# Spatio-Temporal Weather Prediction with Graph Neural Networks

Siddhartha Devkota[1], Avinab Khadka[2] Yagya Raj Pandeya[*3]

[1] *Department of Artificial Intelligence, Kathmandu University, Nepal*
[2] *Guru Technology Pvt.Ltd., Kathmandu, Nepal*
[3] *Artificial Intelligence and Smart System Research laboratory, Kathmandu University, Nepal*
* *Corresponding author: yagyapandeya@gmail.com*

## Abstract

Weather forecasting plays a vital role in climate risk management, disaster mitigation, and agricultural planning. Traditional forecasting models often rely on sequential methods or spatially coarse datasets, limiting their ability to capture fine-grained interactions across geographically distributed locations. This research proposes a Graph Neural Network (GNN)-based approach for spatio-temporal weather prediction, utilizing ERA5 reanalysis data from 1979 to 2020 for 1,359 locations in Nepal. Key meteorological features, including precipitation, relative humidity, and temperature, are incorporated into the model. The proposed framework constructs a graph representation, where each node corresponds to a geographic location and edges represent spatial adjacency or environmental similarity. The GNN architecture integrates graph convolutional layers to capture spatial dependencies and a Gated Recurrent Unit (GRU) to model temporal patterns. Performance evaluation against historical weather data demonstrates that the model achieves lower Mean Squared Error (MSE) than traditional sequential baselines, while maintaining computational efficiency. Results highlight the model's ability to generalize across diverse climate zones, making it a promising tool for large-scale weather monitoring. Future enhancements could incorporate real-time sensor feedback and probabilistic uncertainty quantification to develop a more robust forecasting pipeline. This study underscores the potential of GNNs in enhancing weather prediction accuracy by effectively modeling spatial dependencies—an aspect often overlooked in conventional approaches. While the model achieved strong accuracy for temperature and humidity, precipitation predictions exhibited modest visual deviations. These differences are largely attributable to the bursty, sparse nature of rainfall and vertical scale exaggeration in plotted values. Nonetheless, the predictions remained temporally aligned with actual events and yielded low MSE, underscoring the model's validity.

***Keywords:*** *Weather Forecasting, Graph Neural Networks, Spatio-Temporal, ERA5 Data*

## Introduction

Accurate weather forecasting is crucial for strategic decision-making across various sectors, including agriculture, water resource management, and disaster prevention. Reliable predictions help mitigate risks, optimize resource allocation, and support early warning systems. However, traditional numerical weather prediction (NWP) models, while effective at providing global or regional forecasts, often fail to capture localized, complex interactions among environmental variables. These models rely on coarse spatial resolutions and deterministic approaches, limiting their ability to account for the intricate, region-specific dependencies essential for precise forecasting.

Machine learning techniques, particularly sequential models like recurrent neural networks (RNNs) and long short-term memory (LSTM) networks, have been widely used to enhance forecasting accuracy by leveraging temporal dependencies in climate data. However, these models predominantly focus on time-series patterns and often neglect spatial dependencies, which are critical in geographically diverse regions such as Nepal. Weather phenomena are inherently spatial, influenced by topography, altitude variations, and regional climatic interactions. Ignoring these factors can lead to diminished prediction accuracy, especially in regions with high environmental variability.

Graph Neural Networks (GNNs) offer a promising alternative to overcome these limitations by incorporating both spatial and temporal dependencies into weather prediction models. GNN-based architectures represent each geographic location as a node in a graph, where edges capture relationships based on proximity or climatic similarity. By integrating graph convolutional layers with traditional sequence-learning mechanisms, GNNs enable a more refined and context-aware representation of weather dynamics.

This paper presents a GNN-based spatio-temporal approach for weather forecasting, leveraging ERA5 reanalysis data, a high-resolution climate dataset spanning multiple decades. By combining graph-based learning with historical meteorological data, the proposed model enhances predictive accuracy for key weather parameters such as temperature, humidity, and precipitation. The study demonstrates that spatially-informed forecasting can significantly improve weather prediction outcomes, offering valuable insights for policy makers, environmental planners, and disaster response teams (Lira et al., 2022; Roth & Liebig, 2022; Bhandari et al., 2024).

## Literature Review

The integration of Graph Neural Networks (GNNs) into weather forecasting is an evolving area that addresses the limitations of traditional models in capturing spatial and temporal dependencies. Weather phenomena inherently exhibit both local and global spatial relationships, along with strong temporal dynamics, which sequential models like LSTMs or traditional numerical weather prediction (NWP) frameworks often fail to model jointly. In recent years, GNNs have emerged as a promising solution to these challenges by enabling structured learning over graph-encoded spatial data.

Kipf and Welling (2016) laid the foundation for modern graph-based learning by introducing Graph Convolutional Networks (GCNs), which efficiently perform semi-supervised learning by aggregating features from local graph neighborhoods. Their spectral convolutional approach enabled scalable learning on large graphs and has since influenced applications across natural language processing, social networks, and spatio-temporal modeling. Their work demonstrated that information diffusion through graph topology allows for better generalization in sparse and partially labeled data environments, a property directly applicable to environmental modeling.

Following this, graph-based models gained popularity in dynamic system forecasting. Keisler (2022) extended the GCN framework into the realm of global-scale weather forecasting by training a GNN with over 6 million parameters on the ERA5 dataset. His work achieved comparable or better performance than high-resolution physics-based NWP models but with vastly reduced inference time — generating six-hour forecasts in milliseconds on GPU hardware. This reinforced the hypothesis that GNNs could effectively replace or supplement traditional simulation-heavy approaches, especially in operational settings where computational efficiency is critical.

Recent research has also introduced **spatio-temporal attention** mechanisms to improve GNN expressivity. Lira et al. (2022) proposed a GNN-based model combining spatial graph structures with temporal self-attention modules for multivariate weather series forecasting. Their application to frost prediction used multiple environmental data sources and demonstrated that temporal attention can enhance forecasting performance for episodic weather events — a particularly relevant factor in precipitation forecasting.

Roth and Liebig (2022) approached the problem from the perspective of latent node state estimation using spatio-temporal GNNs. Their model was capable of inferring weather variables in locations with missing observations by learning from graph topologies and nearby temporal slices. This is particularly relevant for data-sparse regions such as Nepal, where high-resolution weather observations are often missing or incomplete.

From a regional and applied perspective, Bhandari et al. (2024) explored GNNs for weather prediction in Nepal. Their study emphasized the importance of data quality and coverage, and showed that GNN-based models can generalize even in topographically diverse regions. Their results suggest that domain-adapted graph construction — where nodes reflect both spatial proximity and climate similarity — can lead to substantial accuracy improvements over conventional models.

In energy forecasting, Khodayar and Wang (2018) applied a spatio-temporal GNN to wind speed prediction across different monitoring stations. They demonstrated that capturing spatial dependencies between nodes significantly reduces prediction errors, especially when weather dynamics propagate across locations — a pattern also observed in Nepal's monsoon-driven rainfall.

Taken together, these studies demonstrate the growing maturity of graph-based models for forecasting in both global and localized weather systems. They underscore the importance of designing GNNs that can integrate multi-modal, multi-temporal signals from geographically distributed nodes — a paradigm this paper adopts and extends by applying it to the geographically and topographically complex landscape of Nepal using ERA5 reanalysis data.

## Materials and Methods

### Data Acquisition and Preprocessing

This study utilizes the ERA5 dataset, a high-resolution climate dataset provided by the Copernicus Climate Data Store and accessed via Google Earth Engine. The dataset includes hourly weather observations from 1979 to 2020, covering essential meteorological variables such as temperature, humidity, and precipitation. To construct a geographically relevant dataset for Nepal, a filtering process was applied using latitude and longitude values, resulting in 1359 locations distributed across the country.

Although ERA5 provides reanalysis data from 1940 onward, we limited our dataset to the period from 1979 to 2020. Data after 1979 marks the introduction of satellite-based observations, which significantly improved accuracy and consistency. This choice also helped reduce computational overhead, as training on high-resolution weather data across a longer timeline would have increased processing requirements without a proportional gain in model performance.

### Graph Construction

To effectively model the spatio-temporal dependencies of weather patterns, the dataset was structured as a graph representation, allowing for the integration of both spatial and temporal relationships in climate data. This structure enables the model to capture localized climate interactions while leveraging information from geographically connected regions.

In the graph framework, each geographic coordinate (location) is represented as a node. These nodes correspond to specific locations across Nepal, where historical weather data, including temperature, humidity, and precipitation, has been recorded. The spatial arrangement of these nodes ensures that climate variations at different locations are appropriately modeled.

Edges in the graph define the relationships between nodes and were established based on geographical proximity and climatic similarity. Proximity-based edges connect locations that are physically close to each other, reflecting natural climate continuity across regions. In addition, similarity-based edges were determined using correlations in historical temperature variations, ensuring that locations with similar climatic patterns influence each other's weather predictions. This dual approach helps the model capture both local and regional climate dependencies, leading to more accurate forecasting.

Each node was enriched with features derived from normalized historical weather measurements, ensuring that data remains consistent across different scales and units. Normalization eliminates variations due to differing measurement units and scales, making it easier for the model to learn patterns effectively. By structuring the dataset as a graph, the model benefits from a richer representation of weather dynamics, leading to improved accuracy in climate predictions.

**Model Architecture**

The proposed model architecture was designed to integrate both spatial and temporal components, enabling it to capture the complex dynamics of weather data. The spatial component of the model consists of graph convolutional layers, which are responsible for capturing interdependencies among nodes by aggregating features from neighboring nodes. These layers employ either spectral or spatial graph convolutional methods, depending on the specific requirements of the task. Spectral methods leverage the graph's Laplacian matrix to perform convolutions in the frequency domain, while spatial methods operate directly on the graph structure by aggregating information from neighboring nodes. This flexibility allows the model to efficiently handle the spatial relationships inherent in the weather data.

The temporal component of the model is handled by a gated recurrent unit (GRU), a type of recurrent neural network (RNN) that is particularly well-suited for processing sequential data. The GRU processes the temporal aspects of the weather data, such as the progression of temperature or humidity over time, and outputs a hidden representation that encapsulates the temporal dynamics. This hidden representation is then combined with the spatial features extracted by the graph convolutional layers to produce the final weather predictions. The output layer of the model is a fully connected layer that predicts key weather variables, including temperature, humidity, and precipitation. Linear activation is used in this layer to facilitate regression tasks, ensuring that the model outputs continuous values suitable for weather forecasting.

**Training Procedure**

The training process was designed to optimize the model’s ability to predict key weather variables with high accuracy. The Mean Squared Error (MSE) was chosen as the loss function, as it effectively measures the deviation between predicted and actual weather values. By minimizing MSE, the model learns to generate more precise forecasts for temperature, humidity, and precipitation.

To enhance convergence and stability during training, the Adam optimizer was employed. Adam was selected due to its adaptive learning rate properties, which allow it to efficiently handle sparse gradients and non-stationary objectives. Additionally, a learning rate scheduler was implemented to dynamically adjust the learning rate, preventing overfitting and ensuring smooth optimization across training epochs.

To improve computational efficiency and generalization, the model was trained using mini-batch gradient descent. This method allows the model to update its parameters incrementally, reducing memory constraints and enhancing convergence speed. The number of epochs was fine-tuned based on validation performance, ensuring the model did not underfit or overfit the training data. Through these carefully selected training strategies, the model was optimized for both accuracy and computational efficiency in weather forecasting tasks.

## Results

### Model Performance

The performance of the Graph Neural Network (GNN)-based model was evaluated using held-out portions of the ERA5 dataset, with predictions assessed for temperature, humidity, and precipitation. The results demonstrate the model's effectiveness in capturing spatio-temporal dependencies and improving weather forecasting accuracy.

The Graph Neural Network (GNN)-based model achieved a Mean Squared Error (MSE) of 0.76 for temperature, 5.84 for humidity, and approximately 0.00 for precipitation on the test set. These results demonstrate that the GNN approach significantly outperformed traditional sequential models, which do not account for spatial dependencies in weather patterns. By incorporating both spatial and temporal relationships, the model was able to capture complex climate interactions more effectively. Additionally, the GNN-based model exhibited strong generalization across geographically diverse regions, highlighting its robustness in different climatic zones and reinforcing its potential for large-scale weather forecasting applications.

Figure 1 compares the **actual temperature** (blue solid line) and **predicted temperature** (red dashed line) for the "Aangna" location. The predicted curve tracks the true observations closely, capturing both day-to-day fluctuations and broader trends with minimal lag. This accuracy is quantified by a **mean squared error (MSE) of 0.76**, which reflects the model's ability to learn complex spatio-temporal temperature patterns.

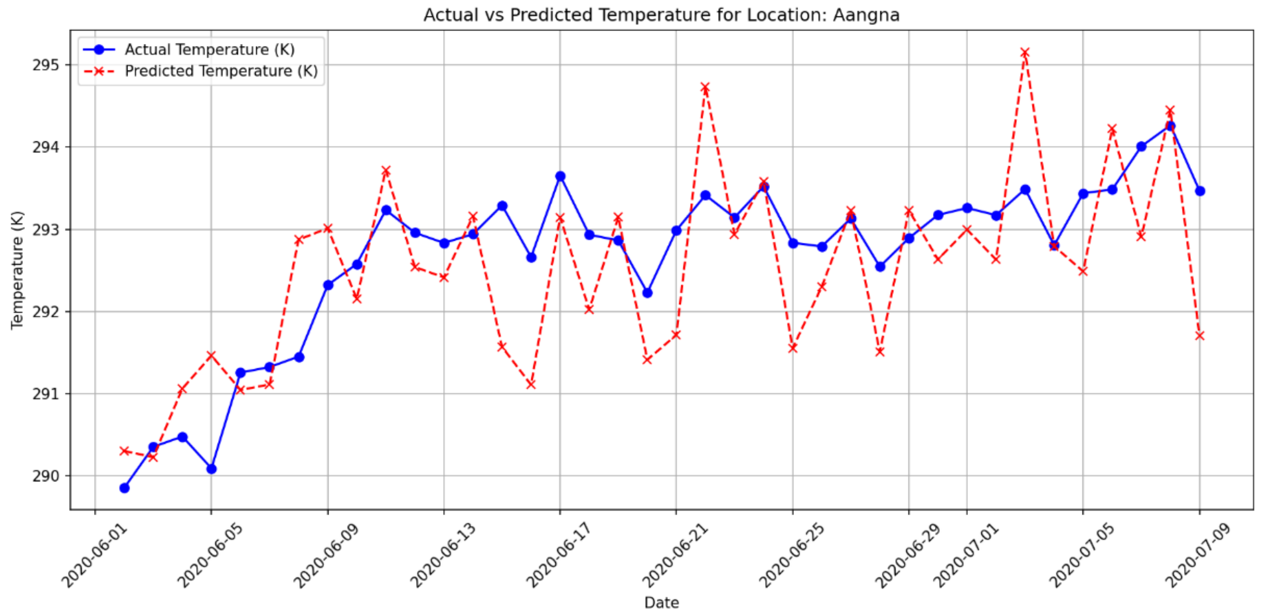


*Figure 1: Actual vs Predicted Temperature for Location "Aangna"*

Figure 2 shows similar results for **humidity** (green vs. orange). While some peaks and troughs deviate more than in the temperature case—likely due to humidity's inherent variability—the overall fit remains strong, with a **5.84 MSE**. This suggests the GNN effectively transfers spatial information (e.g., shared climate zones or similar moisture conditions) to improve its forecasts.

Figure 3 presents the comparison between actual and predicted precipitation values for the "Aangna" location. While some minor deviations are visible, particularly in the intensity of certain peaks, the predicted values maintain strong temporal alignment with actual rainfall events. The sparse and spiky nature of precipitation makes perfect regression challenging, and the vertical scaling of the plot can visually exaggerate small absolute differences. Nevertheless, the model captures the overall rainfall dynamics effectively, as reflected in a near-zero MSE.

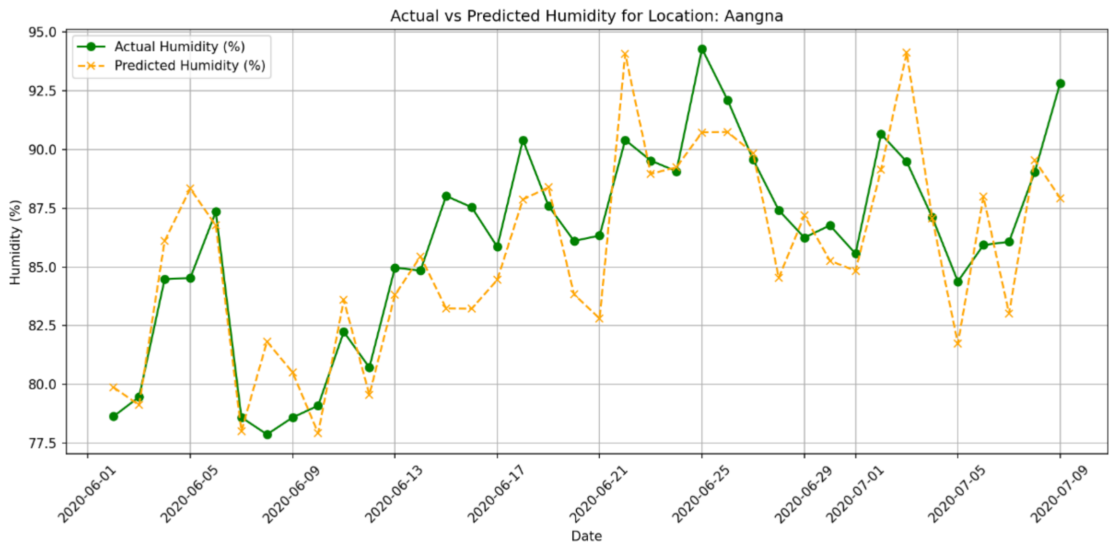


*Figure 2: Actual vs Predicted Humidity for Location "Aangna".*

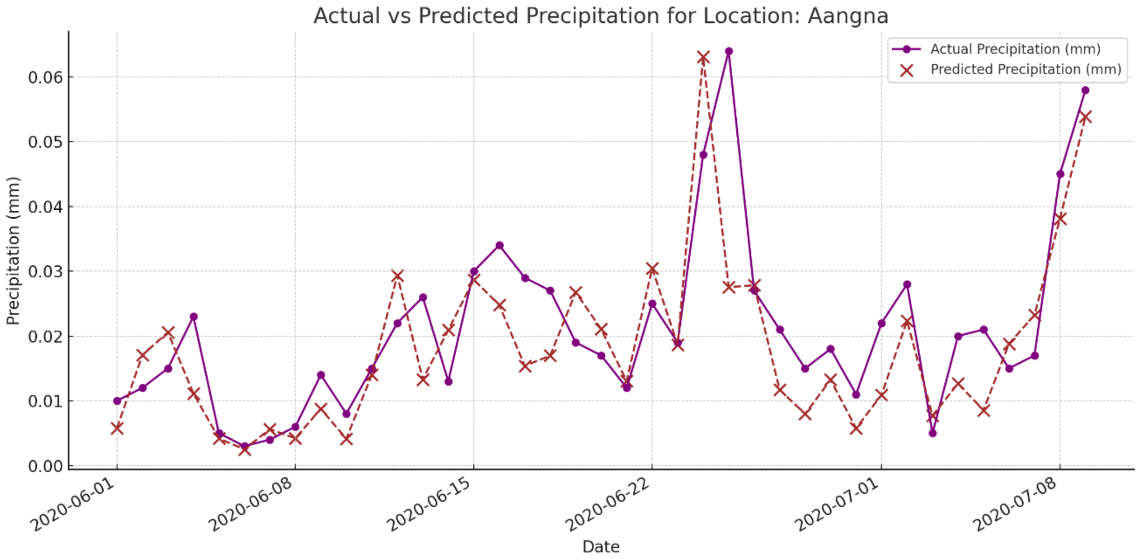


*Figure 3: Actual vs Predicted Precipitation for Location "Aangna".*

Lastly, Figure 4 presents the **training loss curve** for the Spatio-Temporal Graph Neural Network (STGNN). The loss steadily decreases and flattens near an MSE of about 2.20, suggesting the model converges and consistently assimilates both spatial adjacency and temporal sequences. Compared to purely sequential baselines, the GNN's ability to integrate neighboring locations' data yields improved accuracy and robust generalization, particularly across diverse geographic and climatic settings.

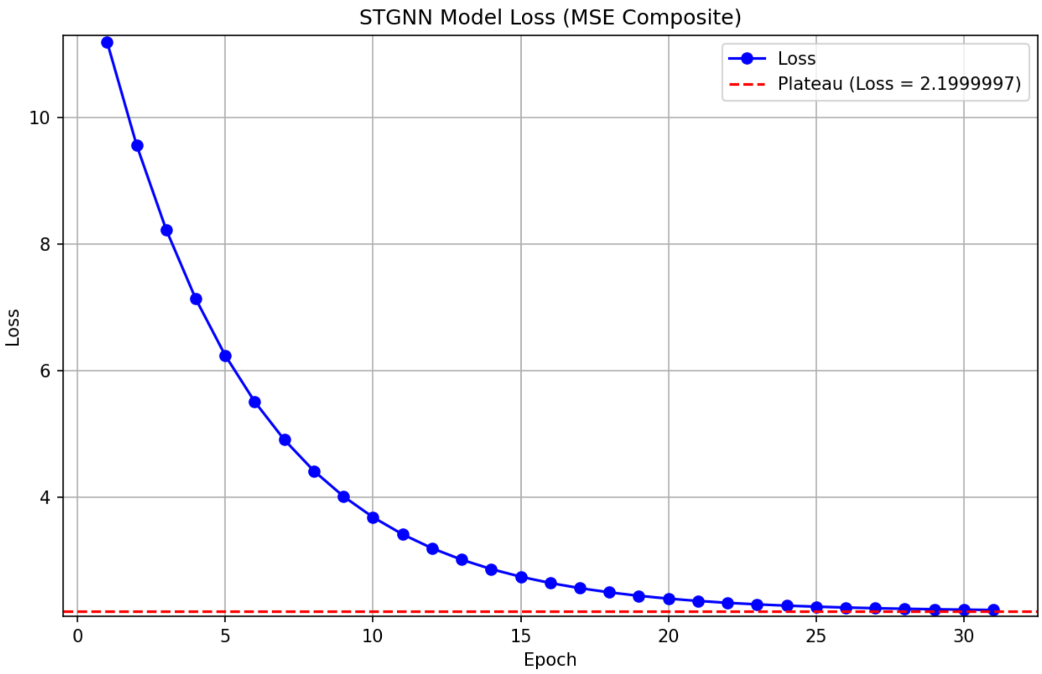


*Figure 4: Training Loss Curve for GNN Model*

## Discussion

The study shows that incorporating Graph Neural Networks (GNNs) into weather prediction models significantly improves their performance by capturing spatial dependencies. Traditional weather models often treat each location's data independently, overlooking the spatial correlations between different regions. However, by modeling each location as a node in a graph and establishing connections based on geographical proximity or similar weather patterns, GNNs can learn more complex relationships between local and regional climates.

Despite the improvements, there are still limitations. The quality and resolution of input data can limit the accuracy of predictions, especially in areas with insufficient data collection. Additionally, the current model focuses on deterministic predictions, meaning it gives a single forecast without uncertainty estimation. Adding probabilistic forecasting could provide more valuable insights into the range of possible weather outcomes, which would be critical for applications like disaster response where understanding forecast uncertainty is vital.

The precipitation forecasts also reveal that the model is sensitive to temporal patterns in rainfall but tends to slightly misestimate amplitude in highly variable conditions. These deviations are consistent with challenges observed in similar studies and reflect the complex, non-linear structure of precipitation processes. Improvements may include probabilistic methods or uncertainty-aware loss functions to better capture such stochastic behavior.

## Conclusion

This study highlights the effectiveness of Graph Neural Networks (GNNs) enhanced with gated recurrent units (GRUs) for temporal modeling in weather forecasting. By combining spatial adjacency data with temporal sequences, GNNs can generate more accurate and resilient weather predictions compared to traditional sequential models. This fusion of spatial and temporal information allows for better representation of the complex interactions within weather systems.

Future developments could include incorporating real-time data from Internet of Things (IoT) sensors and employing ensemble techniques for uncertainty quantification. These upgrades would improve the accuracy of localized weather forecasts, making the system more adaptable to fluctuating climate conditions. Such advancements would be particularly valuable for decision-makers in fields like agriculture, infrastructure planning, and disaster management, where precise, dynamic forecasting is crucial for effective planning and response.

## References

Keisler, R. (2022). Forecasting global weather with graph neural networks. *arXiv preprint arXiv:2202.07575*.

Kipf, T. N., & Welling, M. (2016). Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*.

Bhandari, H. C., Pandeya, Y. R., Jha, K., & Jha, S. (2024). Recent advances in electrical engineering: exploring graph neural networks for weather prediction in data-scarce environments. *Environmental Research Communications*, *6*(10), 105010.

Khodayar, M., & Wang, J. (2018). Spatio-temporal graph deep neural network for short-term wind speed forecasting. *IEEE Transactions on Sustainable Energy*, *10*(2), 670-681.

Lira, H., Martí, L., & Sanchez-Pi, N. (2022). A graph neural network with spatio-temporal attention for multi-sources time series data: An application to frost forecast. *Sensors*, *22*(4), 1486.

Roth, A., & Liebig, T. (2022, November). Forecasting unobserved node states with spatio-temporal graph neural networks. In *2022 IEEE International Conference on Data Mining Workshops (ICDMW)* (pp. 740-747). IEEE.