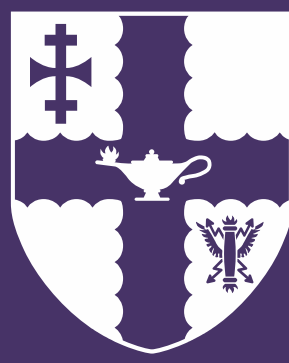


Spatiotemporal Forecasting of PM_{2.5}, PM₁₀ and NO₂ Using a Dynamic Directed Graph Neural Network

Niloufar Esfahanian, Miguel Martínez García, Sarah Dunnett



Loughborough University

Introduction

- Ambient air pollution is a major environmental and public-health concern, contributing to approximately 4.2 million premature deaths worldwide each year according to the World Health Organization [1].
- Key pollutants such as **PM_{2.5}**, **PM₁₀** and **NO₂** show substantial **spatial and temporal variability** due to changing emission sources, meteorological conditions and atmospheric transport.
- Fixed monitoring stations provide measurements only at discrete locations**, which limits the ability to capture pollutant relationships and interactions across the wider monitoring network [2].
- This study develops a **dynamic directed GCN-GRU model** for short-term multi-pollutant forecasting across 12 Beijing monitoring stations [3].

Methodology

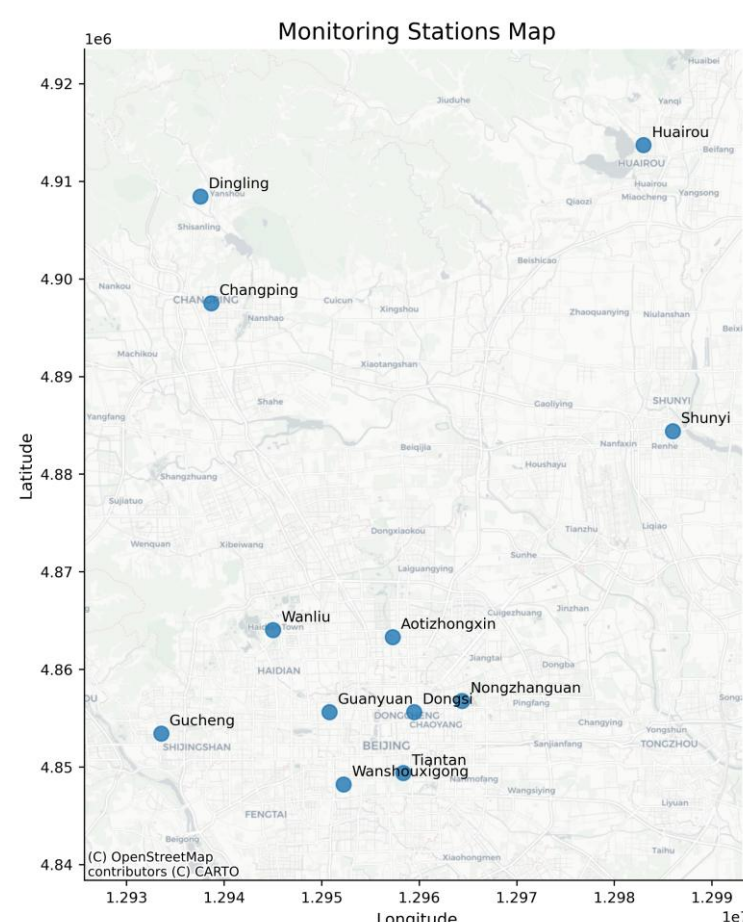
I. Dynamic Graph Construction

Why Graphs?

- Monitoring stations are not independent.
- Pollutant concentrations measured at one station may be related to conditions at nearby or upwind stations because pollutants are transported through the atmosphere.

$G(V, E, A_t)$
 V : 12 monitoring stations
 E : directed candidate connections between stations
 A_t : hourly weighted adjacency matrix

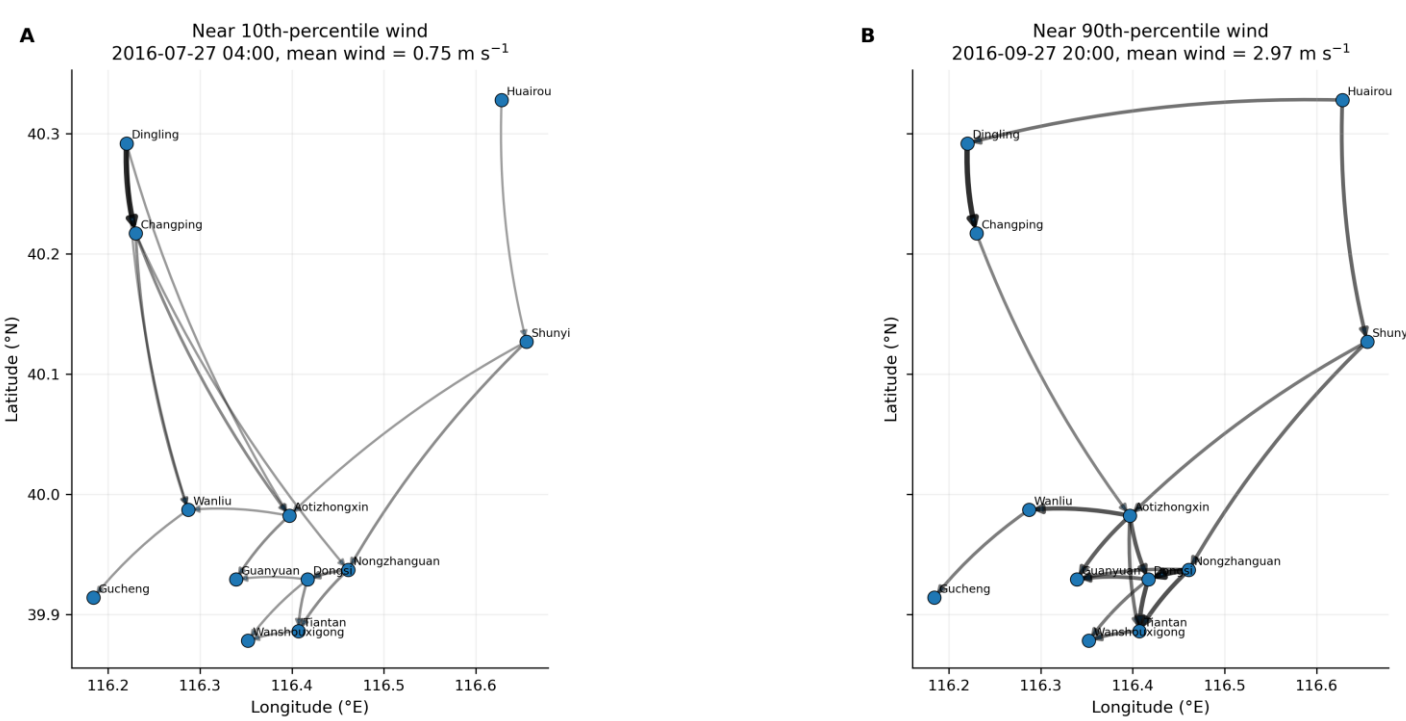
- The adjacency value represents the strength of the relationship from source station i to receiving station j : $A_t[j, i] = a_{j \leftarrow i}(t)$



Why a Dynamic Graph?

- Stations are graph nodes.
- Edges represent possible spatial influence.
- Edge direction follows wind alignment.
- Edge weights change hourly with wind and distance.

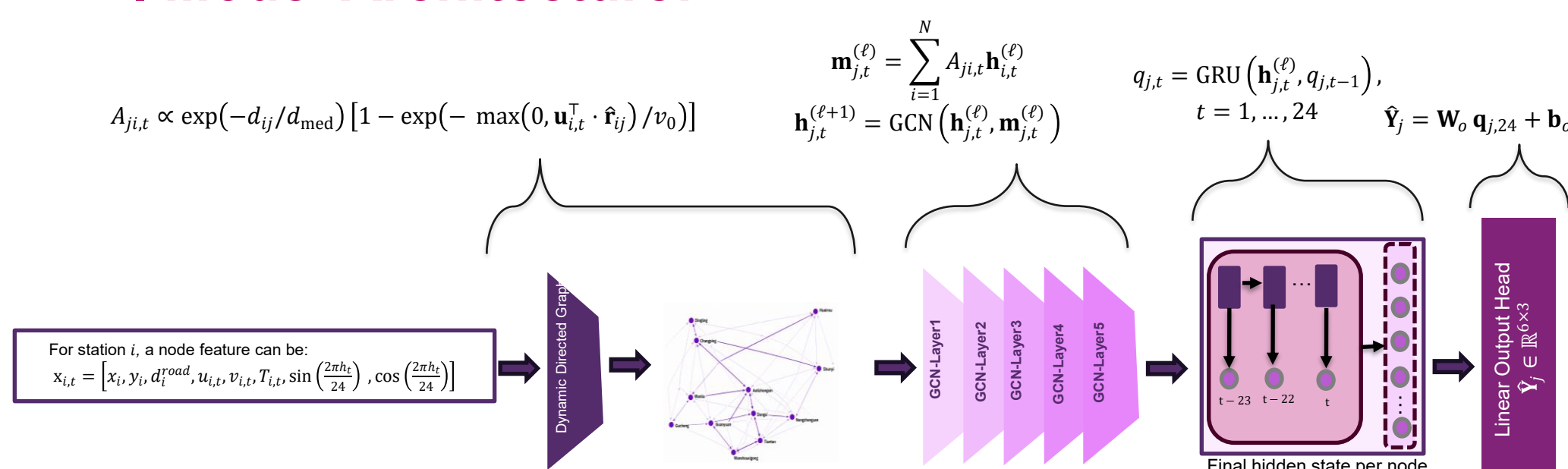
Strongest dynamic directed graph connections at the final input hour



II. Spatiotemporal GCN-GRU model

- Pollutant concentrations are driven by both **spatial interactions between monitoring stations** and **temporal evolution over time**.
- A **GRU** captures temporal dependencies but does not explicitly represent relationships between stations.
- A **dynamic directed GCN** captures spatial dependencies using changing network connections but does not model sequential dynamics alone.
- The proposed **GCN-GRU architecture integrates both components** for joint spatiotemporal forecasting.
- Model architecture:** Hourly distance- and wind-conditioned graph convolution models spatial transport, while the **GRU processes the previous 24 hours** to predict **PM_{2.5}**, **PM₁₀** and **NO₂** for 1–6 h ahead.

III. Model Architecture:



CONTACT INFORMATION:

School of Aeronautical, Automotive, Chemical and Materials Engineering, Loughborough University
Loughborough university
Leicestershire LE11 3TU UK
n.esfahanian@lboro.ac.uk

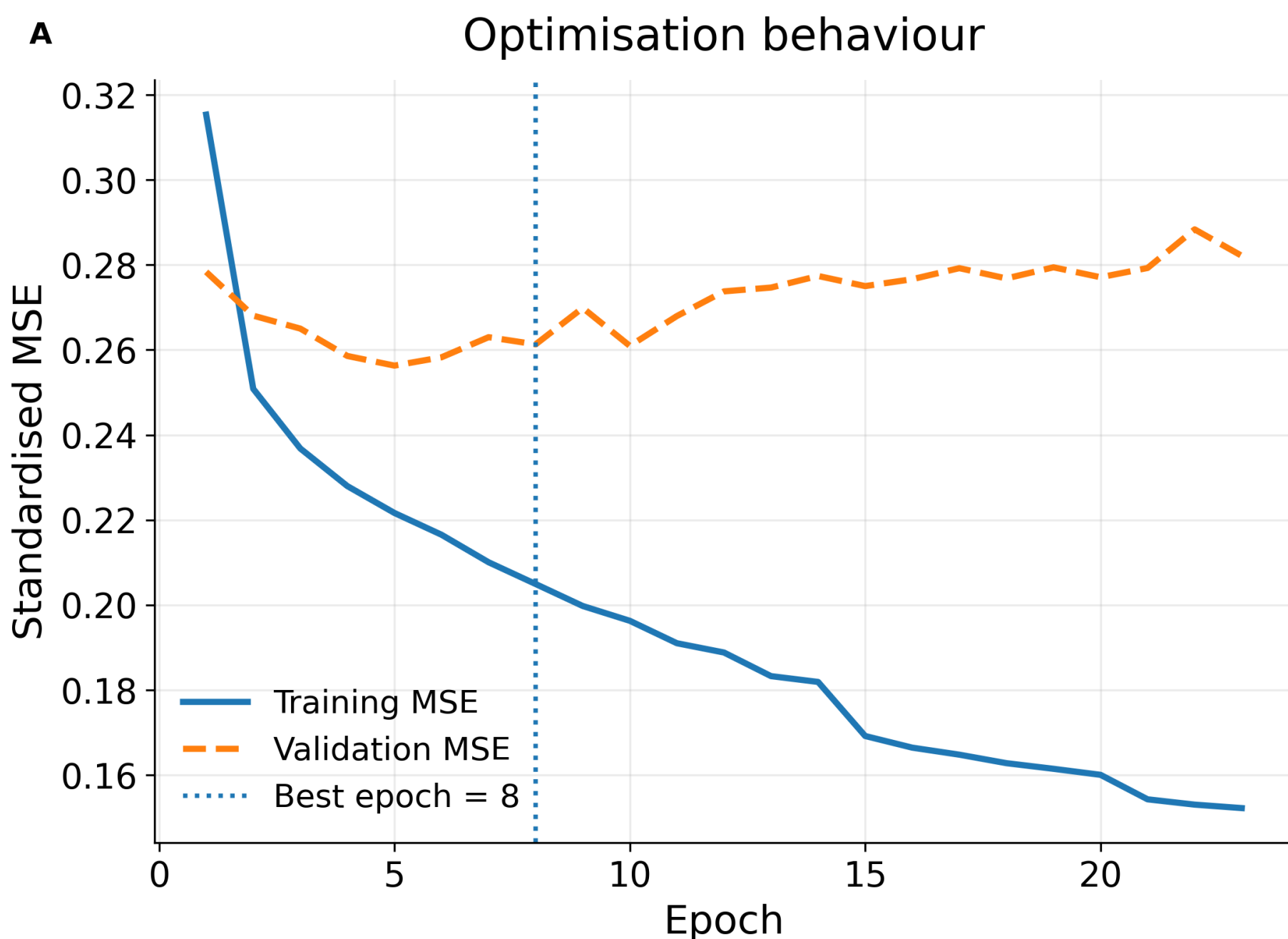


References:

- [1] World Health Organization. (2021). *WHO global air quality guidelines: particulate matter (PM_{2.5} and PM₁₀), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide*. World Health Organization.
- [2] Tang, D., Zhan, Y., & Yang, F. (2024). A review of machine learning for modelling air quality: Overlooked but important issues. *Atmospheric Research*, 300, 107261. <https://doi.org/10.1016/j.atmosres.2024.107261>
- [3] Chen, S. (2017). *Beijing Multi-Site Air Quality [Dataset]*. UCI Machine Learning Repository. <https://doi.org/10.24432/C5RK5G>

Hyperparameter tuning and final model selection

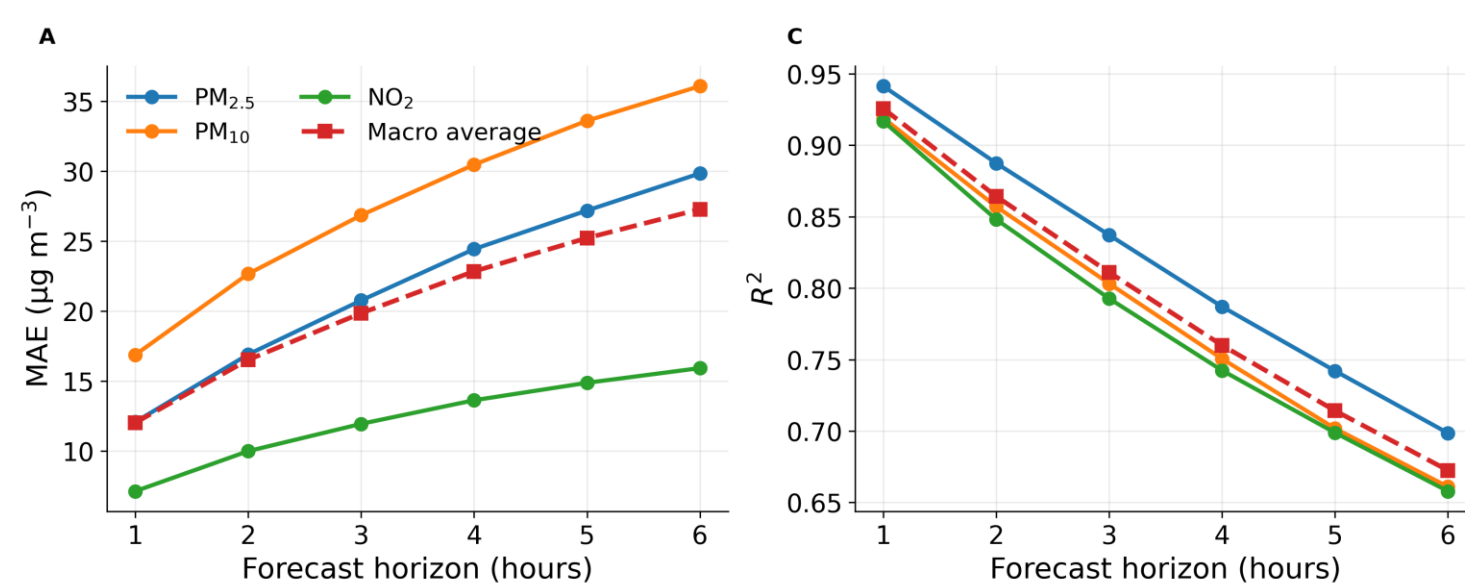
- Optuna tuning:** 150 trials across key GCN and training hyperparameters.
- Best model:** macro-normalised validation MAE $\approx$ **0.288**, with **5 GCN layers**.
- Early stopping:** best checkpoint selected at **epoch 8** to prevent overfitting.



Overall predictive performance

- The optimised **dynamic GCN-GRU** achieved **strong short-term forecasting performance** on the held-out test set.
- At **1 h**, the model achieved a **macro-average $R^2 = 0.926$** across PM_{2.5}, PM₁₀ and NO₂.
- Performance declined with forecast horizon, but remained predictive at **6 h**, with **macro-average $R^2 = 0.672$** .

Test performance degradation with forecast horizon



Conclusion

- Predictive performance declined with lead time, with greater uncertainty at longer horizons and for extreme concentrations.
- Wind-dependent graph connectivity** enabled the model to adapt to changing spatial relationships across the monitoring network.
- The framework could support **short-term air-quality warning systems** by identifying where and when pollutant concentrations are likely to increase.
- Its dynamic directed structure also provides a foundation for modelling **spatial pollutant transport and inter-station influence** across urban areas.

Future Work: Towards a City-Scale Air-Quality Digital Twin

- Enhance long-horizon forecasting** through longer temporal receptive fields, multi-scale temporal modelling.
- Extend the framework towards an **urban air-quality digital twin**, integrating real-time sensor, meteorological and emissions data with the spatiotemporal forecasting model.