東南大學
SOUTHEAST UNIVERSITY

Power forecasting techniques of renewable energy

Ph.D candidate. Nanyang Zhu

東南大學
SOUTHEAST UNIVERSITY

content

FACTS OF OUR GROUP

■ Research Group

Prof.
Kaifeng Zhang

Associate Prof.
Kun Yuan

Associate Prof.
Ying Wang

5 Ph.D candidates and 22 Master students

■ Research Focus：

- Unit commitment and economic dispatch
- Frequency control
- Power market
- Renewable energy

■ Publications

In the last five years, our group has published over 10 papers in prestigious international journals, (such as IEEE Transactions), and more than 20 papers in SCI-indexed journals.

■ Collaborating Institutions

NSFC: Completed 6 projects funded by the NSFC (as the principal investigator) and participated in 6 others.

SGCC: Undertaken more than 10 projects from the SGCC and related enterprises.

NARI: Undertaken and collaborated on numerous projects with NARI

FACTS OF OUR GROUP

東南大學 SOUTHEAST UNIVERSITY

■ About Me

Nanyang Zhu received his master's degree in Electronic Information from Jiangnan University, Jiangsu, China, in 2019. He is currently pursuing his Ph.D. at Southeast University, Nanjing, China. His research interests include artificial intelligence, new energy power prediction, and knowledge-based semantic information grid scheduling

Published (first author)

1. Zhu, N., Dai, Z., Wang, Y., # Zhang, K. (2023). A contrastive learning-based framework for wind power forecast. Expert Systems with Applications, 230, 120619.(IF: 8.5)
2. Zhu, N., Wang, Y., Yuan, K., Yan, J., Li, Y., # Zhang, K. (2024). GGNet: A novel graph structure for power forecasting in renewable power plants considering temporal lead-lag correlations. Applied Energy, 364, 123194.(IF: 11.2)
3. Zhu, N., Liu, X., Liu, Z., Hu, K., Wang, Y., Tan, J., ... # Guo, Y. (2018). Deep learning for smart agriculture: Concepts, tools, applications, and opportunities. International Journal of Agricultural and Biological Engineering, 11(4), 32-44. (IF: 2.4, Cited:230)
4. Zhu, N., Ji, X., Tan, J., Jiang, Y., # Guo, Y. (2021). Prediction of dissolved oxygen concentration in aquatic systems based on transfer learning. Computers and Electronics in Agriculture, 180, 105888.(IF: 8.3)
5. Zhu, N., Xia, Q., Tan, J., Jiang, Y., Xu, G., Chu, D., # Guo, Y. (2019). Model-based prediction of dissolved oxygen content in fish production. Transactions of the ASABE, 62(6), 1417-1425. (IF: 1.5)

■Self-developed software

Deep learning application framework (2021-Now)

Background: the aim is to make deep learning as convenient for researchers in various fields as using Word.

Functions: completing modules for data loading, training, various embedded tricks, and model construction.

significance: there are numerous and messy codes related to deep learning. Our framework can be flexible to any tasks, enabling researchers to easily meet their requirements

ChatSEU project at Southeast University(2022-Now)

Background: to build a ChatGPT with distinctive features of Southeast University.

Functions: besides the ChatGPT, we added search platform embedded semantic question-answering regrading the internal information of Southeast University.

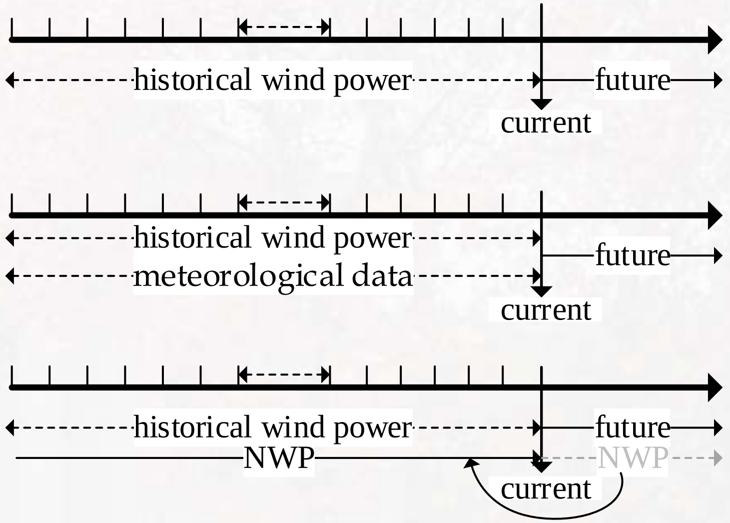
Topic 1: Unsupervised Learning-Based Power forecasting of Renewable Energy for Single Plant

■ Background

For power forecasting of Single renewable plant, there exists following forecasting paradigms.

1. Only using only historical power data
2. Using historical power data along with its related meteorological data (not work well)
3. Using historical power data and NWP data

Or, incorporating Empirical Mode Decomposition, etc.



Summary

all the models seek to enhance the capability of capturing feature representations from historical data, thereby gaining better insight into future wind power.

Challenge

How to accurately capture feature representations from power data

Our work Unsupervised Learning-Based Power forecasting of Renewable Energy for Single Plant[1]

Zhu, N., Dai, Z., Wang, Y., & Zhang, K. (2023). A contrastive learning-based framework for wind power forecast. Expert Systems with Applications, 230, 120619.

Topic 1: Unsupervised Learning-Based Power forecasting of Renewable Energy for Single Plant

■ Our work

Generally, traditional models optimize the parameters of the networks by future data using MSE Loss, called future data-guided optimization. This practice may lead to weak feature representations. Therefore, we propose a space distance-guided optimization based on contrastive learning.

1. Future data-guided optimization

Taking a real-world scenario as an example, S_1 and S_2 in Fig. 1 are quite close in trend and can be considered as "similar days", but the follow-up trends of them exist a larger difference, just shown by ***c*** and ***d*** in Fig. 1.

2. space distance-guided optimization

The space distance-focused optimization can control the feature distribution of similar sequences and non-similar sequences by optimizing the parameters of the network architecture, and its loss function is contrastive loss.

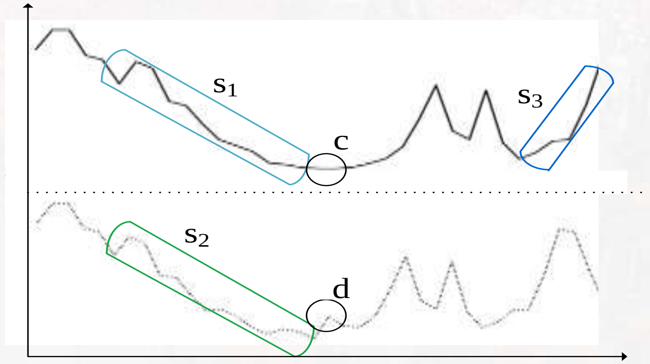
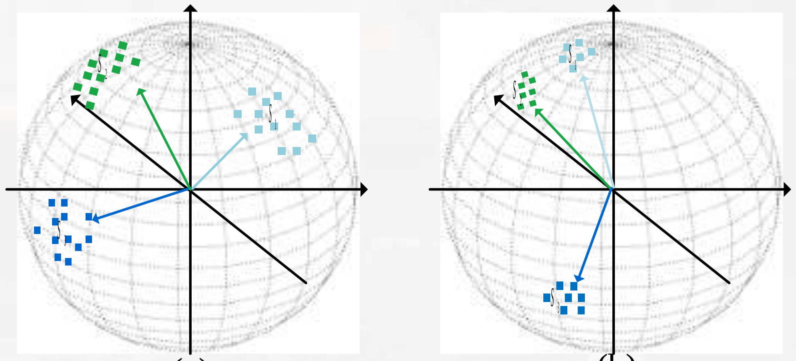


Fig. 1

For example, S_1, S_2, and S_3 in Fig. 1 are all actual wind power sequences (S_1 and S_2 are similar in trend; S_1 and S_3 are not).

If only using future data points (*c* and *d*) to optimize the parameter of the network architecture by future power-focused optimization, the distance of S_1, S_2, and S_3 in the feature space can be seen in Fig. 2(a). Consequently, to pull S_1 and S_2 closer and S_1 and S_3 farther in the feature space, just seen in Fig. 2(b), we use the space distance-focused optimization to further optimize the feature distance of similar and non-similar sequences on the basis of space distance-focused optimization.



(a) (b)

Fig. 2

■ **Our work**

How to do:

we propose a framework based on contrastive learning (CLFNet), consisting of a pre-training stage and a regression stage.

- Pre-training stage: pre-training stage includes feature extraction module, data construction module and loss optimization module.

1. **feature extraction module**

This module is to map the row batch elements into latent space based on any network architectures, including LSTM, TCN, Transformer, etc.

2. **data construction module**

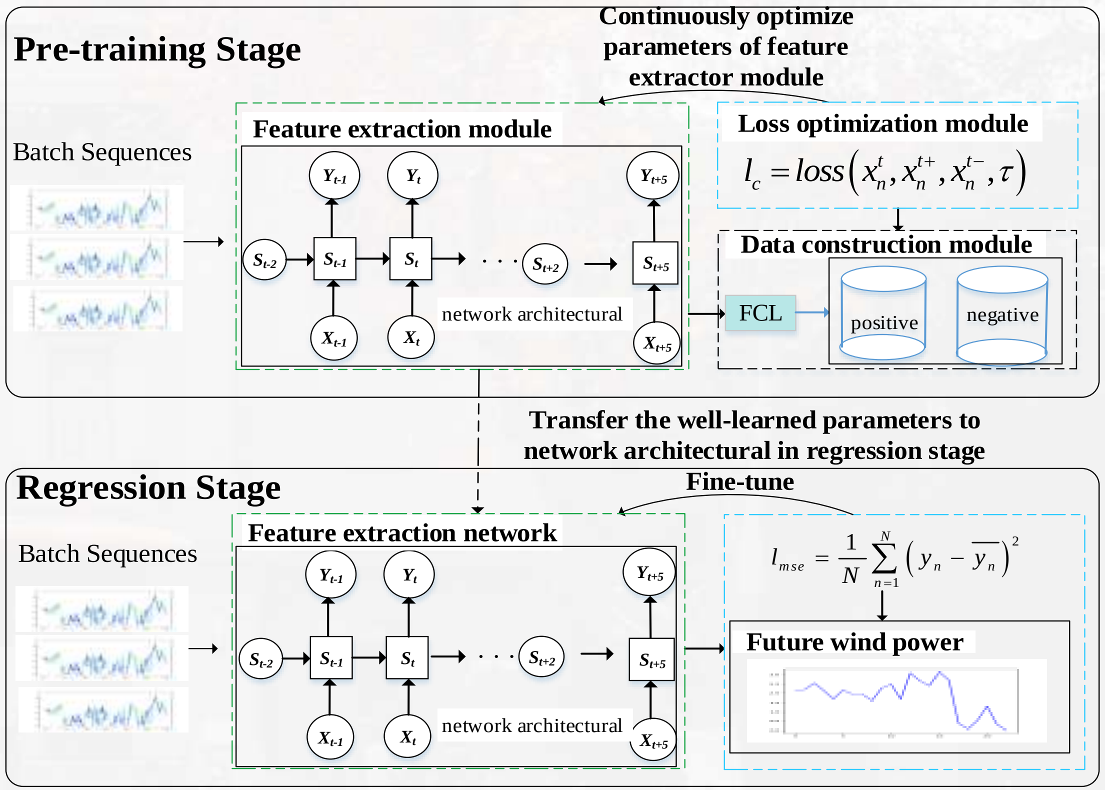
It generate two extra matrices ($\boldsymbol{F}^{+}$ and $\boldsymbol{F}^{-}$) based on $\boldsymbol{F}$ according to "Similar days". $\boldsymbol{F}^{+}$ are positive feature matrices that are similar to $\boldsymbol{F}$, and $\boldsymbol{F}^{-}$ is a collection of negative feature matrices that are not similar to $\boldsymbol{F}$;

3. **loss optimization module**

It uses contrastive loss (space distance-guided optimization) to continuously optimizing the parameters of the network architecture in the feature extraction module, enabling to reduce the space distance between positive pair ($\boldsymbol{F}$ and $\boldsymbol{F}^{+}$) and increase the space distance between negative pairs ($\boldsymbol{F}$ and $\boldsymbol{F}^{-}$)

- Regression stage

To take full use of the well-learned parameters of the network architecture in the pre-training stage, the regression stage needs to take these parameters as initial parameters of feature extraction network, and further fine-tune these parameters using mean squarer error (MSE) loss

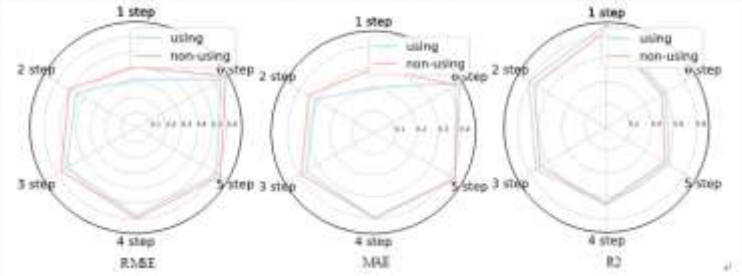


Topic 1: Unsupervised Learning-Based Power forecasting of Renewable Energy for Single Plant

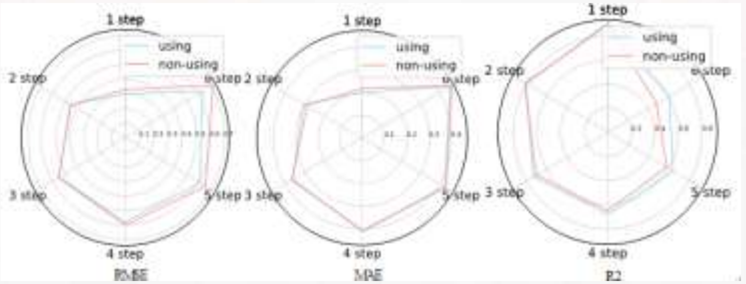
■ Our work

Case studies

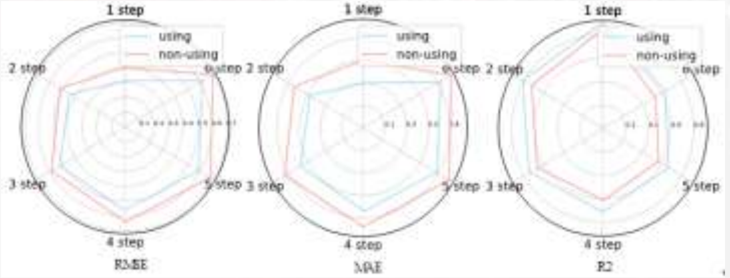
To show that the propose framework can be suitable for various network architectures and can outperform the initial basic network architectures, we apply it for various classic network architectures, including LSTM, CNN, and Transformer, named by CLFNet-LSTM, CLFNet-CNN, and CLFNet-Transformer, respectively



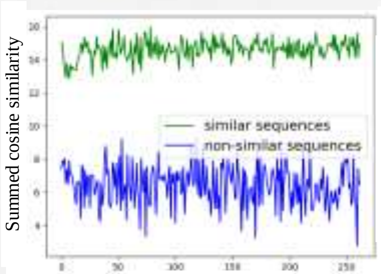
LSTM and CLFNet-LSTM



CNN and CLFNet-CNN

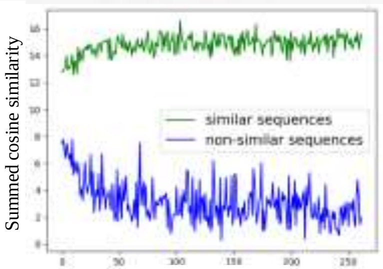


Transformer and CLFNet-Transformer



training steps

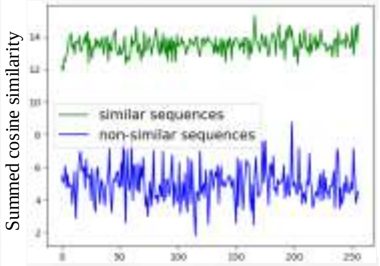
(1) with contrastive learning



training steps

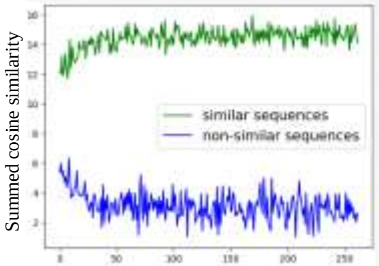
(2) without contrastive learning

change of space-distance for LSTM and CLFNet-LSTM



training process

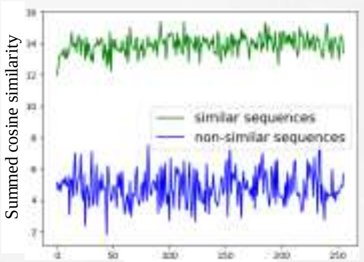
(1) with contrastive learning



training process

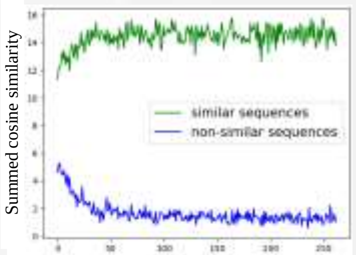
(2) without contrastive learning

change of space-distance for CNN and CLFNet-CNN



training process

(1) with contrastive learning



training process

(2) without contrastive learning

change of space-distance for Transformer and CLFNet-Transformer

Tips: more results regarding parameter sensitivity analysis can be found in our paper

■ Conclusion

This paper proposes a new wind power forecast framework based on contrastive learning, and fills the research gap of wind power forecast using self-supervised learning. The proposed framework consists of the pre-training stage and the regression stage, and its objective is to obtain a better prediction performance for wind power. The pre-training stage uses contrastive learning to extract more precise feature matrices for similar wind power sequences, and then the regression stage constructs a wind power forecast model on the basis of these well-learned feature matrices. Our framework is flexible and can support various network architectures, e.g., LSTM, CNN, Transformer, and more. We perform comprehensive comparative analyzes to demonstrate the effectiveness of our proposed framework. The results of Section 4.2 indicate that the proposed framework is widely suitable for various basis network, and the RMSE, MAE, and R2 can outperform the initial basis network architectures by a mean increase rate of 8.94%, 8.5%, and 8.5%, respectively, among 1-6 hour forecast steps. Additionally, for a network architecture with contrastive learning, the distance of similar sequences is closer in a feature space and the non-similar sequences is farther apart compare to its absence for the network architecture. The results of Section 4.3 indicate that the proposed framework is more powerful than other classic state-of-the-art schemes using different deep learning networks for different forecast steps. Sensitivity analysis of the temperature coefficient used in the contrastive loss can further present the effectiveness of the proposed framework.

Topic 2: Power forecasting of renewable energy clusters

東南大學 SOUTHEAST UNIVERSITY

■ Background

For power forecasting of renewable energy clusters, it is general to cast renewable power plants (RPPs) into a graph-like structure for capturing the correlations among them.

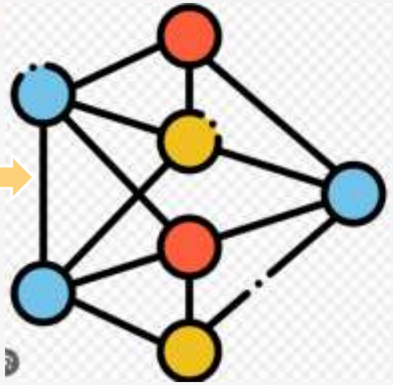
$G = (U, E)$

where U is the set of nodes, and E is the set of edges

cast

1

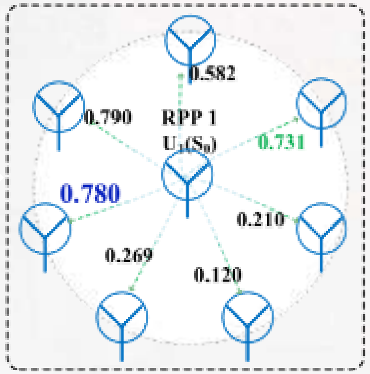
each plant to be the node in a graph



edge correlation

2

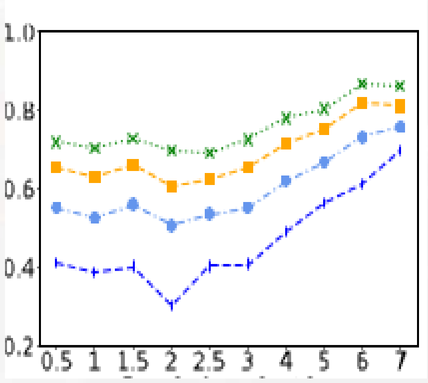
calculating the correlations of the edges in the graph



forecast

3

information aggregation for power forecasting using GNN-based models



Challenge

how to accurately capture the correlations among nodes in the graph structure

Topic 2: Power forecasting of renewable energy clusters

■Adjacency matrix

Generally, we describe the correlations among nodes using adjacency matrix.
there are three manners to calculate adjacency matrix in a graph:

1. physical distance: calculating the correlations among nodes based on geographical location (Geographic Coordinate System)
2. statistical distance: euclidean distance, correlation coefficient, pearson correlation distance, et al.
3. Learning a adjacency matrix in networks[1]: calculating the asymmetric matrices to be the graph adjacency matrix

Limitations In RPPs, due to different arrival time of atmospheric flow among RPPs, there exists lead-lag correlations. Nonetheless, all existing graph-based models ignore this point when constructing adjacency matrix

Our work Research on power forecasting techniques of renewable energy clusters considering Lead-Lag characteristics

[1] Wu, Z., Pan, S., Long, G., Jiang, J., Chang, X., & Zhang, C. (2020, August). Connecting the dots: Multivariate time series forecasting with graph neural networks. In Proceedings of the 26th ACM SIGKDD international conference on knowledge discovery & data mining (pp. 753-763).

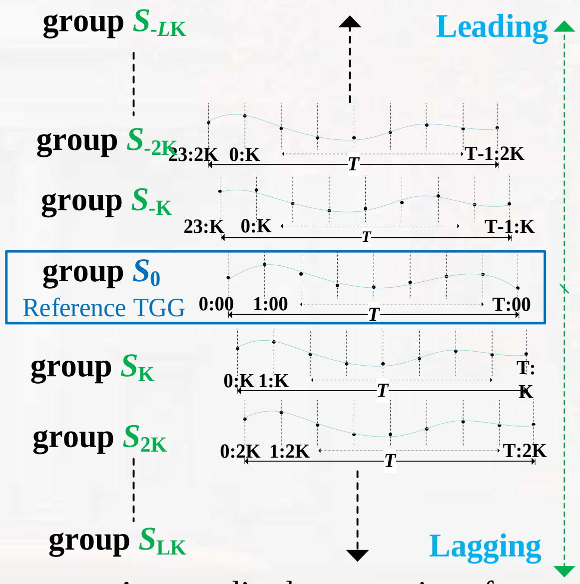
■Our Work[2]

Basic Definitions--- Multiple temporal granularity groups(TGGs)

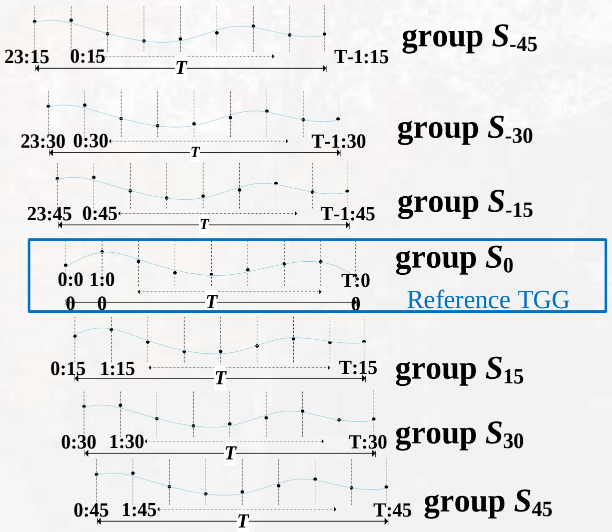
TGGs : power data in each group is distributed across distinct seconds or minutes in an hour

Reference TGG and its neighboring TGGs: **the reference TGG** is characterized by time steps aligned with the hour; **the neighboring TGGs** are positioned forward or backward by a multiple of L in the same hour with the reference TGG.

We can seen that there is lead-lag characteristics between the reference TGG and its **neighboring TGGs**



A generalized presentation of TGGs with resolution K



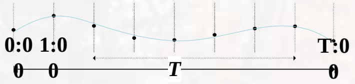
An example of TGGs with 15-minute temporal resolution and 45-minute lead-lag magnitude

[2] Zhu, N., Wang, Y., Yuan, K., Yan, J., Li, Y., & Zhang, K. (2024). GGNet: A novel graph structure for power forecasting in renewable power plants considering temporal lead-lag correlations. Applied Energy, 364, 123194.

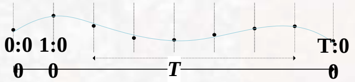
SOUTHEAST UNIVERSITY

■Our Work[2]

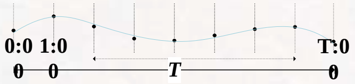
Proposed graph structure (3 plants to be example)

RPP$_1$: 

group S_0
Reference TGG

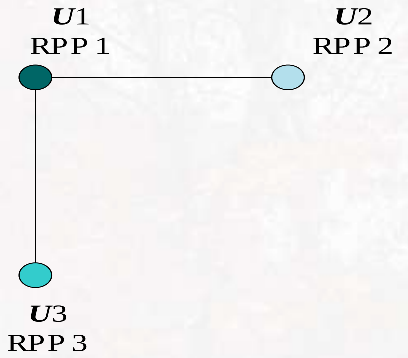
RPP$_2$: 

group S_0
Reference TGG

RPP$_3$: 

group S_0
Reference TGG

tradition graph structure

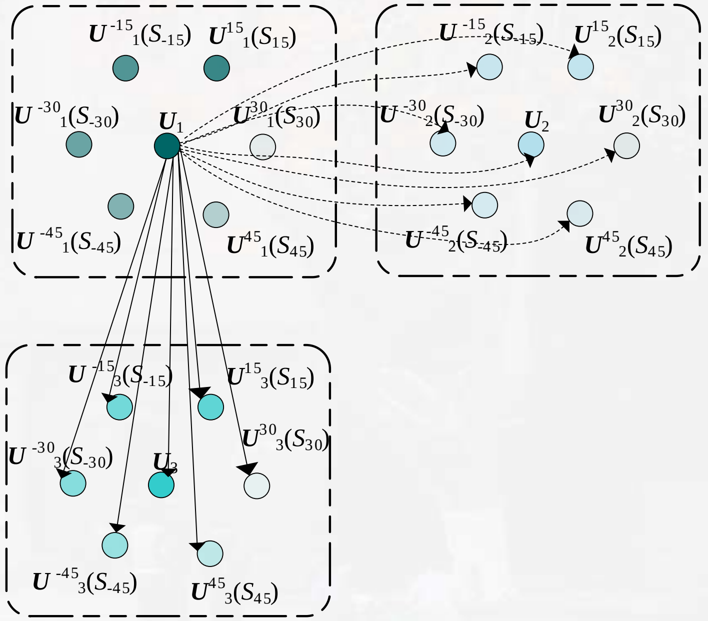


RPP$_1$ ----- **Reference TGG and its neighboring TGGs**

RPP$_2$ ----- **Reference TGG and its neighboring TGGs**

RPP$_3$ ----- **Reference TGG and its neighboring TGGs**

proposed graph structure



another question:

due to the dynamic nature of atmospheric flow, the magnitude and direction of the lead-lag correlation between any two RPPs are uncertainty

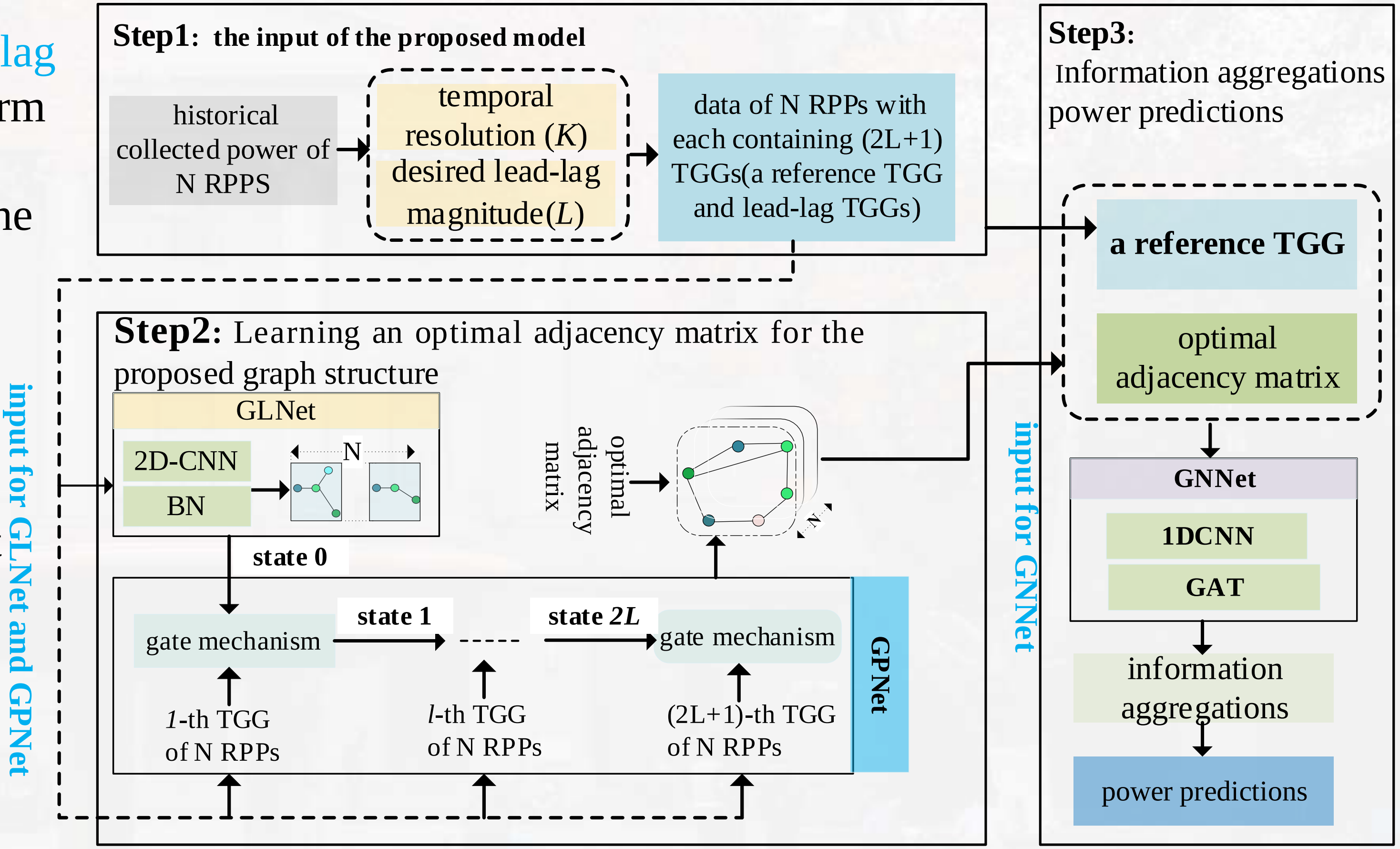
Topic 2: Power forecasting of renewable energy clusters

■Our Work[2]

how to learn adjacency matrix for the proposed graph structure: GGNet

GLNet: GLNet models the uncertainty of the lead-lag correlations with magnitudes and directions in a form of dynamic feature matrix.
GPNet: GPNet uses a gate mechanism to update the weighted coefficients of the GLNet in a iteration process, and obtain final adjacency matrix

GNNet: GNNet use GNN-based models to conduct information aggregation between the adjacency matrix and the data of RPPs with the reference TGG, and subsequently outputs power forecasting of RPPs.



More details regarding these module can be found in our paper

■Our Work[2]

Case studies------Datasets, Training settings

Datasets: the data used in this paper are from a region in North China, including 8 wind power plants and 9 photovoltaic power plants. Specifically, the data of each RPP presents 15-minute resolution and contains two years from January 1, 2019 to December 31, 2020. The data is split into a training dataset spanning from January 1, 2019 to October 30, 2020 and a testing data spanning from October 30, 2020 to December 31, 2020. Subsequently, TGGs with a reference TGG and its related TGGs are obtained based on the method in the Section 2.1 for both the training dataset and the testing dataset. Finally, the TGGs in the training dataset are used to train the proposed model for the proposed graph structure, and the TGGs in the testing dataset is used to evaluate the performance of the proposed model.

Training Settings: during the training process, the training dataset is used to train the proposed model using Pytorch on a 3080 Ti GPU with 64GB RAM, with optimized parameters learned by the backpropagation-based algorithm Adam optimizer (learning rate: 2e-5; batch size: 64; loss function: mean square error (MSE) loss) [58]. During the testing process, the testing dataset is used to evaluate the trained model on a 16 GB RAM-equipped CPU.

Topic 2: Power forecasting of renewable energy clusters

■Our Work[2]

Case studies------Comparing SoTA models

SoTAs: ARIMA, RF, RNN-GRU, LSTNet, DCRNN, StemGNN, MTGNN.

The $RMSE_N$, MAE_N, and $CORR_N$ of the proposed model and SoTA models for wind power plants

Step/h		ARIMA	RF	RNN-GRU	LSTNet	DCRNN	StemGNN	MTGNN	GGNet
1	$RMSE_N$	0.360	0.438	0.428	0.391	0.504	0.481	0.390	**0.300**
	MAE_N	0.330	0.300	0.280	0.220	0.325	0.302	0.221	**0.186**
	$CORR_N$	0.880	0.851	0.905	0.922	0.871	0.881	**0.923**	**0.934**
2	$RMSE_N$	0.673	0.580	0.661	0.626	0.768	0.723	0.601	**0.506**
	MAE_N	0.475	0.411	0.435	0.369	0.504	0.486	0.359	**0.324**
	$CORR_N$	0.764	0.733	0.773	0.800	0.702	0.733	**0.818**	0.805
3	$RMSE_N$	0.950	0.671	0.764	0.778	0.898	0.888	0.737	**0.606**
	MAE_N	0.648	0.492	0.532	0.484	0.602	0.593	0.456	**0.399**
	$CORR_N$	0.597	0.637	0.705	0.691	0.593	0.602	**0.727**	0.684
4	$RMSE_N$	1.232	0.733	0.956	0.887	1.004	0.987	0.846	**0.699**
	MAE_N	0.792	0.554	0.661	0.568	0.684	0.689	0.537	**0.493**
	$CORR_N$	0.404	0.532	0.537	0.598	0.491	0.505	**0.641**	0.590

The $RMSE_N$, MAE_N, and $CORR_N$ of the proposed model and SoTA for photovoltaic power plants

Step/h		ARIMA	RF	RNN-GRU	LSTNet	DCRNN	StemGNN	MTGNN	GGNet
1	$RMSE_N$	0.343	0.333	0.700	0.387	0.387	0.575	0.339	**0.227**
	MAE_N	0.241	0.210	0.519	0.251	0.270	0.353	0.164	**0.137**
	$CORR_N$	0.852	0.890	0.716	0.920	0.923	0.780	0.939	**0.951**
2	$RMSE_N$	0.501	0.534	0.960	0.623	0.563	0.725	0.504	**0.342**
	MAE_N	0.364	0.349	0.649	0.367	0.406	0.464	0.254	**0.214**
	$CORR_N$	0.664	0.684	0.448	0.826	0.840	0.682	0.867	**0.891**
3	$RMSE_N$	0.632	0.653	0.990	0.923	0.591	0.721	0.606	**0.396**
	MAE_N	0.567	0.460	0.662	0.652	0.418	0.487	0.309	**0.268**
	$CORR_N$	0.456	0.463	0.379	0.511	0.822	0.688	0.809	**0.856**
4	$RMSE_N$	0.637	0.734	1.196	1.06	0.700	0.864	0.678	**0.422**
	MAE_N	0.546	0.552	0.829	0.724	0.546	0.578	0.357	**0.294**
	$CORR_N$	0.204	0.180	0.215	0.347	0.707	0.595	0.763	**0.825**

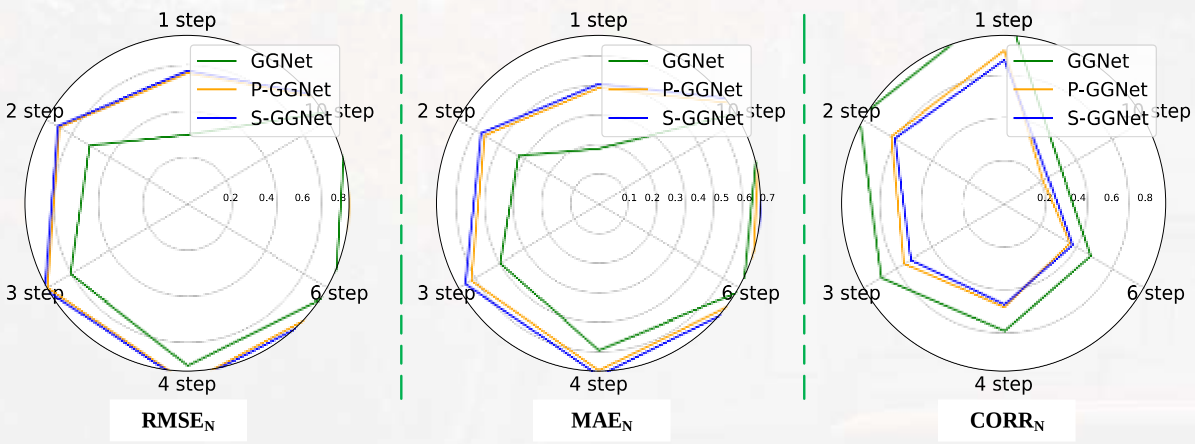
Topic 2: Power forecasting of renewable energy clusters

■Our Work[2]

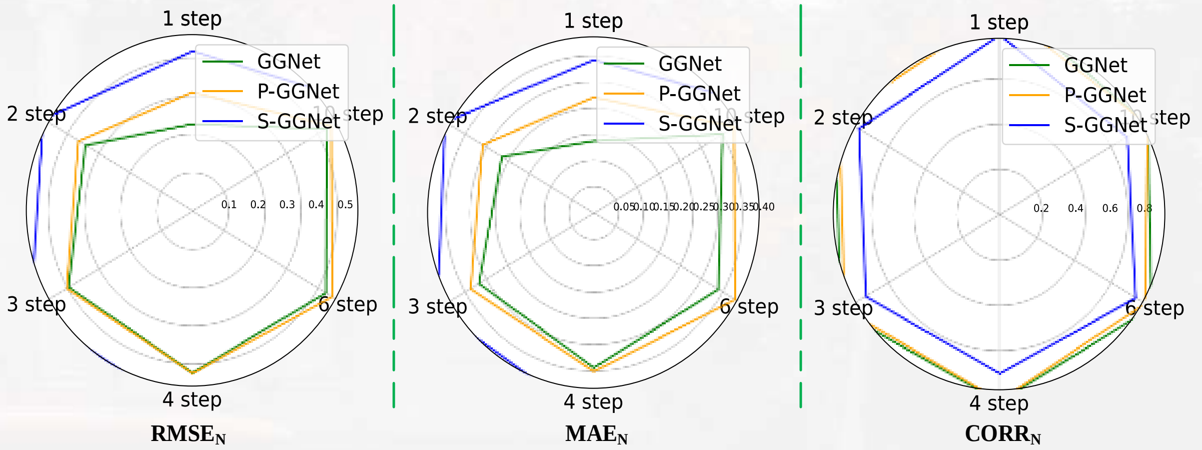
Case studies------Ablation studies

P-GGNet: using a predefined adjacency matrix used in DCRNN to replace the module GLNet;

S-GGNet: abandoning the GPNet module, and just using the feature matrix obtained from GLNet module to generate adjacency matrix.



The radar chart of RMSEN, MAEN, and CORRN of GGNet, P-GGNet, and S-GGNet for wind power plants.

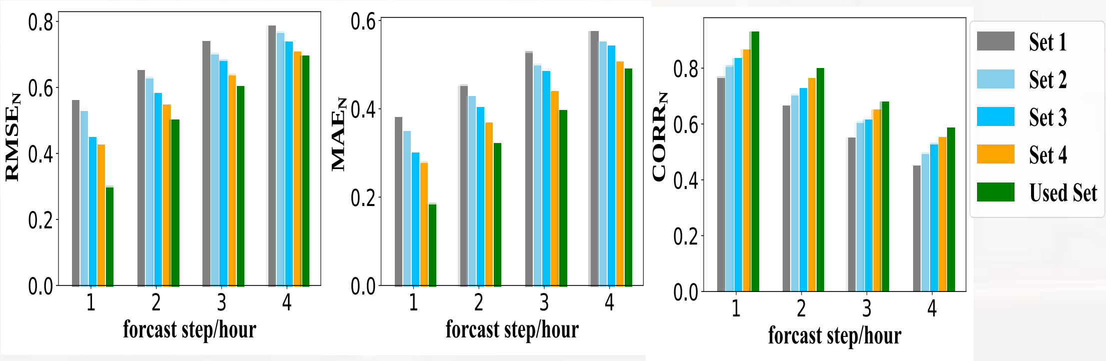


The radar chart of RMSEN, MAEN, and CORRN of GGNet, P-GGNet, and S-GGNet for photovoltaic power plants.

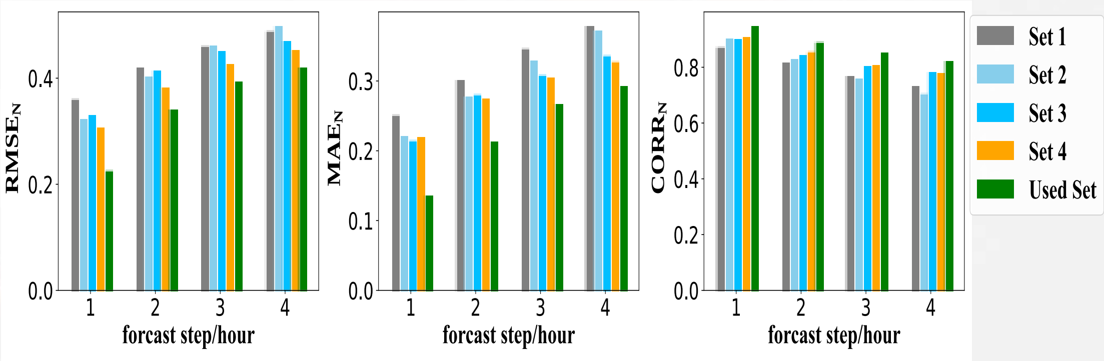
■Our Work[2]

Case studies------Parameter sensitivity analysis

Number of TGGs: the TGGs' Set (S-45, S-30, S-15, S0, S15, S30, S45) in Fig. 1 is used for the proposed model. Apart from the used Set, the experiment of different desired lead-lag magnitudes is performed, including 45-minute leading magnitude (number of TGGs=3), 45-minute lagging magnitude (number of TGGs =3), 15-minute lead-lag magnitude (number of TGGs =2), and 30-minute lead-lag magnitude (number of TGGs= 4), and the used Set (number of TGGs= 6).



The change of RMSEN, MAEN, and CORRN with different TGG combinations for wind power plants.



The change of RMSEN, MAEN, and CORRN with different TGG combinations for photovoltaic power plants.

Tips: more results regarding parameter sensitivity analysis can be found in our paper

■Our Work[2]

Conclusion

This paper proposes a novel dynamic graph structure using multiple TGGs for the power prediction of RPPs. The proposed graph structure can reflect the lead-lag characteristics among RPPs caused by the atmospheric flow, thereby obtaining better correlation representations among RPPs. The proposed model contains the GLNet, and the GPNet, and the GNNet. The former two modules with the GLNet and GPNet can obtain an optimal adjacency for representing the proposed graph structure, and the GNNet can aggregate information on RPPs for power predictions.

Multiple experiments are conducted on wind power plants and photovoltaic power plants to demonstrate the effectiveness of the proposed model. Specifically, the results of the section "Comparing state-of-the art end-to-end models" demonstrate the superior performance of the proposed model in 1-4 hour power prediction for RPPs, surpassing other state-of-the-art models in terms of RMSEN, MAEN, and CORRN. In wind power plants, the RMSEN and MAEN of the proposed model can obtain the best results with an average decreased of 22.92% and 13.18%, respectively, among 1 to 4 hours prediction steps. Particularly for 1 hour prediction step, the proposed model shows more significant improvement in the 1-hour prediction step, with 0.300 for RMSEN, 0.186 for MAEN, and 0.934 for CORRN, respectively; Similarly for photovoltaic power plants, the results of the proposed model can obtain superior results than the compared models, with average decrease by 48.95% for RMSEN, 18.75% for MAEN, and 8.56% for CORRN than the best results among the compared models, respectively. The results of the section "Ablation study" demonstrate that all the designed modules are of great significance for the proposed model. If one of the modules is replaced by other components, the performance of the proposed model shows a great decline, with about decreased of 23.20% for MAEN, 18.69% for RMSEN, and 23.45% for CORRN for all prediction steps. These results demonstrate that the proposed graph structure can ensure remarkable prediction accuracy for power of RPPs.

Thank you!