International Student Family Housing Site Selection: A Spatial Accessibility Analysis in Christchurch

1. Introduction

1.1 Background

Spatial data science is playing an increasingly significant role in urban planning, housing site selection, and accessibility analysis (Zhao, Zhong, & Gao, 2022). The importance lies in its ability to convert abstract, non-spatial urban data into intuitive geospatial information, quantitatively revealing and visualizing complex spatial patterns, thereby helping decision-makers better understand urban challenges, thereby supporting more accurate and evidence-based decisions (Goodchild, 2018). Meanwhile, the reduction of data acquisition costs and the developments in integrating advanced computational techniques have fostered further expansion in spatial data science applications. Benefiting from broad data-sharing platforms (such as OpenStreetMap) and fully functional programming environments (e.g., GIS libraries integrated in Python), people from diverse fields can conveniently conduct spatial analysis, thereby supporting more detailed urban studies (Kasperek & Podpora, 2024).

Specifically, accessibility assessment and site selection are among the most typical applications (Yang, Cao, Wu, Guo, Dong, & Tang, 2022; Loroño-Leturiondo, Hoang, & Cole, 2020). Different groups of people integrate multiple constraints according to their specific needs to identify suitable sites that satisfy all requirements. For example, international student families with children may prioritize access to educational resources (Timmermans, van der Waerden, Molin, & Arentze, 2012; Loroño-Leturiondo, Hoang, & Cole, 2020); commuters may focus on transportation convenience to ensure efficient connections between home and workplaces (Levinson, 1998); while urban planners may tend to evaluate the alignment between population distribution and public service facilities (Feng & Wang, 2025).

1.2 Problem Statement

In New Zealand, a portion of international student families with school-aged children (especially those who have newly arrived) often decide not to buy private vehicles due to the instability of their living situation. For them, housing site selection is constrained not only to prices but also highly relies on the accessibility of educational resources and public transportation (Corney, du Plessis, Woods, Lou, Dewhurst, & Mawren, 2024). This study focuses on families that have **primary** school-aged children.

Their residential choices should simultaneously meet two basic requirements: proximity to educational resources (including primary schools and the university) and public transportation (like bus stops). Unlike typical households, these families' high dependence on walking and public transportation significantly restricts their overall accessibility. The absence of these conditions can significantly reduce their quality of life and academic performance during the initial settlement period. Given their focus on both proximity to educational resources and public transportation accessibility, how can GIS-based spatial analysis quantitatively evaluate the housing suitability for international student families without private vehicles?

2. Research Objectives

The main objective of this study is to design and implement a housing site selection model tailored to the specific needs of international student families with primary school-aged children who do not own private vehicles. Specifically, the study aims to:

1) Build a spatial analysis framework to evaluate the housing suitability of international student families based on accessibility to education and public transport.
2) Apply this framework to identify candidate residential areas that satisfy all three constraints (e.g., proximity to primary schools, the University of Canterbury, and bus stops).
3) Classify and visualize the identified ideal residential areas according to their

level of suitability, providing explicit guidance for housing site selections.

3. Data and Methodology

3.1 Data Sources and Environment

Data Sources: All the data sets used in this project were obtained from open-sourced data, including Stats NZ, OpenStreetMap (OSM), and Wikipedia.

1) *SA2 Boundary Dataset:* Downloaded from Stats NZ and used to define statistical units across the study area.
2) *Meshblock Boundary Dataset:* Downloaded from Stats NZ and used to visualize finer administrative units.
3) *Christchurch Boundary:* Obtained from the OSM through OSMnx to constrain the selection of primary schools and bus stops, reducing the data volume.
4) *UC Campus Boundary & Ernest Rutherford Building:* Obtained from the OSM through OSMnx to demonstrate the campus of the University of Canterbury (UC) and the Ernest Rutherford Building.
5) *Walkable Road Network:* Extracted from the OSM through OSMnx, filtered to include only walkable roads in order to compute walking time and analyze accessibility to primary schools, the University of Canterbury, and bus stops.
6) *Bus Stops:* Obtained from the OSM through OSMnx.
7) *Primary schools:* Obtained from the OSM through OSMnx
8) *Primary school list in Christchurch:* Obtained from Wikipedia for correcting OSM primary school data.

Environment: The project was performed on the ThinkPad laptop (Windows 11 Pro, Intel i5, 16.0 GB RAM), and all the spatial analysis was conducted in Python 3.13.5 (Anaconda) using Jupyter Notebook in Visual Studio Code. These packages were necessary for this project: *NumPy, Pandas, Matplotlib, GeoPandas, OSMnx, Matplotlib-scalebar, JupyterLab, and Contextily.*

3.2 Data preprocessing

1) Uniform CRS: All spatial datasets were reprojected in Python to a common coordinate reference system (EPSG:2193, NZGD2000 / New Zealand Transverse Mercator 2000) for spatial analysis and visualization.

2) Geometry Standardization & Cleaning：For the Ernest Rutherford Building (in UC) and primary schools obtained from OSM, which were represented as a mixture of polygons, multipolygons, and points, all geometries were converted into points (using either centroids or representative points). Primary school names were standardized. As the OSM primary school dataset contained inaccuracies, including redundant schools that were closed, misclassified institutions, and incorrect names, it was merged with the Wikipedia list of Christchurch primary schools to produce a more accurate dataset.

3.3 Variable Selection

This study selected three variables that (1) are highly relevant to the needs of a newly arrived international student family, (2) can be obtained from open-source spatial data, and (3) can be consistently integrated to produce quantifiable outcomes. Accordingly, primary schools, the University of Canterbury (UC), and bus stops were chosen as the key factors. These variables reflect essential accessibility needs for both education and public transportation.

Since many international student families do not initially own private vehicles, and buses are not always convenient for commuting to the University of Canterbury or to primary schools, walking time was adopted as a measure of proximity in this study. Thresholds of 15 minutes to primary schools, 40 minutes to UC, and 15 minutes to bus stops were set based on urban accessibility benchmarks (Geurs & van Wee, 2004; Daniels & Mulley, 2011; Moreno et al., 2021).

These variables were used to generate isochrones, which were then weighted (0.50 for proximity to primary schools, 0.35 for proximity to UC, and 0.15 for proximity to bus stops) to compute suitability scores. The weighting scheme reflected the relative importance of each component. Primary school accessibility received the highest weight, due to fixed school hours and children's limited walking tolerance. University accessibility was ranked second, as academic study is the main aim for settlement and benefits from flexible schedules. Finally, public transport access was given the lowest weight.

3.4 Methodology

3.4.1 Study Area

To define the study area, a circular buffer with a radius of 2,880 m (equivalent to 40 minutes of walking time at 1.2 m/s) was generated around the Ernest Rutherford Building. This buffer was then intersected with the Statistical Area 2 (SA2) boundaries to ensure consistency with official statistical units. The resulting SA2s were adopted as the study area mask.

Since both primary and bus stop suitability zones were required to overlap with the UC Suitability zone to get the resulting suitability outcome, this clipping step can significantly reduce computational demands by excluding irrelevant parts of Christchurch. The raw OpenStreetMap (OSM) pedestrian road network data for the entire city is too large to be processed by personal laptops. Therefore, applying the UC 40-minute walking buffer and intersecting it with SA2 units to reduce analyzed data volume provided a computationally feasible condition.

3.4.2 Network Construction and Travel Time Computation

To prepare for the pedestrian network for subsequent accessibility analysis in Christchurch, raw OSM road data were filtered to retain only walkable edges (e.g., footways or steps). The length (in meters) of each edge was computed using OSMnx. A walking speed of 1.2 m/s, consistent with urban accessibility studies (Giannoulaki & Christoforou, 2024), was assumed to derive the travel_time attribute for each edge, calculated as *travel_time = length / 1.2* (in seconds), forming the basic pedestrian network.

3.4.3 Network-based Accessibility Analysis

Using this pedestrian network, walking accessibility to primary schools, the University of Canterbury (UC), and bus stops were respectively evaluated through a 25-meter grid-based network analysis.

1) UC accessibility (single-source shortest path):
   A 25 m resolution grid was generated within the study area, with each grid centroid snapped to its nearest pedestrian network node. The Ernest Rutherford Building served as the origin, from which the shortest walking times to all nodes were computed using the travel_time attribute. Isochrones were demonstrated using a 40-minute threshold and classified into five bands (0–5, 5–10, 10–20, 20–30, and 30–40 minutes).
2) Primary school accessibility (multi-source shortest path):
   The same 25 m resolution grid was generated, and the grids' centroids were linked to their nearest pedestrian network nodes. All primary schools were represented by their nearest nodes and treated as multiple origins. The *multi_source_dijkstra_path_length* function was applied to compute the shortest walking times from all nodes to their nearest primary school. A 15-minute threshold was applied, generating isochrones which were then classified into three bands (0–5, 5–10, and 10–15 minutes).
3) Bus stop accessibility (multi-source shortest path):
   The same procedure as primary schools was followed for bus stops using a 15-minute threshold, producing three isochrone bands (0–5, 5–10, 10–15 minutes).

All the procedures were implemented in Python using OSMnx and NetworkX libraries.

3.4.4 Overlay analysis

To identify candidate areas in Christchurch that simultaneously satisfy three accessibility constraints, the 40-minute UC grid, 15-minute primary school grid, and 15-minute bus stop grid (derived from Section 3.3.3) were standardized and scored. Spatial intersection was conducted using the overlay function to combine these three layers into a single candidate area layer which meet all the conditions.

Walking times in each grid were normalized to 0 to 1, with short times corresponding to high scores. A weighted sum was then applied to compute the final composite suitability score:

$$Total\_score = w_s \times school_{norm} + w_u \times UC_{norm} + w_b \times bus_{norm}$$

where $w_s = 0.50$, $w_u = 0.35$, $w_b = 0.15$, representing the weights of primary schools, UC, and bus stops, respectively.

The composite scores were classified into five suitability categories (Very Low, Low, Mid, High, Very High) using quantile thresholds (0.20, 0.40, 0.60, 0.80). Polygons were dissolved to improve cartographic clarity.

3.5 Statistical Techniques

First, the area of suitability regions within each SA2 unit was computed, categorized by suitability grades (Very Low to Very High). The total suitability area and high-suitability area (including 'High' and 'Very High' regions) were then summarized to detect which SA2 units contain the most favorable locations. These descriptive statistics provide a quantitative basis for comparing spatial variation across SA2 units within the study area. Although additional statistical analysis, such as linking suitability regions with corresponding housing prices, can be explored in the future, this study focused on the suitability areas' spatial coverage and distribution patterns.

3.6 Cartography and Visualization Approach

All maps were presented as choropleth maps to visualize the analytical outcomes, emphasizing both accuracy and cartographic clarity. Sequential ColorBrewer palettes were applied to reflect accessibility gradients. Standard cartographic elements, including legends, north arrows, scale bars, and inset maps, were incorporated for better readability. Coordinates labels were intentionally omitted, as the focus was on the spatial distribution of accessibility regions rather than precise georeferencing. Administrative boundaries (SA2) and key landmarks (e.g., the University of Canterbury, the Ernest Rutherford building, and primary schools) were annotated to provide contextual reference. All maps were produced using Python.

3.7 Workflow Overview

This study follows a seven-step workflow, as shown in Figure 1.

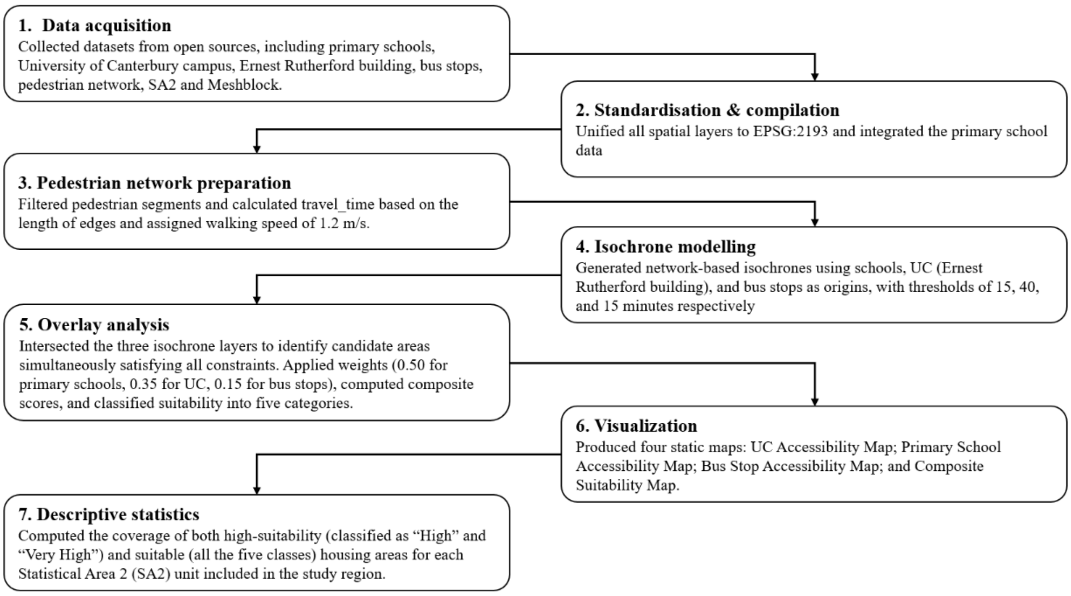

Figure 1. Analytical workflow for spatial accessibility and housing suitability assessment in Christchurch.

4. Results

4.1 Map Visualization of UC Accessibility

The classified UC accessibility isochrones form irregular concentric ellipses centered on UC, with the major axis oriented northwest–southeast and the minor axis northeast–southwest. Accessibility is highest around Ilam, Jelly Park, and Deans Bush, declining sequentially the farther away from UC, as shown in Figure 2. A notable discontinuity appears in the southern sector of UC, where the isochrones deviate from the expected concentric pattern. This is due to the missing pedestrian roads in the OSM data. In this study, these issues were not corrected due to the large amount of verification and editing work. However, it underscores that the accuracy of accessibility outcomes is highly dependent on the completeness and quality of the road network data.

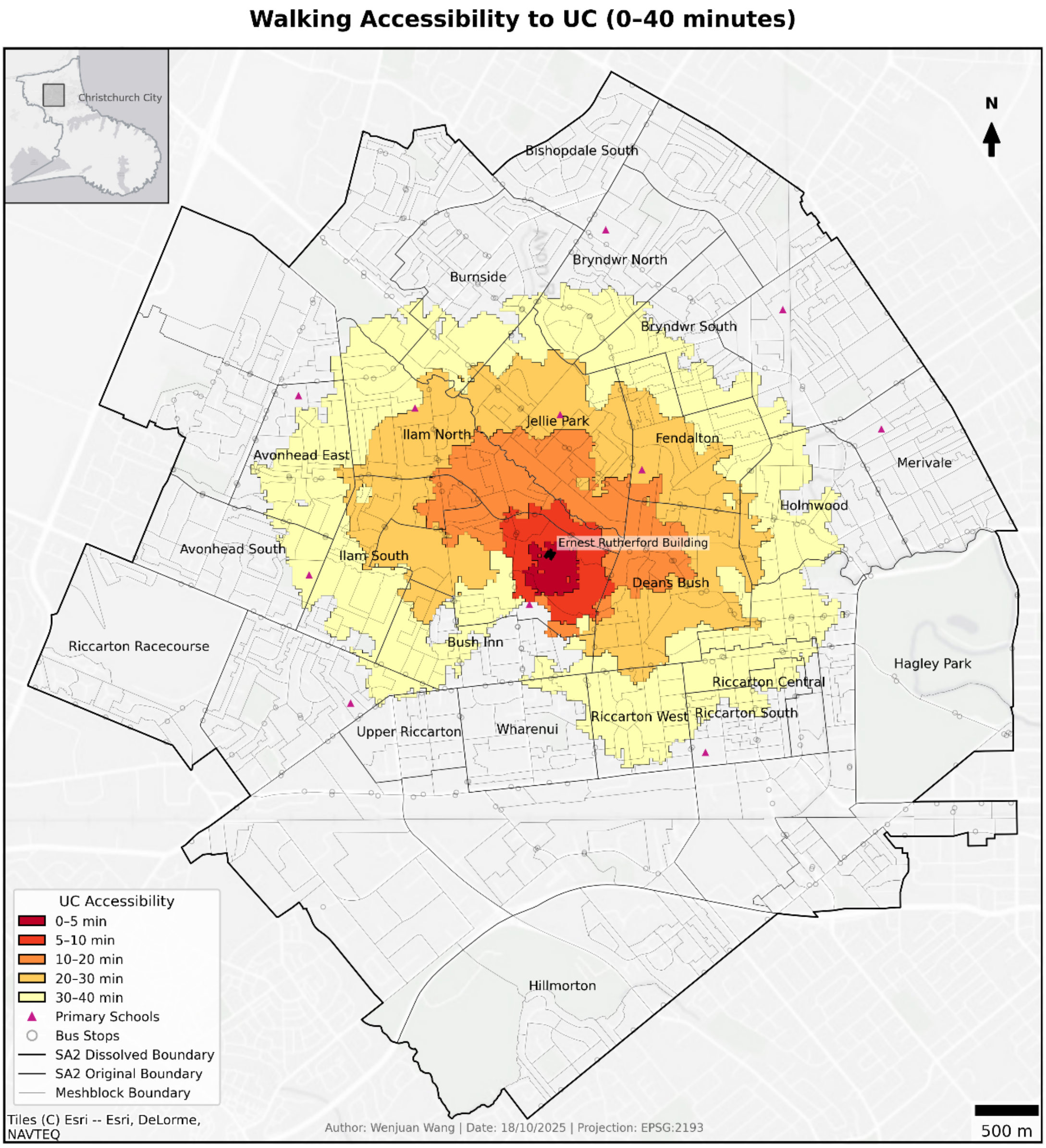

Figure 2. Walking Accessibility to the University of Canterbury (0–40 minutes)

4.2 Map Visualization of Primary School Accessibility

The classified primary school accessibility isochrones generally form clusters centered on individual school sites. As illustrated in Figure 3, in the northeast, north, northwest, and west of UC, where primary schools are relatively dense, the individual isochrones overlap and merge into a semi-elliptical ring. Additional clusters appear along the northeastern and southeastern edges of the study area; however, since these fall outside the UC accessibility isochrones, they were not included in further consideration.

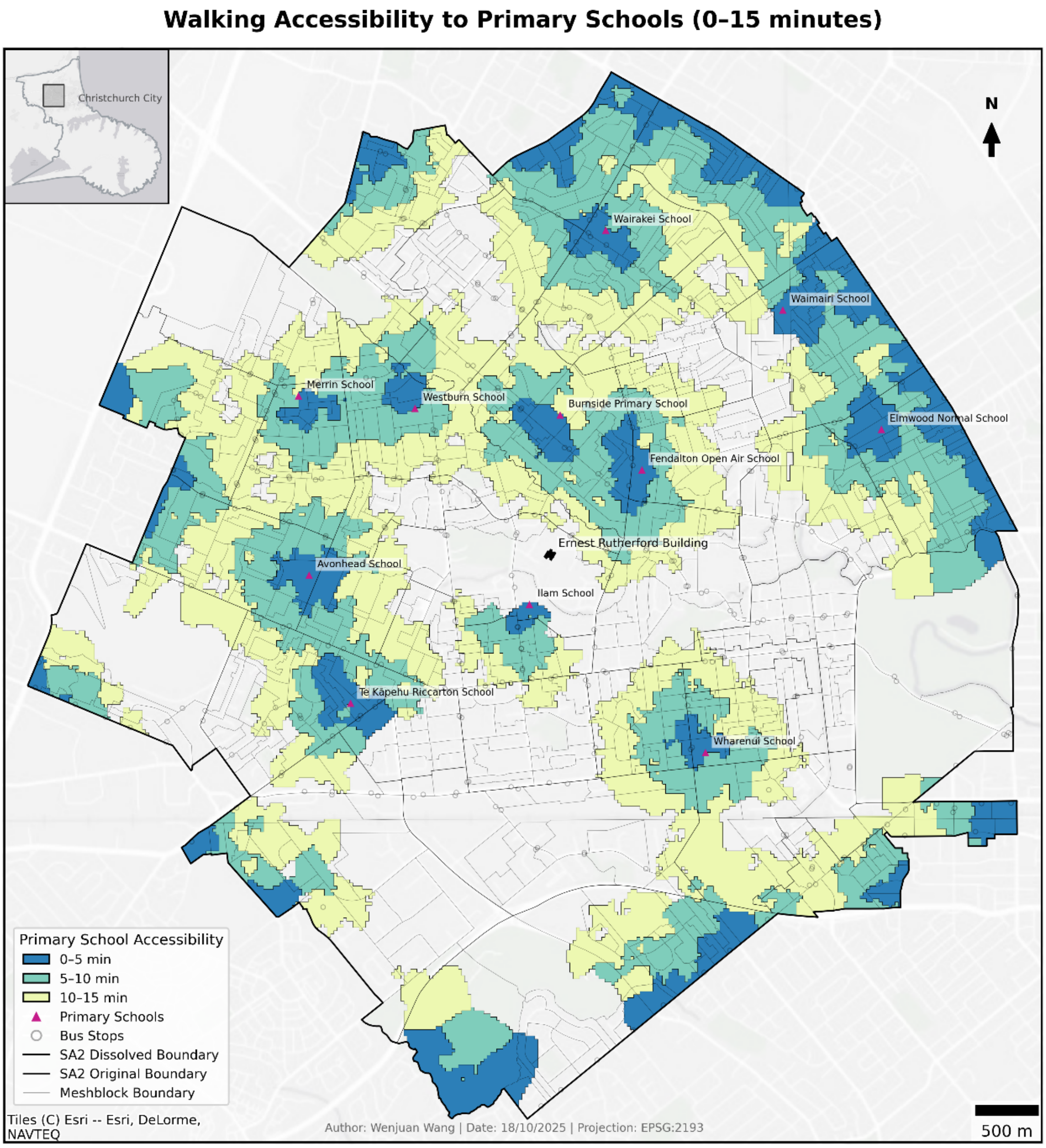

Figure 3. Walking Accessibility to Primary Schools (0–15 minutes).

4.3 Map Visualization of Bus Stop Accessibility

As shown in Figure 4, the classified bus stop accessibility isochrones almost cover the entire study area, with high-accessibility zones (0-5 minutes) concentrated along major roads. Accessibility declines progressively with increasing distance from the roads. Only limited gaps remain, primarily in the southernmost parts of the study area, such as Sockburn South and Hillmorton, where fewer stops result in slightly longer walking distances.

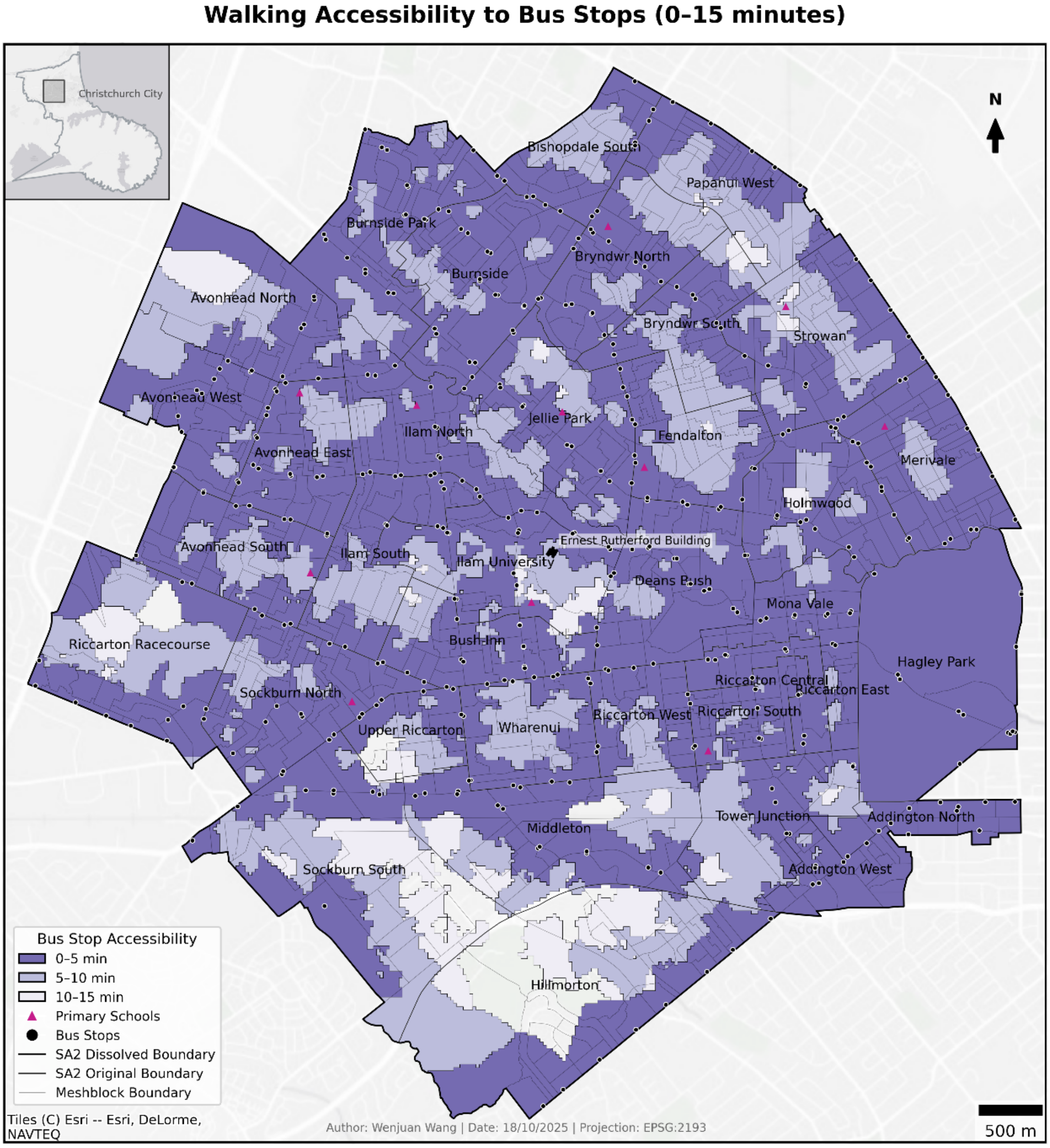

Figure 4. Bus Stop Walking Accessibility Map

4.4 Final Composite Suitability Map

As shown in Figure 5, the high-suitability areas are concentrated to the northeast, north, west, and northwest of UC campus, particularly in Fendalton, Jellie Park, and Ilam, with smaller clusters in Avonhead and Riccarton. Mid-, low-, and very low-suitability regions surround these core areas. The overall pattern closely resembles the primary school accessibility isochrones, forming a semi-ring of suitability around UC.

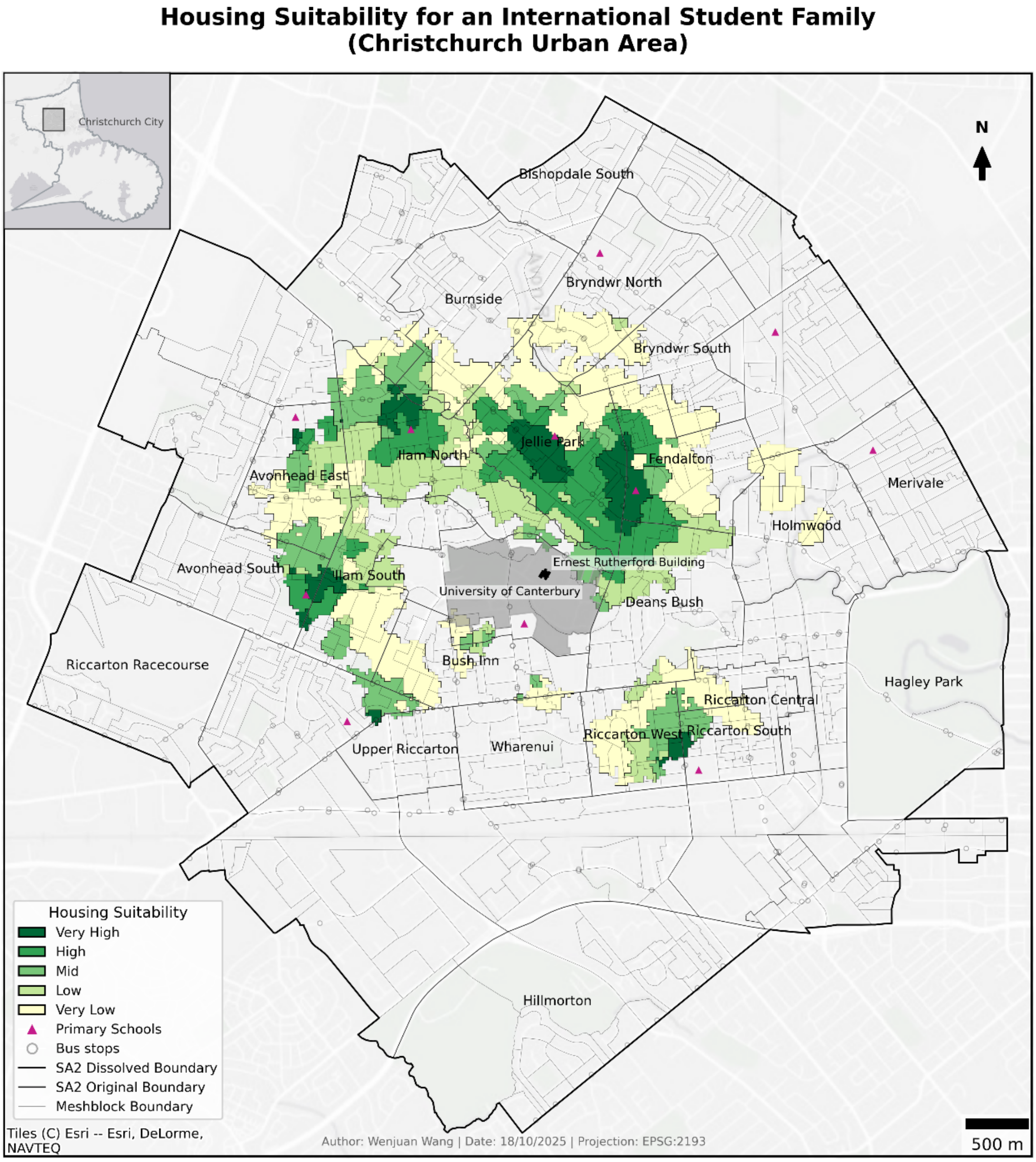

Figure 5. Composite Suitability Map.

4.5 High-Suitability Areas by SA2

Table 1 summarizes the ten SA2s with the largest high-suitability areas (classified as High and Very High), ranked by total area of high and very high-suitability. The results indicate that Jellie Park contains the largest concentration, with around 731,147 m² of high-suitability land, followed by Ilam North (259,853 m²) and Fendalton (218,734 m²). Together, these three SA2s account for the majority of the most suitable residential areas, reflecting their proximity to UC, primary schools and bus stops. Limited pockets are found in Deans Bush, Ilam South, Avonhead South, Avonhead East, Riccarton West, Riccarton South, and Burnside (ranging from 130,153–9,908 m²).

Table 1.

Statistical summaries of high-suitability areas by SA2.

SA2 Name	Total high area (m²)	High (m²)	Very high (m²)
Jellie Park	731,147	466,793	264,354
Ilam North	259,853	190,203	69,650
Fendalton	218,734	140,631	78,103
Deans Bush	130,153	119,485	10,668
Ilam South	94,587	57,565	37,022
Avonhead South	75,409	43681	31,728
Avonhead East	47,922	40000	7,922
Riccarton West	23,063	0	23,063
Riccarton South	11,937	0	11,937
Burnside	9,908	0	9,908

5 Discussion

The results demonstrate that the spatial distribution of suitability areas is primarily shaped by the locations of primary schools, while is also influenced by the accessibility to UC and bus stops. This is partly due to the higher weight assigned to primary schools, which makes them dominant in determining the composite suitability. Additionally, well-connected road networks in these regions also contribute to high suitability by minimizing detours and thereby reducing walking times.

The results broadly align with the residential choices of a number of international student families in Christchurch, who tend to live near UC such as Ilam, Riccarton, and Fendalton. These areas are widely perceived as convenient, safe, and well-served by schools and public transport, which corresponds with the high-suitability clusters identified in the model.

There are several limitations in this study. First, the completeness of OSM data is uncertain, which may affect the accuracy of walking-time calculations. Second, the assumption of a uniform walking speed (1.2 m/s) is a simplification; in reality, children and adults may walk at various speeds, altering accessibility outcomes. Third, the model focuses solely on walking accessibility and does not incorporate housing cost, rental availability, safety and other factors, all of which are critical in real-world housing decisions.

6 Conclusions

This study develops a housing site selection model for international student families with primary school-aged children attending the University of Canterbury, assuming no access to private vehicles. The model builds a GIS-based, road network-driven spatial analysis framework implemented with Python techniques and OpenStreetMap data. It quantitatively evaluates residential suitability by integrating three critical determinants: walking time to primary schools, the University of Canterbury, and bus stops. The results indicate that northern parts of UC contain more high-suitability areas than the remaining parts. The research demonstrates the value of spatial data science in supporting housing site selection, offering a practical approach for decision-makers. Future work could incorporate housing costs, safety and other socioeconomic factors to enhance the model.

7. Gen-AI acknowledgment

This proposal made limited use of Generative Artificial Intelligence tools in accordance with course guidelines. ChatGPT and Grammarly were used for minor tasks, including proofreading, improving grammar and readability, and checking APA reference formatting.

Prompts/questions to ChatGPT included requests such as:

“What is the correct format of APA for this article?”

“Check if there is a reference cited in the proposal but not listed in the references.”

“Check if there are spelling or grammar mistakes.”

“Refine this paragraph to reduce translation awkwardness while keeping the meaning unchanged.”

I carefully evaluated the accuracy of ChatGPT’s suggestions. Only language refinements and formatting adjustments were adopted, without content generation or data analysis by AI, ensuring all research ideas and interpretations remain my own.

8. References

Corney, T., du Plessis, K., Woods, B., Lou, C., Dewhurst, A., & Mawren, D. (2024). 'If you are feeling alone and you are not feeling safe, it impacts everything': A mixed-methods exploration of international students' accommodation, subjective wellbeing and mental health help-seeking. *BMC Public Health, 24*, 1262. https://doi.org/10.1186/s12889-024-18122-0

Daniels, R., & Mulley, C. (2011). Explaining walking distance to public transport: The dominance of public transport supply. *World Symposium on Transport and Land Use Research.* https://ses.library.usyd.edu.au/handle/2123/7814

Feng, Y., & Wang, Y. (2025). Study on the spatial matching between public service facilities and the distribution of population—An example of Shandong Province. *Sustainability, 17*(17), 7866. https://doi.org/10.3390/su17177866

Geurs, K. T., & van Wee, B. (2004). Accessibility evaluation of land-use and transport strategies: Review and research directions. *Journal of Transport Geography, 12*(2), 127–140. https://doi.org/10.1016/j.jtrangeo.2003.10.005

Giannoulaki, M., & Christoforou, Z. (2024). Pedestrian walking speed analysis: A systematic review. Sustainability, 16(11), 4813. https://doi.org/10.3390/su16114813

Goodchild, M. F. (2018). Geospatial analysis and decision-making in urban planning: A review of GIS applications. *Annals of GIS, 24*(2), 83–93. https://doi.org/10.1080/19475683.2018.1510911

Kasperek, D., & Podpora, M. (2024). Optimizing geospatial data for ML/CV applications: A Python-based approach to streamlining map processing by removing irrelevant areas. *Applied Sciences, 14*(24), 11978.

https://doi.org/10.3390/app142411978

Levinson, D. (1998). Accessibility and the journey to work. *Journal of Transport Geography, 6*(1), 11–21. https://doi.org/10.1016/S0966-6923(97)00036-7

Loroño-Leturiondo, M., Hoang, A., & Cole, D. (2020). International students' housing choices in the UK: A qualitative study. *Journal of International Students, 10*(3), 723–742. https://doi.org/10.32674/jis.v10i3.1198

Moreno, C., Allam, Z., Chabaud, D., Gall, C., & Pratlong, F. (2021). Introducing the "15-minute city": Sustainability, resilience and place identity in future post-pandemic cities. *Smart Cities, 4*(1), 93–111. https://doi.org/10.3390/smartcities4010006

Timmermans, H. J. P., van der Waerden, P., Molin, E., & Arentze, T. (2012). Residential location choices of immigrant households: A discrete choice model with endogenous income. *Journal of Transport Geography, 22*, 1–8. https://doi.org/10.1016/j.jtrangeo.2011.12.004

Yang, X., Cao, Y., Wu, A., Guo, M., Dong, Z., & Tang, L. (2022). Accessibility assessment of buildings based on multi-source spatial data: Taking Wuhan as a case study. *ISPRS International Journal of Geo-Information, 11*(12), 631. https://doi.org/10.3390/ijgi11120631

Zhao, Q., Zhong, C., & Gao, S. (2022). GIS and urban data science. *Annals of GIS, 28*(4), 347–359. https://doi.org/10.1080/19475683.2022.2070969