Urban Climate and Mobility: A GNN-based Approach to Predict the Effect of Urban Climate on Personal Mobility Choices in Seoul, South Korea

Michele Giampaolo
1st supervisor: Azarakhsh Rafiee
2nd supervisor: Martín Mosteiro Romero

1 Introduction

As the role of urban climate becomes increasingly central to urban planning and changing climates push governments to adapt to new realities, predictive models are becoming ever-important for developing context-aware planning (Mosteiro-Romero et al., 2024; Liu et al., 2024). In particular, the flow of people's movements in the urban landscape can be studied to better understand a city's dynamics and critical issues. Although countless factors affect an individual's daily mobility choices — like daily commute, nearest public transport stops, terrain morphology, location of sports amenities, etc. — air temperature and other climate factors are among some of the most impactful when understanding city-wide mobility patterns (Jeon et al., 2023; Gehl, 2010).

The relationship between climate and citizen behaviour has already been widely studied in the urban context. However, these studies have often relied on deterministic models, such as the Predicted Mean Vote (PMV), that can simplify the relationship between human action and climate data through generalized assumptions (Liu et al., 2024; Jayathissa et al., 2020). This approach, while helpful in producing widely applicable models, overlooks the human nature of mobility. Personal factors that impact how each person experiences —and reacts to— the climate around them are therefore not considered (Upasani et al., 2024). Activity level, body composition, metabolism, eating habits and individual preferences are some of the many human factors that are difficult to account for in these frameworks, leading to a limited understanding of how individuals respond to their immediate climate (Gehl, 2010).

As data on people's daily movements are easier than ever to acquire due to the widespread use of portable, GPS-enabled devices — such as smartphones and smartwatches — and recent advancements in machine learning, new methods are emerging that allow to study mobility in greater depth (Mosteiro-Romero et al., 2024; Rico et al., 2021; Ignatius et al., 2024; Gottkehaskamp, 2024). By leveraging personal mobility data, it is therefore possible to directly study human behaviour, which in turn means indirectly incorporating the complex individual factors that it is a product of. Machine learning applied to such datasets can find hidden patterns that more traditional methods may miss (Liu et al., 2023; Wang et al., 2024).

It is in this context that this research aims to explore patterns between mobility and climate over a 7-week period in Seoul, South Korea, using data from 22 participants. A Graph Neural Network (GNN)-based approach will be used to analyse participants' position data to carry out a next-location prediction based on climate parameters.

1.1 Scientific Relevance

This project explores the use of GNNs as a novel tool to analyse the relationship between urban climate and mobility. Unlike traditional models that can often simplify these interactions, GNNs are capable of representing spatial and temporal dependencies more effectively. A key aspect of the project is its use of diverse datasets, including real-time position, urban morphology, and climate parameters, within a GNN framework. With Seoul as a case study, the research takes advantage of the city's dense urban structure and available open data. This combination of innovative methods and rich data sources aims to provide practical insights for improving climate-adaptive urban design.

2 Related work

2.1 Graph Neural Networks for Mobility Prediction

Graph Neural Networks (GNNs) have gained significant traction for their ability to model complex relationships in structured data.

Wu et al. (2021) have provided a comprehensive survey on GNNs, providing a taxonomy of their different types and their applications, among which spatial-temporal networks were highlighted for traffic and mobility-related objectives.

Rico et al. (2021) offers a detailed review of the application of GNNs for traffic forecasting, which are founf to surpass traditional methods.

Liu et al. (2023) utilised GNNs for next-location prediction, combining long- and short-term preferences to achieve high prediction accuracy. Similarly, Wang et al. (2024) explored GNNs in travel mode choice prediction, showing the possibility of integrating mobility data with urban context for enhanced prediction reliability.

2.2 Thermal Comfort and Human-Centric Models

Urban climate and its impact on human thermal comfort have been extensively studied. Recently, more studies have been tackling ways to include personal perceptions of climate and offer more reliable models.

Liu et al. (2024) performed predictions on outdoor thermal comfort through machine learning in Singapore. They incorporated multiple climate and urban morphology datasets to offer a more human-based method to traditional models.

An alternative to traditional indoor thermal comfort models was developed by Upasani et al. (2024). By taking into consideration building-specific paramaters and person-specific characteristic, such as clothing and mood, the proposed regression and classification models reached high accuracy rates.

Novel forms of urban data, like street-view images and wearables, were integrated by Ignatius et al. (2024) in the context of an urban digital twin. Personal feedback from the user, paired with a variety of weather and personal measurements were combined to study the relationship between outdoor thermal comfort and the built environment.

Jayathissa et al. (2020) introduced the "humans-as-sensors" approach for indoor comfort modeling, which sees human feedback as the primary source for analysing and assessing indoor climate comfort.

2.3 Urban Mobility Patterns

Jonietz (2016) investigated pedestrian profiles using movement trajectory data, emphasizing the importance of contextual and temporal factors in understanding urban mobility patterns.

Jonietz and Bucher (2017) proposed an analytical framework for enriching movement trajectories with spatio-temporal data, contributing to the methodological advancements in mobility research.

2.4 Urban Morphology and Impact on Climate

The study conducted by Jeon et al. (2023) analysed the relationship between variable land surface temperature across seasons and urban morphology. Areas with higher vegetation density and proximity to water were found to have more moderate temperatures, compared to more extensively urbanised areas

Peng et al. (2022) conducted an assessment of outdoor urban microclimate along urban waterfronts. The study highlighted the importance of variability in urban morphology for thermal variability and comfort.

3 Research Goals

3.1 Research Questions

The main research question that will drive the project is:

How do urban climate factors impact people's mobility choices in the urban landscape of Seoul?

The research aims to analyze the relationship between climate and mobility in an urban context. It proposes a GNN-based method for predicting a person's next locations by leveraging their historical position data, enriched with additional climate and urban morphology attributes. The following secondary questions will further explore the plausibility of the GNN-based method as well as comparing it to currently used climate comfort models:

- To what extent and how reliably can GNNs be used to automatically predict mobility?
- How is mobility defined within a GNN framework and how does this impact the findings of the research?
- How do user-reported thermal preference levels differ from those as estimated by climate factor-based comfort models?
- Are mobility patterns more indicating than comfort models in assessing thermal preference levels?

3.2 Scope

The primary objective of the research is to develop and evaluate a GNN framework for predicting mobility patterns. The implementation and assessment of the framework are therefore central to achieving the research goals and will address the primary research question and the first two secondary questions.

A secondary objective, as reflected by the third and fourth secondary research questions, is to compare the performance of the developed GNN method with existing climate comfort models. This comparison will be useful in assessing whether the proposed GNN method could serve as a viable alternative for current models. Although this comparison is important to the overall research themes, the primary focus remain on development of the GNN method. Therefore, the research will still be considered successful in the case that this secondary goal is not completed in total.

The scope is further defined by the dataset used, specifically the survey responses that will be the base of the GNN framework. As described in the Tools and Data section, these responses are limited to participants (mostly university students) in Korea and the research to Seoul. Although evaluating outside this geographical context and with different user bases is valuable to assess the applicability of the method, this is outside the scope of this study. The MoSCoW chart below shows an overview of the expected outcomes of the research.

Must Have	**Should Have**
• Develop GNN framework for predicting mobility patterns • Evaluate GNN method on accuracy of predictions	• Extend GNN to predict thermal preferences • Implement an existing climate comfort model and compare it to the GNN model's predictions
Could Have	**Won't Have**
• Test additional configurations of the GNN to optimize performance	• Validation of the method in different regions and user bases • Comparison with an exhaustive list of currently used comfort models

4 Methodology

This chapter outlines the methodology followed in this research. The diagram in Figure 1 shows an overview of the phases and their connections.

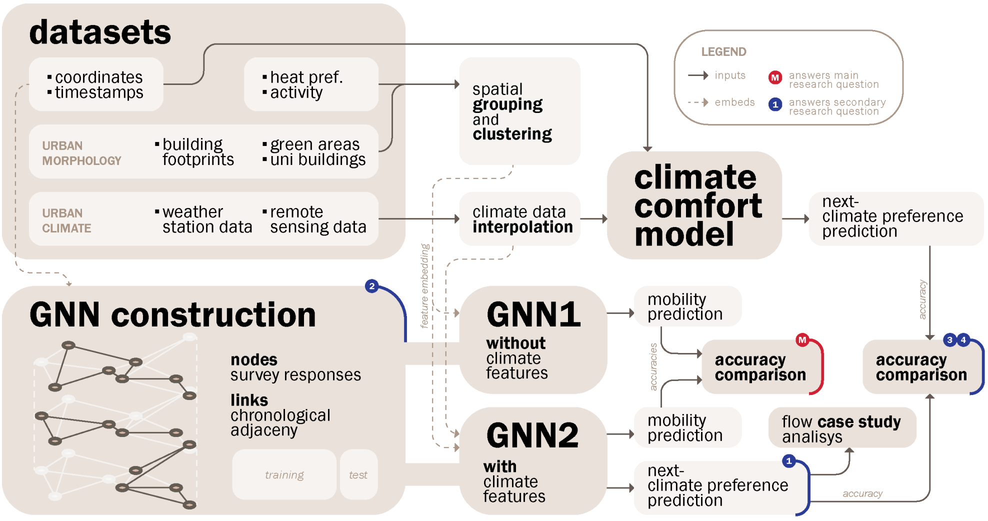

Figure 1: The database report after the CityGML file was loaded

4.1 Data Acquisition and Processing

The data required for constructing the GNN will be gathered as outlined in the Tools and Data section. The datasets are categorized into four groups: participant positions, survey responses, urban morphology and climate parameters.

The participant positions and survey responses will be cleaned in order to retain only valid and pertinent entries. Responses with errors, such as missing or incorrect corresponding coordinates, and those with coordinates outside Seoul will be excluded .

The urban morphology datasets will also be cleaned to ensure that they can be integrated without error. The cleaned data will then be merged with the participants' positions to add a *place* attribute to each one that ties it to a specific physical location. This will be done in different ways depending on the urban element that each coordinate intersects. If a coordinate intersects or is within a set buffer zone of a building footprint, the building's unique id will be assigned as the response's location. The same process will be applied for open and green areas. If a coordinate does not intersect with a predefined urban element, they will be grouped through a clustering method (e.g. K-means) to identify other recurring and significant locations that are not mapped in the morphological datasets.

Climate parameters will also be mapped to survey responses based on their geographical coordinates. For raster data, such as LANDSAT land surface temperature, the value at each response's coordinate will be directly assigned. For point data, such as meteorological measurements from weather stations, kriging interpolation will be used to estimate values at each response coordinate.

4.2 GNN Construction

Two GNN models will be developed: GNN1 and GNN2. These networks will have identical structures, with the only distinction being the inclustion of climate attributes in GNN1. Their basic structure will have each node represent a survey response, with their attributes including the corresponding temporal, geographical, activity and (for GNN1) climate data. Each edge will represent the chronological relationship between responses.

The structure shared by the two GNNs will be divided into the same training and test sets to ensure that both networks are evaluated under equal conditions.

4.3 Location Prediction

The GNNs will both perform a prediction on a participant's next few locations based on prior node information. A similar method used by Liu et al. (2023) will be implemented, where a GNN-based framework uses both long-term and short-term user preferences by modeling spatial and temporal dependencies.

The performance of the GNNs wil be assessed using the Mean Average Precision (MAP). Subgroups of the model will be also be analysed to find significant differences in accuracy under specific conditions, such as differing climate conditions. Additional metrics may be employed to provide further detail into model performance. The GNN will then be applied to a case study to output flow analyses in the covered area.

4.4 Thermal Preference Prediction

GNN1 will be used to predict participant's thermal preferences on a three-level scale: prefers warmer, prefers no change, and prefers colder. The prediction framework will follow a similar approach to that described in the location prediction task.

For comparison to widely used comfort model, thermal preference will also be predicted using the the Predicted Mean Vote Model (PMV), which will use the same climate data input as GNN1. The MAP scores of the two models will again be calculated to assess their performance. These scores will be derived from comparing the models' predictions to the actual thermal preferences recorded in participant's responses, which follow the same three-level scale. As with the location predictions, additional metrics may be employed.

5 Time planning

5.1 Schedule

The schedule that follows lays out a proposed time frame for the goals of the research in the context of the five evaluation periods.

Start	End	Activity
23 Sep	15 Nov	Topic and supervisory team formation
4 Nov	15 Nov	Initial definition of research questions
		P1 - Registration
18 Nov	30 Nov	Response data acquisition and cleaning
18 Nov	13 Dec	Literature review (urban climate and mobility)
9 Dec	4 Jan	Literature review (GNN and ML for prediction tasks)
6 Jan	17 Jan	Final definition of research questions
20 Jan	28 Jan	Drafting of P2 document and presentation
		P2 - Kick-Off
29 Jan	14 Feb	Data processing and location grouping
15 Feb	28 Feb	Development of initial GNN model framework
1 Mar	20 Mar	Implementation of location prediction
20 Mar	14 Apr	Testing and evaluation
		P3 - Midterm
21 Apr	14 May	Implementation of thermal preference prediction
15 Apr	25 May	Comparative analysis with traditional models
		P4 - Green Light
20 May	20 Jun	Conclusion of results
16 Jun	29 Jun	Drafting of final document and presentation
		P5 - Finalisation

The evaluation dates of the graduation calendar are shown in the table below. Please note that the dates following P2 still need to be decided .

Date	Event
15 Nov	**P1 - Registration**
29 Jan	**P2 - Kick-Off**
14 - 20 Apr	**P3 - Midterm**
19 - 25 May	**P4 - Green Light**
23 - 29 Jun	**P5 - Finalisation**

5.2 Meetings

Regular meetings will be held biweekly with the supervisory team on Thursdays at 10:00. They will take place in the BK building or, if meeting in-person is not possible, online through a Teams meeting. During the meetings the progress made will be presented and feedback will be shared. Additional meetings can be scheduled if deemed necessary.

6 Tools and data

6.1 Data

The datasets used by this research can be divided into four categories: participant positions, survey responses, urban morphology and climate parameters.

The participant positions and survey responses datasets originates from the study performed by Mosteiro-Romero et al. (2024), where each participant's geographical coordinates were collected on an hourly basis over the study period or when a signifiant change in position was detected. Additionally, it includes responses to questions on their thermal preference and current activity.

Urban morphological data refers to geographic datasets of the built environment of Seoul, such as building footprints and the location of public and green areas, that will be gathered in part from the city's official data platform as well as other open sources, such as OpenStreetMap.

Finally, the datasets pertaining to urban climate parameters will be gathered from local weather stations that periodically give information on the air temperature, wind speed and humidity at their location. Additionally, remote sensing datasets will be used to gather additional climate data, in particular LANDSAT and ECOSTRESS for land surface temperatures.

6.2 Tools

The research will be performed using the Python language. The cleaning and interpolation of the datasets will use the Pandas and NumPy libraries, while the construction and implementation of the GNN will use thePyTorch Geometric (PyG) library.

References

J. Gehl. *Cities for people*. Island Press, Washington, DC, 2010. ISBN 9781597269841. Description based on publisher supplied metadata and other sources.

B. G. Gottkehaskamp. Precision monitoring for outdoor thermal comfort: Cozie and spatial data synthesis for subjective and physiological feedback. Master's thesis, National University of Singapore, 2024.

J. Ignatius, Marcel Ignatiusand Lim, B. Gottkehaskamp, K. Fujiwara, C. Miller, and F. Biljecki. Digital twin and wearables unveiling pedestrian comfort dynamics and walkability in cities. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, X-4/W5-2024:195–202, 2024. ISSN 2194-9050. doi: 10.5194/isprs-annals-x-4-w5-2024-195-2024. URL https://isprs-annals.copernicus.org/articles/X-4-W5-2024/195/2024/.

P. Jayathissa, M. Quintana, M. Abdelrahman, and C. Miller. Humans-as-a-sensor for buildings—intensive longitudinal indoor comfort models. *Buildings*, 10(10):174, Oct. 2020. ISSN 2075-5309. doi: 10.3390/buildings10100174.

G. Jeon, Y. Park, and J.-M. Guldmann. Impacts of urban morphology on seasonal land surface temperatures: Comparing grid- and block-based approaches. *ISPRS International Journal of Geo-Information*, 12(482), Nov. 2023. ISSN 2220-9964. doi: 10.3390/ijgi12120482.

D. Jonietz. Learning pedestrian profies from movement trajectories. In G. Gartner and H. Huang, editors, *Proceedings of the 13th International Conference on Location-Based Services*, pages 14–16, Nov. 2016.

D. Jonietz and D. Bucher. Towards an analytical framework for enriching movement trajectories with spatio-temporal context data. In A. Bregt, T. Sarjakoski, R. Van Lammeren, and F. Rip, editors, *Societal Geo-Innovation: Short Papers, Posters and Poster Abstracts of the 20th AGILE Conference on Geographic Information Science*, number 133, Wageningen, May 2017. Association of Geographic Information Laboratories for Europe.

J. Liu, Y. Chen, X. Huang, J. Li, and G. Min. Gnn-based long and short term preference modeling for next-location prediction. *Information Sciences*, 629: 1–14, June 2023. ISSN 0020-0255. doi: 10.1016/j.ins.2023.01.131.

X. Liu, Z. Gou, and C. Yuan. Application of human-centric digital twins: Predicting outdoor thermal comfort distribution in singapore using multi-source data and machine learning. *Urban Climate*, 58:102210, Nov. 2024. ISSN 2212-0955. doi: 10.1016/j.uclim.2024.102210.

M. Mosteiro-Romero, Y. Park, and C. Miller. Converging smartwatch and urban datasets for sustainable city planning: A case study in seoul, south korea. *E3S Web of Conferences*, 562(03004), 2024. ISSN 2267-1242. doi: 10.1051/e3sconf/202456203004.

Z. Peng, R. Bardhan, C. Ellard, and K. Steemers. Urban climate walk: A stop-and-go assessment of the dynamic thermal sensation and perception in two waterfront districts in rome, italy. *Building and Environment*, 221 (109267), Aug. 2022. ISSN 0360-1323. doi: 10.1016/j.buildenv.2022.109267.

J. Rico, J. Barateiro, and A. Oliveira. Graph neural networks for traffic forecasting. 2021. doi: 10.48550/ARXIV.2104.13096.

N. Upasani, O. Guerra-Santin, and M. Mohammadi. Developing building-specific, occupant-centric thermal comfort models: A methodological approach. *Journal of Building Engineering*, 95(110281), Oct. 2024. ISSN 2352-7102. doi: 10.1016/j.jobe.2024.110281.

K. Wang, Z. Peng, M. Cai, H. Wu, L. Liu, and Z. Sun. Coupling graph neural networks and travel mode choice for human mobility prediction. *Physica A: Statistical Mechanics and its Applications*, 646:129872, July 2024. ISSN 0378-4371. doi: 10.1016/j.physa.2024.129872.

Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and P. S. Yu. A comprehensive survey on graph neural networks. *IEEE Transactions on Neural Networks and Learning Systems*, 32(1):4–24, Jan. 2021. ISSN 2162-2388. doi: 10.48550/ARXIV.1901.00596.