118, 2019.
- [15] M. Gori, G. Monfardini, and F. Scarselli, "A new model for learning in graph domains," in *Proceedings. 2005 IEEE International Joint Conference on Neural Networks, 2005.*, vol. 2, pp. 729–734, IEEE, 2005.
- [16] F. Scarselli, M. Gori, A. C. Tsoi, M. Hagenbuchner, and G. Monfardini, "The graph neural network model," *IEEE transactions on neural networks*, vol. 20, no. 1, pp. 61–80, 2008.
- [17] Z. Cui, K. Henrickson, R. Ke, and Y. Wang, "Traffic graph convolutional recurrent neural network: A deep learning framework for network-scale traffic learning and forecasting," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 11, pp. 4883–4894, 2019.
- [18] T. Dettmers, P. Minervini, P. Senetorp, and S. Riedel, "Convolutional 2d knowledge graph embeddings," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 32, 2018.
- [19] T. Bian, X. Xiao, T. Xu, P. Zhao, W. Huang, Y. Rong, and J. Huang, "Rumor detection on social media with bi-directional graph convolutional networks," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, pp. 549–556, 2020.
- [20] P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Liò, and Y. Bengio, "Graph attention networks," in *International Conference on Learning Representations*, 2018.
- [21] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," *arXiv preprint arXiv:1412.3555*, 2014.