This is the Accepted Manuscript Version of an article published by Sage in the journal *Environment and Planning B: Urban Analytics and City Science* in 2025, which is available at: https://doi.org/10.1177/23998083251358527

Living upon Networks: A Heterogeneous Graph Neural Embedding Integrating Waterway and Street Systems for Urban Form Understanding

Pengyuan Liu[a,b], Yuan Wang[c], Stef De Sabbata[d], Binyu Lei[c], Filip Biljecki[c,e], Jing Tang[c], Rudi Stouffs[c]

[a]*Urban Analytics Subject Group, Urban Studies & Social Policy Division, University of Glasgow, United Kingdom*

[b]*Future Cities Lab Global, Singapore-ETH Centre, Singapore*

[c]*Department of Architecture, National University of Singapore, Singapore*

[d]*School of Geography, Geology and the Environment, University of Leicester, United Kingdom*

[e]*Department of Real Estate, National University of Singapore, Singapore*

Abstract

Cities are supported by multiple, interacting networks, most prominently streets, which channel movement and economic exchange, and, in many contexts, waterways, which regulate flows of goods, people, and environmental amenities. Conventional quantitative studies of urban form have tended to privilege streets alone, limiting their ability to capture the full spatial logic of the urban fabric. This paper introduces a Heterogeneous Graph Autoencoder (HeterGAE) that jointly embeds street and waterway systems, providing a unified, graph-based representation of urban form. Using Singapore as a case study, we train HeterGAE embeddings and employ them in two downstream tasks: predicting daytime and night-time land-surface temperature (LST) and estimating resale prices of public housing. Relative to a baseline model that encodes streets only, the dual-network embeddings improve predictive accuracy by about 20% for both tasks, confirming that natural and built infrastructures make complementary contributions to urban socio-environmental processes. By capturing the interaction between street junctions and waterway nodes within a single latent space, the proposed approach provides a flexible template for GeoAI-assisted urban analytics in diverse settings. The results underscore the value of integrating heterogeneous urban networks in evidence-based planning and highlight the potential of

graph-neural techniques for developing more nuanced and sustainable urban strategies.

Keywords: Urban Form, Neural Embedding, Graph Neural Network, Urban Climate, Socio-economics

1. Introduction

Urban form is shaped by a complex interplay of both engineered and natural networks, primarily street systems and, in many contexts, waterways (Gandy, 2004). In many urban contexts, waterways function as essential arteries, facilitating the flow of resources into and out of urban centres. Consequently, they often form the backbone of a city's infrastructure, influencing the location of industries, residential areas, and commercial districts, and hence, dictating the layout of the cities' structures (Haeffner et al., 2017).

Other than waterways, street networks are widely recognised as fundamental and often primary carriers that shape and define the structure and identity of cities (Boeing, 2022; De Sabbata et al., 2023). These networks serve as another artery through which the lifeblood of the city (i.e., its people, goods, and services) flows, influencing everything from the placement of buildings to the patterns of daily movement and interaction (Wang et al., 2012). Therefore, both waterways and street networks are crucial for a comprehensive understanding of urban form, as they are interconnected and interact in ways that support urban activities (Cai et al., 2018) and create a dynamic infrastructure where roads often intersect with or run parallel to waterways, facilitating seamless connectivity across urban landscapes (Bell et al., 2021).

Quantitative analysis of urban form has traditionally focused on street network analysis, utilising various predefined measures to understand the structure and functionality of cities. Common approaches include topological measures, such as connectivity and centrality (Boeing, 2022), which evaluate the network's accessibility and the significance of individual streets within the broader urban fabric. Geometric measures, including street length, width, and intersection density, are also employed to understand how these physical characteristics influence urban commute patterns (Xie and Levinson, 2007). In recent years, there has been a notable shift towards integrating advanced artificial intelligence (AI) methods into urban form analysis. Graph neural networks (GNNs), in particular, have gained prominence for their ability to model street networks as graphs, capturing both the spatial proximity and the intricate interconnections among roads. These AI techniques have been widely adopted to study urban form (De Sabbata et al., 2023) and function (Zhang et al., 2023), the built environment (Lei et al., 2024), and the spatial homogeneity of road networks (Xue et al., 2022), offering deeper insights into the complexities of urban systems.

However, although street networks have provided critical insights into the structural dynamics of cities, the significant influence of waterways is often overlooked in urban analysis (Peng et al., 2024). Recognising street junctions as pivotal anchors within their local networks, which serve as crucial nodes in the flow of urban movement and activity (Batty, 2013; De Sabbata et al., 2023), our research seeks to address this gap by integrating waterways as an essential component. This integration transforms street junctions into multi-dimensional spaces that reflect a richer understanding of urban spatial dynamics. We employ a heterogeneous graph autoencoder (HeterGAE) that integrates both street and waterway networks, allowing us to capture the complex interplay of these key urban infrastructures. By learning node embeddings that reflect network connectivity and spatial relationships, HeterGAE enables a more comprehensive view of how roads and waterways jointly shape urban form. Throughout this paper, we show how these embeddings can predict land surface temperature (LST) and public housing prices, thereby illustrating the approach's versatility in real-world urban analytics.

To summarise, this paper:

- introduces a heterogeneous graph framework for representing the urban physical environment, capturing the multi-layered nature of cities;
- develops a neural embedding technique using HeterGAE, which simultaneously encodes street and waterway connections to enrich our understanding of urban form;
- demonstrates the embeddings' utility in downstream tasks—specifically, LST prediction and public housing price estimation—showing their practical value in complex urban analyses.

2. Background

2.1. Urban Form, Networks, and Junctions

Urban form, the physical structure of cities, has been a subject of extensive study, evolving from early morphological investigations to more computationally driven analyses. Early work in urban morphology focused on how historical forces shape built environments by examining town plans, building fabrics, and plot patterns (Conzen, 1960; Moudon, 1997). As scholars sought more systematic ways to quantify spatial relationships, the development of space syntax introduced graph-based measures (e.g., integration, choice) to describe how the configuration of interconnected urban spaces influences pedestrian movement, land use, and social interaction (Hillier et al., 1976; Van Nes, 2014). By situating movement and visibility at the heart of urban form, space syntax builds on morphological principles, bringing a network-oriented lens that highlights how spatial structure reflects and guides socio-economic processes. Hence, networks have long been central in understanding the city's form (Sevtsuk and Mekonnen, 2012).

Among the urban networks, street systems have been widely recognised as the primary frameworks through which cities are organised (Sharifi, 2019; Boeing, 2022). By providing essential routes for transportation and circulation, streets not only affect land-use patterns and real estate values but also serve as carriers of population flow, shaping how people navigate and experience urban spaces (Fleischmann and Arribas-Bel, 2022; Balsa-Barreiro et al., 2021). Within this framework, street junctions, points where multiple roads intersect, stand out as pivotal nodes of interaction (Ahmed et al., 2013), frequently becoming hubs of economic and social activity due to their accessibility and visibility (Bird, 2001). Major intersections often host commercial centers, public squares, and transit nodes, acting as anchors of place that influence both local and city-wide spatial organisation (Whitelegg, 1994; Batty, 2013).

In addition to street networks, waterways form another vital, though comparably less studied, type of network within cities, significantly contributing to our understanding of urban structure and layout (Haeffner et al., 2017). Streets often align with or intersect waterways, forming crucial junctions that serve as nodes of urban activity. These intersections frequently become hubs of commerce, industry, and social interaction, playing a central role in shaping the overall structure and vitality of the city (Luo et al., 2022). Although manifest in various forms, the intertwined connections between streets and waterways create integrated networks that enhance urban mobility, economic activity, and spatial organisation. For instance, bridges or overpasses where roads cross rivers or canals transform the adjacent junctions into critical nodes supporting vehicular and pedestrian traffic (Kondolf and Pinto, 2017). Similarly, parallel connections occur where roads run alongside waterways, such as riverside roads or canal towpaths, with junctions providing access points to recreational spaces or waterfront amenities (Gobster et al., 1998). Such integrated networks are essential for understanding how cities function and thrive, illustrating the importance of considering both street and waterway connections in urban planning and design.

Building on the interplay between street networks and waterways, our research proposes a spatial modelling approach that integrates these two critical urban elements to study how street junctions function as key urban forms and structure determinants to facilitate our understanding of the underlying socio-economic patterns and environmental conditions.

2.2. Graph Theory and Neural Embedding

Graph theory has long been a foundational tool in the computational modelling of urban flows and networks, offering a framework for analysing the complex interconnections that define urban spaces (Thomson and Richardson, 1995; Anderson and Dragićević, 2020). By representing cities as graphs, where nodes correspond to intersections or significant points and edges represent the connections between them, graph theory allows researchers to quantitatively assess the structure and efficiency of urban networks (Marshall et al., 2018) and examine how different components of a city's infrastructure, such as streets and waterways, interact to shape urban form (Boeing, 2018, 2022).

Recent years have witnessed increasing interest in adopting GNNs for urban-related studies (Liu and Biljecki, 2022). The key driver behind the adoption of GNNs is the fact that they allow the creation of spatially-explicit models (Mai et al., 2023) due to their ability to algorithmically process graph-structured data. As street networks can be represented in graph formats, and GNNs can be used to learn numerical representations (embeddings) of nodes (or graphs), GNNs hold the potential to enable a numerical analysis of urban form through the encoding of the street network graph into an embedding space (Fan et al., 2024). A recent study by De Sabbata et al. (2023) is a fundamental inspiration for our research, which proposed an unsupervised graph representation learning framework, specifically using Graph AutoEncoder (GAE) models, to generate numerical embeddings of street junctions that capture the urban form of cities. The study demonstrates the framework's effectiveness through a case study of Leicester by training the model on a sample of street junctions from various UK cities. The results showed that these embeddings can represent the transition from urban to suburban forms and capture meaningful spatial patterns related to, but also different from, traditional metrics.

Based on their research, our paper explores new urban modelling and analysis methods that can integrate the interplay among urban networks within a HeterGAE framework.

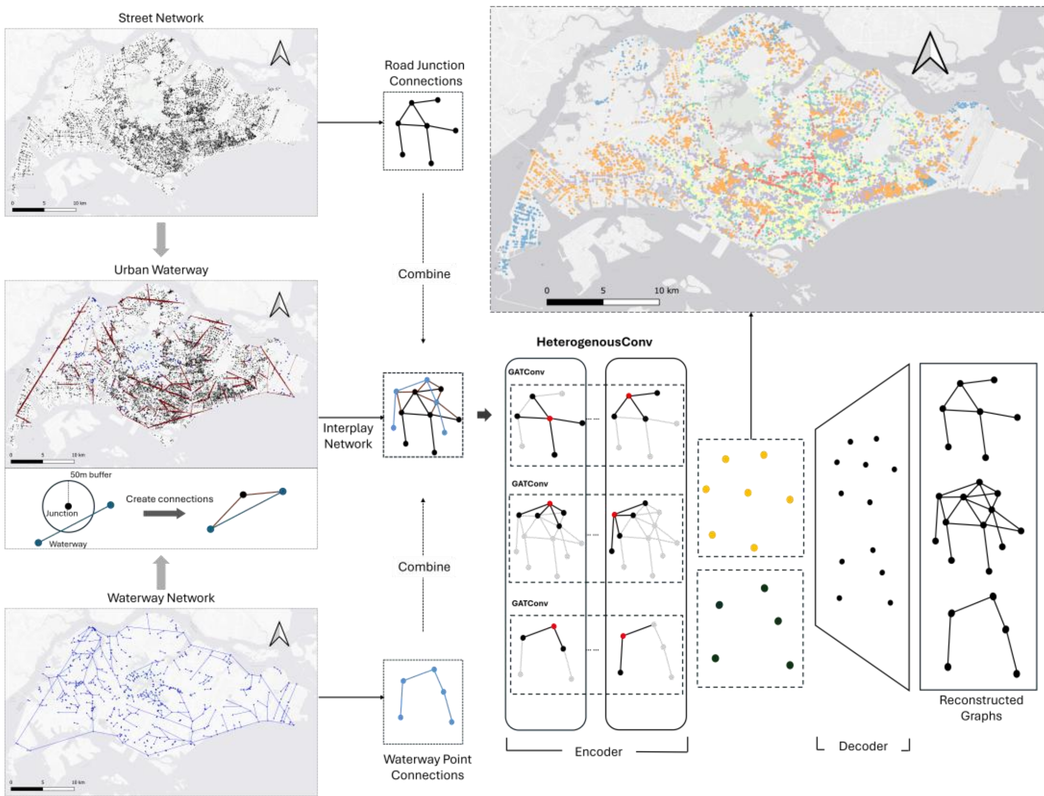


Figure 1: An overview of the framework. Basemap data source: Esri, HERE, Garmin, INCREMENT P, © OpenStreetMap contributors, and the GIS user community.

3. Methods

Figure 1 presents an overview of the proposed framework, which consists of two main components: a spatial graph modelling of urban networks and a HeterGAE processing the input graph into output neural embeddings for downstream analytics. More details are provided in the following subsections.

3.1. *Spatial Graph Modelling of Urban Networks*

As illustrated in Figure 1, our spatial modelling of the urban environment integrates two key urban networks: the street network and waterway connections. It is important to note that our analysis focuses on the structural connectivity of urban networks; that is, the topological and spatial configuration of nodes and edges, rather than on their functional characteristics such as traffic volume or pedestrian counts. Following the methodology outlined by De Sabbata et al. (2023), we sourced the street network data from the Global Urban Street Networks dataset (Boeing, 2020), which is annually updated. In constructing the street network, we represented street junctions as nodes within the graph, while the street segments that connect these junctions were modelled as edges, as shown at the top of Figure S1 in the Supplementary Material. This graph-based representation captures the essential structural elements of the urban street network, allowing us to analyse the topological closeness centrality (the average shortest distance from each node to each other node), betweenness centrality (quantifies how often a node acts as a bridge along the shortest path between two other nodes), and degree centrality (the number of direct connections a node has with other nodes in the network) within the city. These metrics, which are widely used in urban network analysis (Boeing, 2022), are each encoded as node-level features within the model, thereby enabling it to learn from the inherent structural roles of different junctions. By incorporating these measures, the graph modelling process takes into account not only the physical layout of the network but also the relative importance of each node in facilitating overall connectivity. In addition, consistent with the approach used in De Sabbata et al. (2023), we enriched our model by assigning the lengths of the streets as features for the edges in the graph.

Meanwhile, the waterway data was sourced from OpenStreetMap[1], which provided comprehensive information on the various water-based elements within the urban environment. The dataset includes not only the primary waterways, such as rivers and canals that traverse the city, but also coastlines that define the city's boundaries. To represent these features accurately within our spatial model, we constructed a graph depicting each waterway as an edge, as shown at the bottom of Figure S1. These waterways' endpoints, typically junctions where one waterway connects with another, were

[1]https://www.openstreetmap.org

represented as nodes within the graph. For the coastlines, depending on the dataset that we have, the coastline is divided into a number of sections. The connecting points of those sections are also nodes in the graph, and the lines between them are the edges. They all form a part of the waterway connection graph. In line with our approach to street networks, we assigned topological metrics, closeness centrality, betweenness centrality, and degree centrality as features to the nodes within the waterway connection network. Additionally, the length of each waterway was incorporated as an edge feature in the graph, providing an analytical framework for studying the structural and functional roles these water-based networks play in the urban landscape.

Beyond simply integrating street networks and waterway connections into our spatial modelling, we introduced a crucial step to capture and analyse the spatial interactions between these two networks by creating an additional spatial graph. As illustrated in Figure 1, we implemented a method to model these interactions by establishing connections between street junctions and nearby waterways. Specifically, for each street junction in the city, we generated a 50-metre buffer zone around the junction to capture immediate, meaningful interactions (Carthy et al., 2020). Within a 50-metre radius, waterways and street networks are close enough to experience direct environmental or social impacts, such as pedestrian access to the waterway or flood risk areas around street infrastructure. If a waterway intersects this buffer, we create connections between the street junction and the two endpoints of the waterway. The distance between the junction and the endpoints is assigned as the edges' features. Consistent with the previous steps, closeness centrality, betweenness centrality, and degree centrality of this artificially created network were used as nodes' features. Such an approach not only models the direct spatial proximity between streets and waterways but also simulates the potential for interaction and influence between these networks. It is also worth mentioning that the choice of 50 meters is a hyperparameter, subject to changes depending on the specific modelling tasks.

After creating spatial graphs of the urban networks, we fed the constructed graphs into a HeterGAE network.

3.2. Heterogeneous Graph Autoencoder

A Heterogeneous Graph Autoencoder (HeterGAE) is an advanced neural network architecture designed to learn representations from graphs that consist of multiple types of nodes and edges, capturing the complexities of heterogeneous networks. Unlike traditional graph autoencoders, which typically operate on homogeneous graphs with uniform node and edge types, a HeterGAE can handle the diversity of entities and relationships that exist within heterogeneous graphs, which allows us to simultaneously integrate and analyse different types of networks, such as street networks, waterway connections, and their interplay network in this study.

For the street network introduced in the previous section, let V_s denote the set of nodes representing street junctions, and E_s the set of edges representing streets. The graph is associated with an adjacency matrix $A_s^{l_s}$, where each entry corresponds to the scalar edge attribute l_s as a feature representing the street length for the respective connection. Additionally, nodes V_s are characterised by features X_s such as degree centrality C_d, betweenness centrality C_b, and closeness centrality C_c. Together, the street network graph G_s is formally defined as:

$$G_s = (V_s, E_s),\ X_s,\ A_s^{l_s} \tag{1}$$

Similarly, the waterway connection network G_w in the city is denoted as:

$$G_w = (V_w, E_w),\ X_w,\ A_w^{l_w} \tag{2}$$

where V_w represents the set of nodes corresponding to waterway endpoints, and E_w denotes the set of edges representing waterways. The adjacency matrix $A_w^{l_w}$ encodes the connectivity between nodes, with L_w specifying the waterway length as an edge feature. Nodes V_w are also characterised by centrality measures denoted by X_w, which are similar to those used in the street network.

The interplay network captures the interactions between street junctions and waterways. Nodes in this network comprise both street junctions (V_s) and waterway endpoints (V_w), while edges E_{sw} represent the connections as processed as introduced in Section 3.1. The interplay network is formally defined as:

$$G_{sw} = (V_s \cup V_w, E_{sw}),\ X_{sw},\ A_{sw}^{l_{sw}} \tag{3}$$

where $A_{sw}^{l_{sw}}$ represents the adjacency matrix of the interplay network, with scalar edge features l_{sw} denoting the distances between street junctions and waterway endpoints. The nodes in $V_s \cup V_w$ are characterised by centrality measures computed within the context of the interplay network denoted as X_{sw}.

The HeterGAE uses an encoder-decoder structure where the encoder learns to compress the information from these heterogeneous graphs into a lower-dimensional latent space, capturing the relationships within and between the different types of networks. The decoder then reconstructs the original graph from this latent representation, ensuring that the learned embeddings preserve the structural and feature-based properties of the urban networks.

The encoding process for a heterogeneous graph is represented as:

$$H = GAT_{street}(G_s) + GAT_{waterway}(G_w) + GAT_{interplay}(G_{sw}) \tag{4}$$

where $GAT_{street}(G_s)$, $GAT_{waterway}(G_w)$, and $GAT_{interplay}(G_{sw})$ correspond to single-layer Graph Attention Network (GAT) operations (Velickovic et al., 2017) applied to the street, waterway, and interplay networks, respectively. Each GAT layer computes a latent representation for its input graph, producing feature embeddings with a dimensionality of 32 for the nodes in the respective graph. The resulting embeddings from the GAT layers are aggregated using a summation operation, combining information from the street network, waterway network, and their interactions into a unified intermediate latent representation for each node type.

The encoder consists of two layers of GAT operations. The first layer performs GAT convolutions separately for each graph, producing intermediate latent representations of dimension 32 for the nodes. Then, the second layer re-applies GAT operations using the intermediate latent features as input, integrating information across the three graphs into the final latent representation H, which has a dimensionality of 16.

The decoder part of the HeterGAE is implemented as an inner product decoder, designed to reconstruct the graph structure from the latent space embeddings:

$$\hat{A}_{uv} = \sigma(H_u^T \cdot H_v) \tag{5}$$

where u and v represent pairs of nodes in the graph; $\hat{A}_{uv}$ is the adjacency matrix of the reconstructed graph; H_u and H_v are the latent embeddings of nodes u and v, respectively. $H_u^T \cdot H_v$ represents the inner product between the embeddings of nodes u and v, and σ is a non-linear activation function, typically a sigmoid function, which ensures that the reconstructed adjacency values are between 0 and 1. The goal of the HeterGAE is to minimize the reconstruction loss, $Loss$, which measures the difference between the original three graphs and the reconstructed ones from the latent embeddings:

$$Loss = \sum_{g \in G} \sum_{(u,v) \in g} |A_{uv}^{g} - \hat{A}_{uv}^{g}| \tag{6}$$

where A_{uv} is the adjacency matrix of the original graph; G is the set of graphs, and g denotes each graph (i.e., street network, waterway connection and the interplay network).

After the neural network's learning process, we extracted the latent embedding H for our further downstream tasks. Model implementation can be found in the Model Implementation section of the Supplementary Materials.

4. Case Study

We selected Singapore as our study area; the reasons and introductions to the downstream tasks can be seen in the Case Study Area section of the Supplementary Materials.

5. Results

The results of this research are organised into three key sub-sections. First, we present a visual analysis of the embeddings produced by the HeterGAE, offering insights into the latent patterns captured by our proposed heterogeneous graph-based spatial modelling approach. This visualisation helps to illuminate the underlying structure and relationships within the urban environment as identified by the model. Next, we demonstrate how these embeddings can be effectively utilised in downstream tasks, specifically for predicting LST and estimating the resale prices of HDB flats.

5.1. Neural Embedding

To demonstrate the effectiveness of our proposed neural embedding, we included a baseline method that focuses exclusively on the road network, providing a comparative analysis against our more comprehensive approach. In this baseline scenario, we employed a standard GAE model, which utilises a two-layer graph convolutional network (GCN) (Kipf and Welling, 2017) as its encoder and a deterministic inner product setup as its decoder. This GAE processes the street network introduced in Section 3.1, where the nodes represent street junctions, and the features consist of topological metrics such as degree centrality, betweenness centrality, and closeness centrality. The edges in this network correspond to the streets, with their lengths serving as the feature of the edges. By isolating the street network in this baseline model, we aim to highlight the additional value provided by incorporating the interplay between street networks and waterways in our HeterGAE approach.

Figure S2 in the Supplementary Material provides a comparative analysis of the embeddings generated by the GAE and HeterGAE models for street junctions, with the results categorised using the K-means clustering algorithm. This comparison shows that the HeterGAE embeddings capture a more nuanced urban representation than those generated by GAE, as evidenced by their respective Silhouette Scores (Rousseeuw, 1987) of 0.73 versus 0.31. Specifically, the HeterGAE model, which integrates both street and waterway networks, offers a more detailed characterisation of street junctions by incorporating the influence of major waterways within the city. The embeddings result in clusters that reflect the underlying urban structure, particularly in areas where the presence of waterways significantly impacts the organisation and function of the surrounding streets.

While the clustering algorithm provides initial evidence that the embeddings can capture the underlying geographic patterns of urban networks, relying solely on clustering results may not be sufficient to fully assess the effectiveness and practical utility of the embeddings. Therefore, we incorporated two downstream tasks: LST prediction and estimating HDB resale prices, which test how the embeddings can be applied in different urban contexts and assess their ability to capture relevant urban features and patterns beyond what is evident from clustering alone.

5.2. Land Surface Temperature Prediction

The daytime LST data, integrated from 2015 to 2020, was extracted from Landsat 8 OLI via the Google Earth Engine (GEE) platform. The statistical mono-window model (SWM) was applied to retrieve LST from Landsat 8 OLI's band 10 (Ermida et al., 2020). Furthermore, fractional vegetation cover derived from the normalised difference vegetation index (NDVI) was used to refine the calculation of emissivity, addressing the issue of vacant emissivity data in the original algorithm (Wang et al., 2020). This LST data was then resampled to a resolution of 1 km using bilinear interpolation. Meanwhile, the nighttime data, due to the lack of a nighttime Landsat data source integrated from 2015 to 2020, was produced from Level 2 Gridded Moderate Resolution Imaging Spectroradiometer (MODIS) intermediate LST product

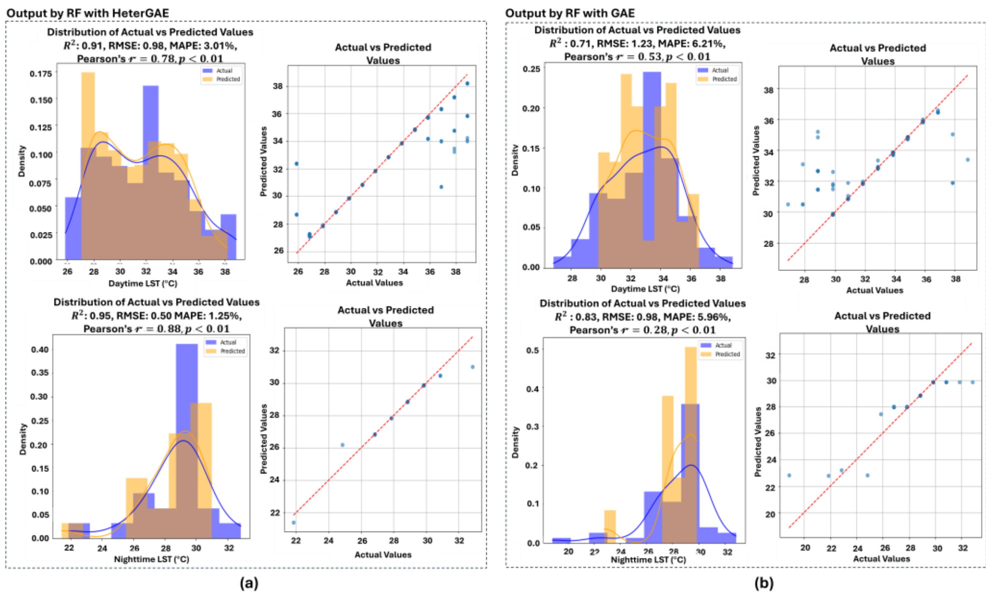


Figure 2: Results for LST predictions using HeterGAE vs. GAE. Reported metrics (e.g., R^2, RMSE, MAPE, and Pearson's r indicate that HeterGAE-based predictions (left) generally align more closely with observed LST than do GAE-based predictions (right).

at a resolution of 1 km. Both data obtained were in the form of raster data; we then vectorised the raster data into polygons for the ease of further analysis, as shown in the example of the daytime data process in Figure S3 in the Supplementary Material.

The generated embedding output by HeterGAE was mapped and aggregated into each polygon in the map. Then, a random forest (RF) model with 500 forest trees was implemented for the regression task on the Daytime LST prediction. The data was split into conventional 7:3 training-test data ratio; that is, 70% percentage of the polygons were randomly sampled as the training data, and the remaining 30% data was the test data. The same process was also applied to the nighttime LST predictions, and results are summarised in Figure 2 with compared results of the same RF model using GAE embedding as input.

As demonstrated in Figure 2, the regression tasks performed by the RF model using latent embeddings from HeterGAE outperform those using embeddings from GAE. Moreover, the nighttime LST predictions are more ac-

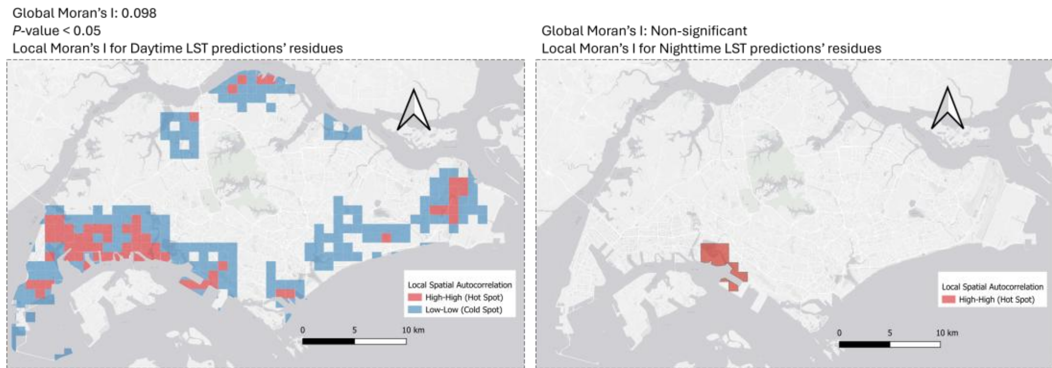


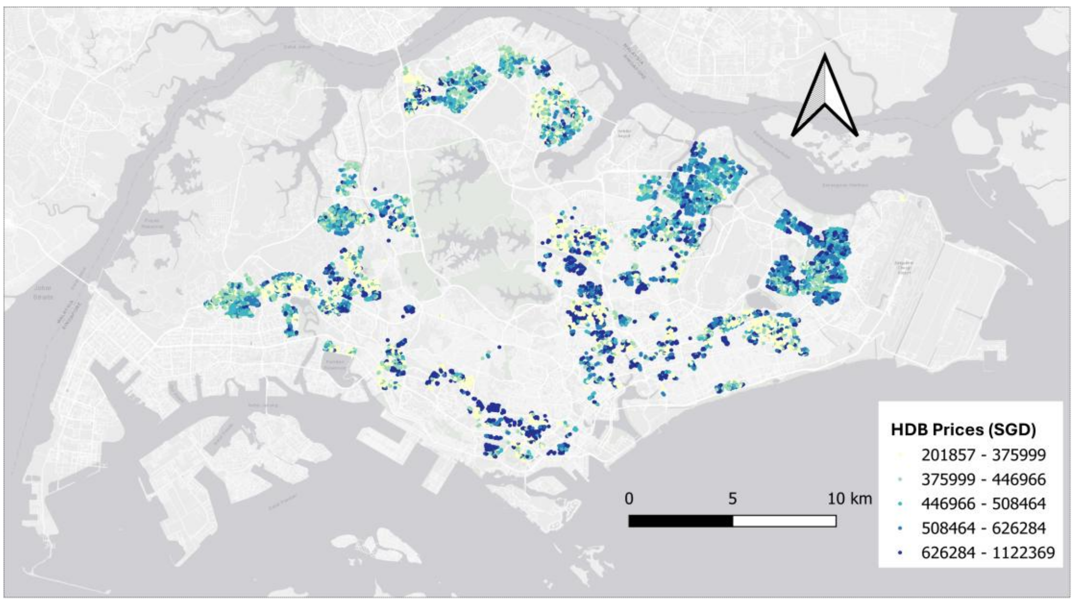
Figure 3: Spatial autocorrelation analysis for the prediction residuals. Areas shaded red indicate contiguous high-valued residuals (hot spots), whereas blue regions highlight contiguous low-valued residuals (cold spots).

curate than those for daytime LST, aligning with previous findings on LST predictions in tropical cities (Hua and Ping, 2018; Harshan et al., 2018). However, the results also reveal that the model struggles to accurately predict extreme temperatures, both high and low, across the city. Instead, it tends to produce more conservative predictions, which are clustered closer to the median of the temperature distribution. Such an issue is not uncommon for machine learning regression predictions (Fouedjio and Klump, 2019), where the bias-variance trade-off often leads to an underestimation of extreme values.

To better understand these discrepancies, we conducted a spatial analysis of the model's residuals. Using the trained RF model, we re-predicted LST across all polygons in Singapore and computed the spatial autocorrelation of the residuals, comparing predicted values against ground truth measurements. Figure 3 reveals that, for daytime LST predictions, the residuals exhibit weak global spatial autocorrelation, suggesting that the RF model with HeterGAE embeddings may not fully capture the complex spatial heterogeneity of LST distributions across the city. Notably, these mis-predictions are concentrated in industrial areas (shown on the left and top of the figure), near the airport (right section), and in downtown regions (bottom-right section). These areas are characterised by distinct microclimatic factors and urban heat island effects, which may not be adequately represented by the model's latent features. Thus, future research may need to include additional microclimatic features to enhance the model's performance.

In contrast, the nighttime LST predictions do not show significant global spatial autocorrelation. However, local Moran's I analysis highlights specific clusters of mis-predictions, particularly around the Pandan Reservoir in the western part of Singapore. This area, known for its large water body and surrounding recreational spaces, likely introduces local microclimatic variations that are challenging for the model to capture using embeddings that focus on urban networks.

5.3. Public Housing Price Estimation



Metrics	HeterGAE with building-specific features	GAE with building-specific features	Only building-specific features	HeterGAE	GAE	GWR
RMSE	79400.60	85634.83	98007.19	138926.91	140214.45	79412.54
MAPE	8.88%	9.76%	10.79%	25.16%	26.23%	8.87%
R^2	0.74	0.69	0.65	0.33	0.35	0.73
Pearson's r	0.67***	0.52***	0.50***	0.18***	0.11***	0.66***

Figure 4: The mapping of HDB flat prices to the buildings' footprints and the results summarised using the proposed HeterGAE-generated embedding with RF and baseline methods; bar inset reports test-set accuracy (R^2, RMSE, MAPE). SGD in the figure refers to Singapore Dollars. $***$ denotes for $p < 0.01$.

HDB resale price data from 2022 to 2024 was sourced from Singapore's open data platform[2] and spatially linked to building footprints using polygon data retrieved from OpenStreetMap, as depicted in Figure 4. The resale prices of HDB buildings were aggregated by calculating the mean resale price of all flats within each building as the target values (i.e., dependent variable), forming the basis for building-level predictions in the regression task. HDB resale prices represent a complex socio-economic indicator shaped by multiple factors, such as the flats' location, average floor area, and construction year (lease commencement year in the Singapore context). In Singapore, a particularly critical factor is the remaining lease years, as HDB flats are limited to a 99-year lease. To enhance the accuracy of public housing price predictions, we integrated not only the embeddings generated by HeterGAE but also additional building-specific features, including the aforementioned factors, which aim to better capture the nuances that influence resale prices, providing a better predictive model.

For the experiments, we utilised the same RF model implemented in the previous section but tested six different input configurations: (1) HeterGAE embeddings combined with building-specific features (such as average floor area, remaining lease years, and construction year), (2) GAE embeddings combined with the same building-specific features, (3) building-specific features only, (4) HeterGAE embeddings alone, (5) GAE embeddings alone, and (6) Geographically Weighted Regression (GWR) (Fotheringham et al., 2009) with building-specific features. The RF's out-of-bag error (tracked during training) differs by < 10% from the held-out test-set error, suggesting minimal over-fitting with all embeddings employed. For experiments involving either HeterGAE or GAE embeddings, we represented each HDB building by summing the embeddings of the 15 nearest road junctions (set as a hyperparameter that can be adjusted based on urban settings) to the building, thus enriching the feature set with spatial context from the surrounding urban network. Note that the predictions are at the building scale; therefore, buildings are not part of the graph. Instead, road junction embeddings produced by the GNNs are used as part of building features. All numerical performance metrics reported below are computed on a 30% hold-out test set that was not used during model training.

The results summarised in Figure 4 demonstrate that both the Heter-

[2]https://data.gov.sg/

GAE and GAE embeddings, when combined with building-specific features, outperform models using only building-specific features or embeddings alone, underscoring the value of integrating urban form with building characteristics for housing price predictions. Notably, HeterGAE, combined with building-specific features, yielded the highest accuracy in price estimation (R^2=0.74), validating the hypothesis that incorporating waterway-influenced urban form improves the model's ability to capture local socio-economic patterns. This confirms that the spatial interactions between street networks and waterways contribute meaningfully to understanding housing market dynamics in urban environments. It is worth highlighting that the RF models leveraging HeterGAE and GAE embeddings performed poorly when building-specific features were excluded. In contrast, the RF model, which relies solely on building-specific features, delivered reasonable performance (R^2 = 0.65), underscoring the importance of incorporating spatial context in housing price predictions (Soltani et al., 2021). Although building-specific features alone provide a solid baseline, the lack of urban network information limits the model's overall accuracy.

Additionally, we employed GWR with the same building-specific features as a classic spatial baseline. Interestingly, GWR achieves a predictive performance (R^2=0.73) comparable to the RF model using HeterGAE embeddings, suggesting that a coordinate-based location approach can likewise capture local market variations. Nevertheless, the heterogeneous graph-based perspective of HeterGAE encodes the connectivity and topology of multiple urban networks (e.g., streets and waterways), offering a network-centric lens that extends beyond housing price estimation to other tasks, such as environmental or socio-economic analyses. In principle, HeterGAE can be adapted to broader or multiple regions if a sub-graph training strategy is adopted (De Sabbata et al., 2023), a flexibility not easily matched by GWR. Hence, while building-specific attributes remain essential, integrating them with a heterogeneous, network-structured representation of urban form remains a promising avenue for more comprehensive housing market modelling.

After the model evaluation, the best predictive model trained on the 70% training portion was applied to the full dataset to generate the predictions shown in Figure S4 in the Supplementary Material. Figure S4 presents a two-part visual analysis of the HDB price prediction task. The left panel maps predicted resale prices across the entire city, derived from applying the trained HeterGAE-based model to all buildings. This output illustrates the spatial variability in housing prices and highlights general trends captured by the model. The right panel shows a local spatial autocorrelation map of prediction residuals, revealing statistically significant clusters of over- and under-prediction. Visual inspection of the residual indicates that systematic under-predictions are concentrated in the city's network-dense or centrally located precincts, which suggests that while the model successfully encodes structural connectivity, it may not fully account for localised premiums associated with accessibility, amenity concentration, or historically valued urban character.

6. Integrating Urban Blue and Green Networks

Beyond urban road and waterway networks, green spaces are key drivers in shaping urban forms (Zhang et al., 2022). We further included a study on integrating urban blue and green networks into understanding the urban form. Details can be seen in the Integrating Urban Blue and Green Networks section of the Supplementary Materials.

7. Discussion and Conclusion

In this study, we underscore the value of integrating additional urban natural environments, such as waterways, alongside street networks for a more holistic understanding of urban environments and their socio-economic and environmental outcomes. Our HeterGAE framework captures the interplay between road junctions and waterways, offering deeper insights into how urban form shapes phenomena like land surface temperature and housing resale prices. However, it is worth noting that waterways are only one example of how natural or alternative infrastructure can complement conventional roads. In cities lacking extensive water-based infrastructure, such a multi-network approach can be adapted to include other relevant systems, ensuring both flexibility and wide applicability.

The first key outcome was the usefulness of the proposed embedding in LST prediction. By incorporating waterway networks, we were able to accurately capture temperature variations across the city. Waterways complement street networks in contributing to urban cooling, mitigating the UHI effect by lowering surface temperatures in their vicinity. This cooling effect, particularly during the daytime, is crucial for urban environments nowadays, facing the increasing challenges of global warming, where managing heat is a major concern for urban sustainability and livability. Including waterway-related features enhanced the model's ability to predict LST, illustrating the value of integrating both natural and built environments in environmental modelling (Batty, 1976).

The second important result concerns housing resale price estimation. Our findings showed that proximity to water not only influences urban temperatures but also determines housing values. The HeterGAE model, which integrates waterway influences, demonstrated higher accuracy in predicting housing resale prices than models that only used building-specific or street network data. Such a finding reinforces the idea that urban form, shaped by both streets and waterways, has a profound impact on socio-economic dynamics, particularly in real estate markets (Walsh et al., 2011). Interestingly, while the HeterGAE embeddings alone did not perform as well without building-specific features, models using only building-specific data achieved reasonable performance. Such a finding highlights that, while spatial context provided by urban networks is critical, intrinsic building characteristics are equally fundamental in predicting housing prices. However, the superior performance of models that integrated both spatial and structural data underscores the importance of combining diverse data types for more comprehensive urban analysis.

Additionally, we demonstrated the generalisability of our method by extending HeterGAE to include an urban green space network, thereby creating a cohesive framework that links urban blue (waterways) and green spaces. This integrated approach allows for a more comprehensive understanding of urban forms, capturing the interplay between these two vital environmental elements. By jointly considering the roles of both blue and green spaces, the model exhibits enhanced performance in housing price estimation, highlighting the close relationship between greenery and urban socio-economics (Dai, 2011).

Several directions will be pursued in our future studies. First, future research will expand this framework by incorporating additional urban networks, such as public transportation systems, to offer a more comprehensive view of urban environments. Public transportation networks, including bus and rail systems, often shape the spatial structure of cities and influence patterns of mobility, accessibility, and land use. Incorporating these networks may hold the potential to provide deeper insights into how transport connectivity interacts with urban form, influencing socio-economic factors such as housing prices and the distribution of economic activities. Second, to test the robustness and generalisability of the HeterGAE framework, we aim to apply it to cities with diverse geographic and climatic conditions. Urban areas in different climates, such as temperate, arid, or cold regions, may exhibit distinct interactions between natural and built environments. Applying the HeterGAE framework across multiple cities will help determine whether the inclusion of waterways, street networks, and other urban features leads to consistent patterns in LST prediction and housing price estimation, which allows us to identify specific regional factors or variations that may influence the model's performance and provide insights into how urban form adapts to different environmental challenges. Third, we plan to refine our predictive models by incorporating temporal data. The dynamic changes in urban infrastructure, housing markets, or environmental conditions can be integrated into the model, allowing us to predict how urban form evolves and assisting urban planners in forecasting future urban growth patterns, mitigating climate risks, and making long-term decisions that foster sustainable urban development. Last but not least, one overarching limitation of most advanced predictive models, including ours in this paper, is their limited explainability and interoperability, which are the key factors contributing to ethical and responsible urban AI systems (Liu et al., 2024). Our future work intends to address this gap by developing techniques that make our models more transparent and readily interpretable for researchers, policymakers, and the general public.

References

Ahmed, I., et al., 2013. Road infrastructure and road safety. Transport and Communications Bulletin for Asia and the Pacific 83, 19–25.

Anderson, T., Dragićević, S., 2020. Complex spatial networks: Theory and geospatial applications. Geography Compass 14, e12502.

Balsa-Barreiro, J., Morales, A.J., Lois-González, R.C., 2021. Mapping population dynamics at local scales using spatial networks. Complexity 2021, 8632086.

Batty, M., 1976. Urban modelling. Cambridge University Press Cambridge.

Batty, M., 2013. The new science of cities. MIT press.

Bell, S., Fleming, L.E., Grellier, J., Kuhlmann, F., Nieuwenhuijsen, M.J., White, M.P., 2021. Urban blue spaces: Planning and design for water, health and well-being. Routledge.

Bird, R.N., 2001. Junction design, in: Handbook of transport systems and traffic control. Emerald Group Publishing Limited. volume 3, pp. 399–412.

Boeing, G., 2018. Measuring the complexity of urban form and design. Urban Design International 23, 281–292.

Boeing, G., 2020. Global Urban Street Networks GraphML. doi:10.7910/DVN/KA5HJ3.

Boeing, G., 2022. Street network models and indicators for every urban area in the world. Geographical analysis 54, 519–535.

Cai, Z., Han, G., Chen, M., 2018. Do water bodies play an important role in the relationship between urban form and land surface temperature? Sustainable cities and society 39, 487–498.

Carthy, P., Lyons, S., Nolan, A., 2020. Characterising urban green space density and footpath-accessibility in models of bmi. BMC Public Health 20, 1–12.

Conzen, M.R.G., 1960. Alnwick, northumberland: a study in town-plan analysis. Transactions and Papers (Institute of British Geographers) , iii–122.

Dai, D., 2011. Racial/ethnic and socioeconomic disparities in urban green space accessibility: Where to intervene? Landscape and urban planning 102, 234–244.

De Sabbata, S., Ballatore, A., Liu, P., Tate, N., 2023. Learning urban form through unsupervised graph-convolutional neural networks, in: Geospatial Knowledge Graphs and GeoAI Workshop.

Ermida, S.L., Soares, P., Mantas, V., Göttsche, F.M., Trigo, I.F., 2020. Google earth engine open-source code for land surface temperature estimation from the landsat series. Remote Sensing 12, 1471.

Fan, C., Yang, Y., Mostafavi, A., 2024. Neural embeddings of urban big data reveal spatial structures in cities. Humanities and Social Sciences Communications 11, 1–15.

Fleischmann, M., Arribas-Bel, D., 2022. Geographical characterisation of british urban form and function using the spatial signatures framework. Scientific Data 9, 546.

Fotheringham, A.S., Brunsdon, C., Charlton, M., 2009. Geographically weighted regression. The Sage handbook of spatial analysis 1, 243–254.

Fouedjio, F., Klump, J., 2019. Exploring prediction uncertainty of spatial data in geostatistical and machine learning approaches. Environmental Earth Sciences 78, 38.

Gandy, M., 2004. Rethinking urban metabolism: water, space and the modern city. City 8, 363–379.

Gobster, P.H., Westphal, L.M., Nilon, C., 1998. People and the river: perception and use of Chicago waterways for recreation. National Park Service, Rivers, Trails, and Conservation Assistance Program.

Haeffner, M., Jackson-Smith, D., Buchert, M., Risley, J., 2017. Accessing blue spaces: Social and geographic factors structuring familiarity with, use of, and appreciation of urban waterways. Landscape and Urban Planning 167, 136–146.

Harshan, S., Roth, M., Velasco, E., Demuzere, M., 2018. Evaluation of an urban land surface scheme over a tropical suburban neighborhood. Theoretical and Applied Climatology 133, 867–886.

Hillier, B., Leaman, A., Stansall, P., Bedford, M., 1976. Space syntax. Environment and Planning B: Planning and design 3, 147–185.

Hua, A.K., Ping, O.W., 2018. The influence of land-use/land-cover changes on land surface temperature: a case study of kuala lumpur metropolitan city. European Journal of Remote Sensing 51, 1049–1069.

Kipf, T.N., Welling, M., 2017. Semi-supervised classification with graph convolutional networks, in: International Conference on Learning Representations (ICLR).

Kondolf, G.M., Pinto, P.J., 2017. The social connectivity of urban rivers. Geomorphology 277, 182–196.

Lei, B., Liu, P., Milojevic-Dupont, N., Biljecki, F., 2024. Predicting building characteristics at urban scale using graph neural networks and street-level context. Computers, Environment and Urban Systems 111, 102129.

Liu, P., Biljecki, F., 2022. A review of spatially-explicit geoai applications in urban geography. International Journal of Applied Earth Observation and Geoinformation 112, 102936.

Liu, P., Zhang, Y., Biljecki, F., 2024. Explainable spatially explicit geospatial artificial intelligence in urban analytics. Environment and Planning B: Urban Analytics and City Science 51, 1104–1123.

Luo, J., Zhao, T., Cao, L., Biljecki, F., 2022. Water view imagery: Perception and evaluation of urban waterscapes worldwide. Ecological Indicators 145, 109615.

Mai, G., Li, Z., Lao, N., 2023. Spatial representation learning in geoai, in: Handbook of Geospatial Artificial Intelligence. CRC Press, pp. 99–120.

Marshall, S., Gil, J., Kropf, K., Tomko, M., Figueiredo, L., 2018. Street network studies: from networks to models and their representations. Networks and Spatial Economics 18, 735–749.

Moudon, A.V., 1997. Urban morphology as an emerging interdisciplinary field. Urban morphology 1, 3–10.

Peng, R., Ding, K., Guo, H., Liu, X., Liu, Y., Weng, H., Li, R., 2024. The impact of rivers and lakes on urban transportation expansion: A case study of the century-long evolution of the road network in wuhan, china. Plos one 19, e0298678.

Rousseeuw, P.J., 1987. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. Journal of computational and applied mathematics 20, 53–65.

Sevtsuk, A., Mekonnen, M., 2012. Urban network analysis. Revue internationale de géomatique–n 287, 305.

Sharifi, A., 2019. Resilient urban forms: A review of literature on streets and street networks. Building and Environment 147, 171–187.

Soltani, A., Pettit, C.J., Heydari, M., Aghaei, F., 2021. Housing price variations using spatio-temporal data mining techniques. Journal of Housing and the Built Environment , 1–29.

Thomson, R.C., Richardson, D.E., 1995. A graph theory approach to road network generalisation, in: Proceeding of the 17th international cartographic conference, pp. 1871–1880.

Van Nes, A., 2014. Space syntax in theory and practice. Geodesign by integrating design and geospatial sciences , 237–257.

Velickovic, P., Cucurull, G., Casanova, A., Romero, A., Lio, P., Bengio, Y., et al., 2017. Graph attention networks. stat 1050, 10–48550.

Walsh, P.J., Milon, J.W., Scrogin, D.O., 2011. The spatial extent of water quality benefits in urban housing markets. Land Economics 87, 628–644.

Wang, P., Hunter, T., Bayen, A.M., Schechtner, K., González, M.C., 2012. Understanding road usage patterns in urban areas. Scientific reports 2, 1001.

Wang, Y., Zhao, Y., Wu, J., 2020. Dynamic monitoring of long time series of ecological quality in urban agglomerations using google earth engine cloud computing: A case study of the guangdong-hong kong-macao greater bay area, china. Acta Ecol. Sin 40, 8461–8473.

Whitelegg, J., 1994. Roads, jobs and the economy. Citeseer.

Xie, F., Levinson, D., 2007. Measuring the structure of road networks. Geographical analysis 39, 336–356.

Xue, J., Jiang, N., Liang, S., Pang, Q., Yabe, T., Ukkusuri, S.V., Ma, J., 2022. Quantifying the spatial homogeneity of urban road networks via graph neural networks. Nature Machine Intelligence 4, 246–257.

Zhang, Q., Zhou, D., Xu, D., Rogora, A., 2022. Correlation between cooling effect of green space and surrounding urban spatial form: Evidence from 36 urban green spaces. Building and Environment 222, 109375.

Zhang, Y., Liu, P., Biljecki, F., 2023. Knowledge and topology: A two layer spatially dependent graph neural networks to identify urban functions with time-series street view image. ISPRS Journal of Photogrammetry and Remote Sensing 198, 153–168.