The Wheels of Life

Yin Junren, Ryan, Chow Xing Yu, Wong Yee Hern
Hwa Chong Institution

Abstract

This project investigated three questions about cycling infrastructure in Bukit Panjang and Bukit Batok. Firstly, how to optimise environmental factors that cyclists find important in the cycling path? Secondly, how to determine the most efficient placement of cycling path? Thirdly, how to determine the most efficient placement of bicycle parking lots?

We first identified the four most important environmental factors. Fuzzy membership functions were used to generate suitability values for the footpath based on each factor. The suitability maps were combined using Weighted Linear Combination.

Next, graph theory was applied to the combined suitability map to derive the most connected segments of the footpath to nearby facilities that also best fulfill the four environmental factors, generating a map of the optimal cycling path.

Finally, suitability maps were plotted based on the mean proximity of each bicycle parking lot to three nearest facilities. The parking lots were modelled as a graph. The average local and weighted clustering coefficients were calculated and compared to determine the level of clustering.

We have found that the usage of existing cycling path is not maximised in relation to nearby facilities, and some segments are not optimal when analysed based on the four environmental factors.

Almost all yellow boxes in Bukit Batok are positioned close to a facility. The average local clustering coefficient was calculated to be 0.690, while the average weighted clustering coefficient was 0.590, indicating that some level of clustering exists.

Our approach allows transport planners to ensure cycling paths and bicycle parking lots are built in environmentally favorable areas, while also maximising their placement based on proximity to nearby facilities so as to reduce space wastage in a land-constrained Singapore.

1. Introduction

Singapore's transport network has been recognized globally as a leading model in providing a wide range of transport options for its people that are easily accessible from all around the island. Singapore has done well in establishing an interconnected bus and MRT network, but with burgeoning private car usage, it continues to face problems of traffic congestion on its roads especially during peak periods, leading to increased air pollution and reduced productivity. This raises the need for Singapore to search for alternative transport options for its people, and one of such solutions is cycling.

Cycling is deemed to be not only convenient, but also effective in bringing riders closer to nature and outdoors (Ministry of Transport, 2020). However, due to the lack of connectivity of cycling routes in Singapore, not many people would choose to cycle as a means of transport in their daily lives. Considering Singapore's land constraints, planning the most effective cycling route is critical. Hence, our group decided to embark on this topic to devise the best cycling routes and positioning of bicycle parking spaces, such that they benefit people while using minimal space and resources, so as to provide the necessary facilities cyclists need and promote cycling as an alternative to vehicles. This project will be focusing on the Bukit Batok and Bukit Panjang planning areas in the West district, as the number of cycling paths there is the fewest compared to other areas like Jurong East. Furthermore, as the West district has always been regarded as the country's largest manufacturing hub, it becomes more important to provide more interconnected cycling paths and ensure a sufficient number of bicycle parking lots in the district so that travel to work and MRT stations becomes more convenient, as such enhancing work productivity.

2. Terminology

Consistency Ratio	Consistency Index for the set of judgments divided by the Index for the corresponding random matrix. If the ratio exceeds 0.1, the set of judgments may be too inconsistent to be reliable (Saaty, 1990).
Fuzzy Membership	A method that transforms the input data to a 0 to 1 scale based on the possibility of being a member of a specified set.

Geographic Information System (GIS)	A framework that analyzes spatial location and organizes information into visualizations using maps and 3D scenes.
Graph Theory	Mathematical structures that model pairwise relations between objects. A graph is made up of vertices (also called nodes) which are connected by edges (also called lines).
Multi - Criteria Decision Analysis (MCDA)	A process of determining the optimal alternative, considering multiple, conflicting and interactive criteria (Chen et al., 1992).
Paired t-tests	A statistical procedure used to determine whether the mean difference between two sets of observations is zero. Common applications of the paired sample t-test include case-control studies or repeated-measures designs.
Suitability Analysis	The search for locations that are characterized by a combination of certain properties. Often, the result is a suitability map showing the locations suitable for a specific use in the form of a thematic map.
Weighted Graph	A graph in which each edge is given a numerical weight.
Weighted Linear Combination	A process of calculating the total score of an alternative by taking the weighted average of the suitability values of all attributes (Drobne and Lisec, 2009).

3. Research Questions

3.1 How to optimise the environmental factors that cyclists find important in the proposed model of the cycling route?

3.2 How to determine the most efficient placement of cycling paths that meets cyclists' travelling needs while not taking up unnecessary space?

3.3 How to ensure the most efficient placement of bicycle parking lots such that they are sufficient enough to meet cyclists' needs, while not overly clustered together such that some become unused?

4. Literature Review

Shin and Kai (2018) embarked on a case study of the Woodlands Planning Area (WPA) in Singapore. A GIS-MCDA framework was used that gathered the weightings of nine criteria from the members of the public, transport experts and government planners, after which combining suitability maps of the nine criteria based on their weightings to form a preference map for each stakeholder. A suitability map that combined the three stakeholders' preference maps in equal proportion was derived, from which suitable areas for building cycling paths were determined.

Similarly, Koh and Wong (2013) asked members of the public to rate the level of importance of 11 environmental factors in affecting their willingness to select a cycling route, as well as draw on a map, the usual route taken from the station to his/her destination. For each route, the shortest possible route was identified and points were awarded to each of them based on the 11 factors. Paired t-tests were then derived to find out the reasons why a person would choose his/her route over the shorter one, after which the extent of influence of each factor is analysed.

Oh and Jeong (2001) investigated the usefulness of the GIS - fuzzy set approach as compared to a crisp approach in the evaluation of environmental conditions of a residential area. Specific factors under each objective were identified. First, a crisp approach, that awards two or three discrete values between zero and one, was used to evaluate the area. Next, a Fuzzy set approach was used for the same criteria. The correlation coefficients for the mean suitability output under the two approaches were compared. It was found that the fuzzy approach determined suitability values more precisely and displayed more diverse qualitative levels.

5. Methodology

5.1 RQ1: *How to optimise the environmental factors that cyclists find important in the proposed model of the cycling route?*

Firstly, the four environmental factors cyclists find most important in their choice of cycling path were identified. With reference to such data of the significance level of each factor, the pairwise comparison method was used to derive a measurement of the relative weighting of each factor from a value of zero to one. The consistency ratio was checked to ensure the precision of our measurements (Appendix B1). Next, data for the four environmental conditions in the Bukit Batok and Bukit Panjang planning areas were retrieved from government data sources (Appendix

A1). Criteria for each of the environmental factors was determined and fuzzy membership functions were developed using RStudio that standardize the suitability values between zero and one (Appendix B2). After inputting the data of environmental factors, suitability maps were plotted, assessing how well environmental conditions meet the criteria (Appendix B3). Weighted Linear Combination was then used to combine the suitability maps for each factor into one complete map, representing a weighted average of the suitability values of the network (Appendix E2).

5.2 RQ2: *How to determine the most efficient placement of cycling paths that meets cyclists' travelling needs while not taking up unnecessary space?*

Python was first used to plot the edges and nodes of the graph, which represents the footpath segment. The graph of the footpath network was overlaid with the combined suitability map (Appendix C1). Points were then given to each node on the graph based on the overall suitability value of the footpath segment the node is located at. Next, the datasets of housing areas and of facilities was obtained and filtered (Appendix A2). Both datasets were overlaid onto the graph of the footpath network. A nearest node was chosen for each housing area and for each of the facilities, after which a shortest path was generated from each of M housing areas to each of N facilities using Dijkstra's algorithm (Appendix E1) in Python' networkx module. As a result, $M \times N$ paths were generated. They were layered together to obtain a frequency graph, and a map of the most efficient placement of cycling paths was then obtained by filtering the graph with the median value of all the values in the graph (Appendix C2).

5.3 RQ3: *How to ensure the most efficient placement of bicycle parking lots such that they are sufficient enough to meet cyclists' needs, while not overly clustered together such that some become unused?*

The dataset of Bicycle Parking lots was first taken from Land Transport DataMall. The bicycle parking facilities this project focused on were the yellow boxes within the two planning areas. The coordinates of yellow boxes were filtered out from the other bicycle parking facilities as a csv file, which was plotted and analysed in RStudio (Appendix D1). Suitability of placement of yellow boxes in the planning areas was analysed, in relation to surrounding facilities, and the shortest distance between each facility and any yellow box was calculated. An analysis of the suitability of placement of each individual yellow box was also conducted (Appendix D2). To analyse the connectivity of yellow boxes, graph theory was applied by taking the yellow boxes as the nodes, and connecting each yellow box to the four nearest neighbours to form a weighted graph (Appendix D3). The average local clustering coefficient and average weighted clustering coefficient of the graph were then calculated to analyse the connectivity of the network formed (Appendix E3).

6. Results

6.1 RQ1: *How to optimise the environmental factors that cyclists find important in the proposed model of the cycling route?*

Weightage of factors were proven acceptable and reliable by the use of pairwise comparison. Fuzzy membership functions were used to determine values of suitability based on different environmental factors, and the values were successfully used to generate suitability maps. Weighted Linear Combination was applied to combine all four suitability maps into one complete map which was further analysed in RQ2.

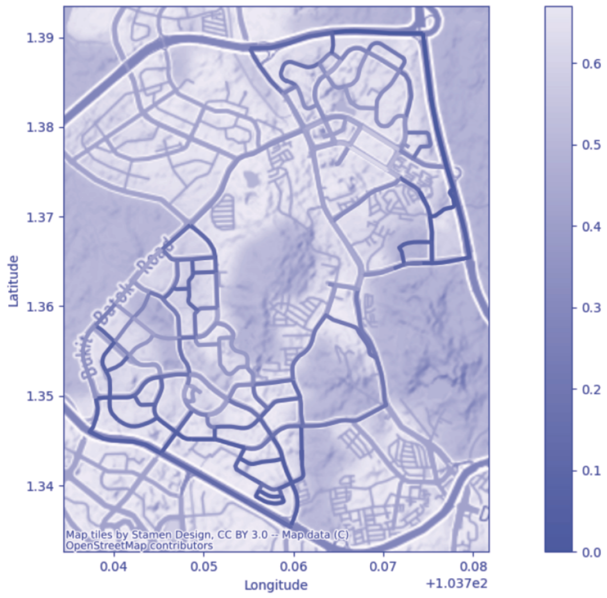


Figure 1. Combined Suitability Map of four environmental factors in Bukit Batok and Bukit Panjang. Made with Python.

6.2 RQ2: *How to determine the most efficient placement of cycling paths that meets cyclists' travelling needs while not taking up unnecessary space?*

Graph theory was applied to the combined suitability map from RQ1. Based on the overall suitability value of the footpath segment the node is located at and potential demand for cycling indicated by external datasets like housing and facilities, points were allocated to the nodes and edges.

The best cycling path was then determined, that optimises the environmental factors from the combined suitability map, and maximises usage of the path in relation to surrounding facilities.

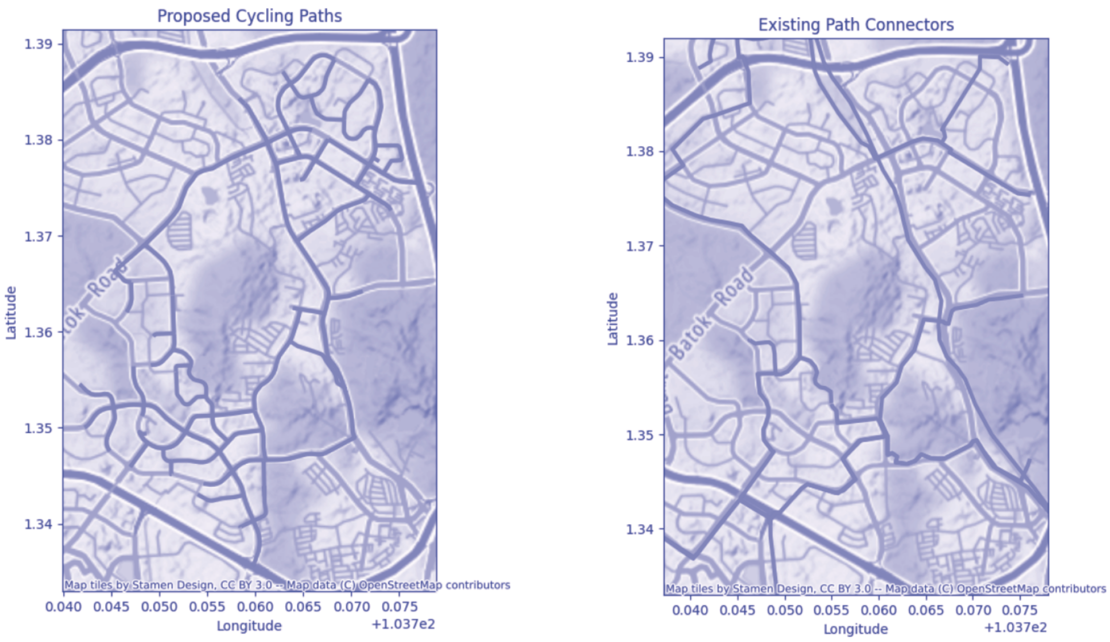


Figure 2. Proposed model of cycling paths (left) and existing park connectors (right) in Bukit Batok and Bukit Panjang. Made with Python.

6.3 RQ3: *How to ensure the most efficient placement of bicycle parking lots such that they are sufficient enough to meet cyclists' needs, while not overly clustered together such that some become unused?*

After measuring the shortest distance from each facility to any bicycle lot and the distance from each bicycle parking lot to the three nearest facilities, the suitability of placement of each individual bicycle parking lot was analysed.

It can be concluded that apart from the parking lots in the outskirts of the planning areas, almost all the bicycle parking lots are positioned close to a facility.

After graph theory has been applied, the average local clustering coefficient was calculated to be 0.690 while the average weighted clustering coefficient was calculated to be 0.590 in Bukit Batok. It has been found that the network of bicycle parking lots in Bukit Batok is one in which there is minimal distance between the bicycle parking lots, indicating some level of clustering.
Thus, this has shown that suitability map plotting must be done in tandem with graph theory to conduct a complete analysis of positioning of bicycle parking lots in terms of proximity to each other, and to nearby facilities.

7. Conclusion

In this project, a GIS-Fuzzy set approach was taken to generate suitability maps of the footpath network based on each individual segment's performance in four environmental criteria. Graph theory was then applied to ensure that segments of the footpath with highest suitability values are utilised, while at the same time have their usage maximised in relation to surrounding community facilities, so as to derive an efficient cycling route that meets cyclists' needs while minimising space wastage. The GIS method was similarly used to generate suitability maps of yellow boxes based on their proximity to surrounding facilities, and graph theory was applied to measure the level of clustering of yellow boxes in the area.
The GIS-Fuzzy set approach, combined with graph theory, have been able to produce realistic results in terms of the optimal cycling path and suitability of placement of yellow boxes, though we have not been able to prove the effectiveness of this approach. Nonetheless, our approach can be considered by transport planners as a thorough approach to not only ensure cycling routes and bicycle parking facilities are built in environmentally favorable areas, but also to maximise placement of bicycle parking spaces based on proximity to surrounding facilities so as to ensure minimal space wastage for the cycling path, and minimal clustering for parking facilities.
One limitation of our project was that the needs and interests of cyclists who travel for recreational purposes have not been considered, as our project assumes people cycle only for commuting purposes. Future versions of our method can include the characteristics of cycling paths recreational cyclists are more inclined to, such as the use of certain bicycle facilities (Larsen and El-Geneidy, 2013), to produce a model that caters to a wider range of cyclist needs. In addition, not all environmental factors or types of surrounding facilities may have been considered in our generation of suitability maps, reducing the reliability of suitability values.

8. References

American Association of State Highway and Transportation Officials. (2018). *A policy on geometric design of highways and streets 7th edition*. Washington, DC. Available: ftp://www.ahtd.state.ar.us/Outgoing/Roadway/TRC1805/AASHTO%20A%20Policy%20on%20Geometric%20Design%20of%20Highways%20and%20Streets%202018,%207th%20Edition.pdf

Barrat A., Barthelemy M., Pastor-Satorras R., Vespignani A. (2004). The architecture of complex weighted networks. Barcelona, Spain. Available: https://doi.org/10.1073/pnas.0400087101 .

Chua H.X. (2017). *Train Stations in Singapore*. [On-Line]. Available: https://data.world/hxchua/train-stations-in-singapore. Retrieved May 2, 2020.

Data.gov.sg. (2017). *Listing of Licensed Supermarkets*. [On-Line]. Available: https://data.gov.sg/dataset/listing-of-licensed-supermarkets. Retrieved May 16, 2020.

Datamall. (2020). *Geospatial Whole Island*. [On-Line]. Available: https://www.mytransport.sg/content/mytransport/home/dataMall/search_datasets.html?searchText=Geospatial. Retrieved May 16, 2020.

Datamall. (2020). *Kerbline*. [On-Line]. Available: https://www.mytransport.sg/content/mytransport/home/dataMall/search_datasets.html?searchText=kerbline. Retrieved May 16, 2020.

Data.gov.sg. (2020). *LTA Bicycle Rack*. [On-line].Available: https://data.gov.sg/dataset/lta-bicycle-rack. Retrieved 10 July, 2020.

Koh P. P., Wong Y. D. (2013). *Influence of infrastructural compatibility factors on walking and cycling route choices*. Singapore. Available: https://doi.org/10.1016/j.jenvp.2013.08.001.

Kapasny. (2016). *Data Science of Visiting Famous Movie Locations in San Francisco.* [On-Line]. Available: https://www.kdnuggets.com/2016/07/visiting-famous-movie-locations-san-francisco.html . Retrieved March 2, 2020.

Land Transport Authority. (2019). *Transport Impact Assessment Guidelines for Developments.* Singapore. Available: https://www.lta.gov.sg/content/dam/ltagov/industry_innovations/industry_matters/develo pment_construction_resources/Street_Work_Proposals/Guidelines/TIA_Guidelines_WCP_Final_07Dec2018.pdf.

Land Transport Authority. (2020). *Cycling Routes*. [On-line]. Available: https://www.lta.gov.sg/content/ltagov/en/map/cycle.html. Retrieved 10 July, 2020.

Larsen, J., El-Geneidy, A. (2013). *A travel behaviour analysis of urban cycling facilities in Montreal Canada*. Canada. Available: http://dx.doi.org/10.1016/j.trd.2010.07.011.

Oh K., Jeong Y. (2001). *The usefulness of the GIS - fuzzy set approach in evaluating the urban residential environment*. Republic of Korea. Available: https://doi.org/10.1068/b2779.

Peidro D., Vasant P. (2010). *Transportation planning with modified S-curve membership functions using an interactive fuzzy multi-objective approach.* Spain. Available: https://doi.org/10.1016/j.asoc.2010.10.014.

Shin H. T., Kai C. (2018). *GIS-MCDA based cycling paths planning: a case study in Singapore.* Singapore. Available: https://doi.org/10.1016/j.apgeog.2018.03.007.

Saaty T. L. (1990). How to make a decision: The Analytic Hierarchy Process. *European journal of operational research*, 48, pg. 9-26. Available: https://doi.org/10.1016/0377-2217(90)90057-I

Urban Redevelopment Authority. (2018). *Walking and Cycling Design Guide*. Singapore. Available: https://www.ura.gov.sg/Corporate/Guidelines/-/media/BD725DB201DB496A93569C8072DD9FD0.ashx

Watts, D., Strogatz, S. (1998). Collective dynamics of ‘small-world’ networks. *Nature,* 393, pg.440–442. https://doi.org/10.1038/30918

9. Appendix

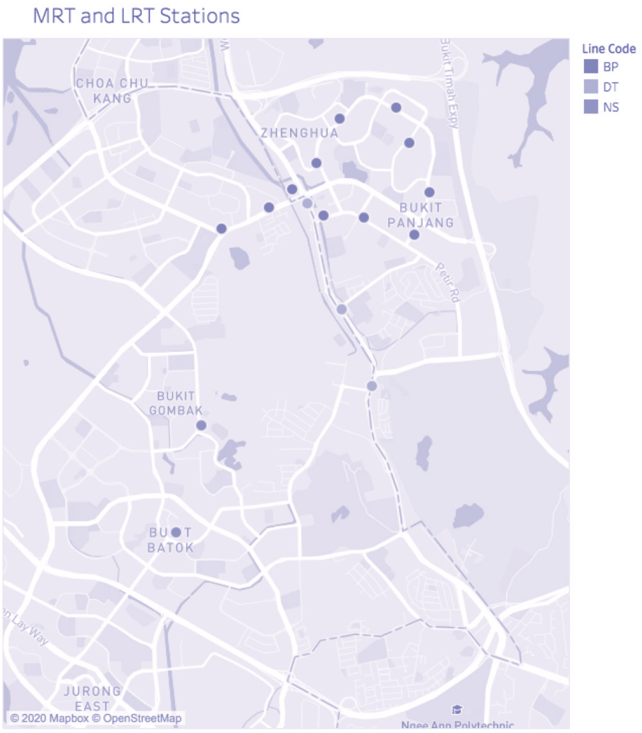
Table of Contents

Appendix A: Data Sources and Data Collation

Appendix A1: Environmental Factors for RQ1

A dataset of the footpaths in Singapore was first retrieved from Datamall, and serve as the model for which suitability values will be calculated for. To obtain the suitability values for the footpath, data must be obtained for the respective environmental factors. This was done through API calls from Datamall for the Major Roads dataset and Shops Dataset, as well as from data.world for the locations of MRT and LRTs. The elevation values in the slopes dataset was obtained by interpolating data from a SRTM Digital Elevation Model using Python libraries “scipy” and “rasterio”, while the distance value was calculated with EPSG code 3414, a geographical coordinate system with meters as units.

Figure 3. Visualization of Footpaths data. Made with Tableau.


MRT and LRT Stations
Line Code
BP
DT
NS
CHOA CHU KANG
ZHENGHUA
BUKIT PANJANG
Bukit Timah Expy
Petir Rd
BUKIT GOMBAK
BUKIT BATOK
JURONG EAST
© 2020 Mapbox © OpenStreetMap


Shops
Business Name
COLD STORAGE
GIANT
HAO MART
HILLVIEW MARKET PLACE
MCP FAIRMART
NTUC FAIRPRICE
PRIME SUPERMARKET
SHENG SIONG
THE SEAFOOD MARKET PLACE BY SONG FISH
YEO EXPRESS
ZENMART
CHOA CHU KANG
ZHENGHUA
BUKIT PANJANG
BUKIT GOMBAK
BUKIT BATOK

Figure 4. Visualization of data for four environmental factors. Made with Tableau.

Appendix A2: Housing and Facilities Dataset for RQ2

Th dataset of housing areas was taken from Data.gov.sg where housing areas in the area of focus were filtered out using Python. The dataset of facilities, namely Community Institutions, Healthcare facilities, and Sports and Recreational facilities was obtained and filtered similarly. Both datasets were overlaid onto the graph of the footpath network, where each of the housing areas were highlighted in red and facilities highlighted in orange.

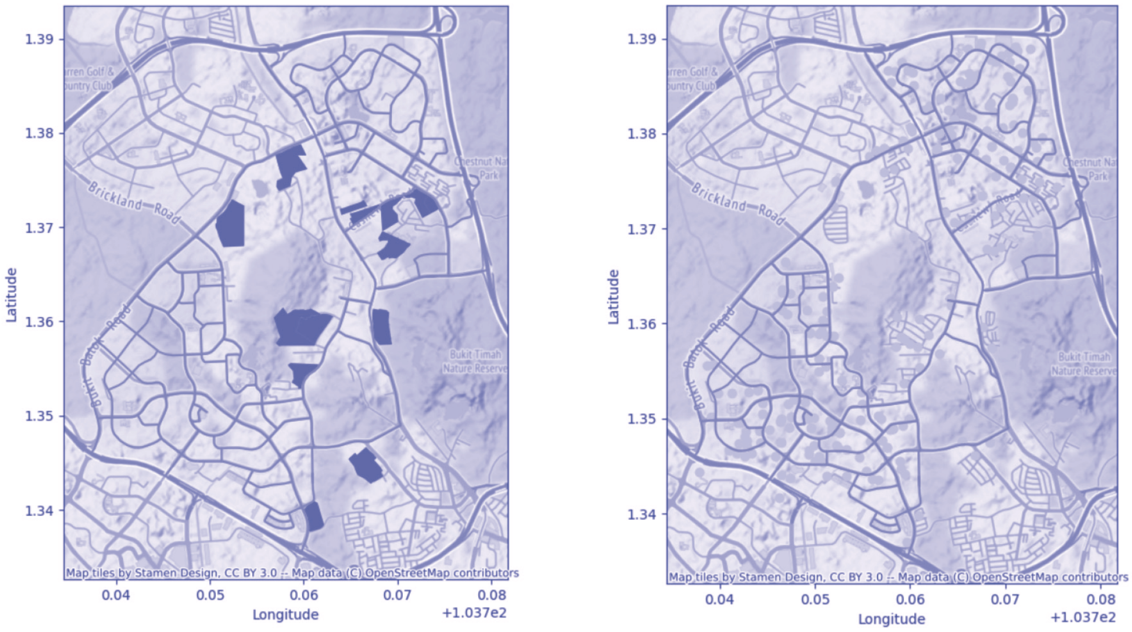


Figure 5. Visualization of housing area dataset (left) and facilities dataset (right). Made with Python.

Appendix B: Results for RQ1

Appendix B1: Pairwise Comparison Method

First, from the paired sample t-tests derived by Koh and Wong (2013), the four top environmental factors generally perceived as the most important by cyclists in Singapore were identified.

Infrastructural compatibility factors	Actual (mean point)	Shortest (mean point)	Paired *t*-tests		
			t statistics	$P(T \leq t)$ 2 tail	Bonferroni test
Detour, F2	4.08	6.20	*−5.46*	*<0.001*	*S*
Road crossing delays, F3	6.72	6.72	–	–	–
Directional signs, F4	0.29	0.36	−0.32	0.75	NS
Comfort, F5	6.91	6.28	*4.68*	*<0.001*	*S*
Weather protection, F6	1.36	1.77	−1.21	0.24	NS
Stairs/slope, F7	8.76	8.50	1.93	0.07	NS
Accident risk, F8	8.05	8.52	*−2.34*	*0.03*	NS
Crowdedness, F9	8.74	9.64	*−2.74*	*0.01*	NS
Shops along route, F10	0.48	0.33	1.24	0.23	NS
Good scenery, F11	2.15	1.66	1.75	0.09	NS

Italic represents significance at 95% confidence level.

Table 1. Paired sample t-tests for comparison of infrastructural compatibility factors (Koh and Wong, 2013)

Excluding comfort, which is difficult to measure as it is subjective to the cyclists' own tastes, and detour, which will be analysed with graph theory later on in the project, the three most important factors are slope / stairs, Accident Risk and Accessibility to Shops along the route. Considering the need to promote first and last mile transport in the West District, proximity to MRT stations was included as the final environmental factor.

1 - Equal

3 - Slightly favours

5 - Strongly favours

7 - Very strongly favours

9 - Extremely favours

	9	7	5	3	1	3	5	7	9	
F1									✓	F2
F1								✓		F3
F1				✓						F4
F2				✓						F3
F2	✓									F4
F3		✓								F4

F1 - Proximity to Shops

F2 - Proximity to Major Roads

F3 - Proximity to MRT / LRT stations

F4 - Gentle Slopes

Table 2. Scoring the factors by comparing relative importance

	F1	F2	F3	F4
F1	1	1/9	1/7	3
F2	9	1	3	9
F3	7	1/3	1	7
F4	1/3	1/9	1/7	1
Sum	17.3333	1.5556	4.2857	20.0000

Table 3. Relative Importance Matrix

	F1	F2	F3	F4	Weight
F1	0.0577	0.0714	0.0333	0.1500	0.0781
F2	0.5192	0.6429	0.7000	0.4500	0.5780
F3	0.4038	0.2143	0.2333	0.3500	0.3003
F4	0.0192	0.0714	0.0333	0.0500	0.0435

Table 4. Normalization and Weight Determination

Appendix B2: Fuzzy Membership Functions

Before plotting the fuzzy membership functions, suitability values and their respective criteria were generated for each of the four environmental factors. For the percentage slope factor, terrain with a percentage slope value of between 2% and 4% is comfortable for cyclists to travel on (American Association of State Highway and Transportation Officials, 2018). For proximity to major roads, research was conducted on guidelines for cycling paths, and it was found that a roadside verge ranging from 3m to 5m is necessary to be left between a development and a road, for tree planting to sustain a pervasive sense of greenery and the safety of cyclers (Urban Redevelopment Authority, 2018). As for the accessibility to MRT Stations and Shops factors, guidelines for the maximum walking distance to bus and LRT services were stated to be 400m, which approximates to a 5 minutes walk (Land Transport Authority, 2019). As such a distance of more than 400m was awarded 0 points for suitability, and a distance of less than 80m, a 1 minute walk, was considered to be the optimum for the placement of the cycling route.

<table>
<tr><th>Factor</th><th>Minimum (0)</th><th>Intermediate (between 0 and 1)</th><th>Optimum (1)</th></tr>
<tr><td>Amount of slope (%)</td><td>> 4</td><td>4 - 2</td><td>< 2</td></tr>
<tr><td>Proximity to major roads (m)</td><td>< 3</td><td>3 - 5</td><td>> 5</td></tr>
<tr><td>Accessibility to MRT stations (m)</td><td rowspan="2">> 400</td><td rowspan="2">400 - 80</td><td rowspan="2">< 80</td></tr>
<tr><td>Accessibility to shops (m)</td></tr>
</table>

Table 5. Criteria for allocation of suitability values

After determining the criteria for allocation of suitability values, fuzzy membership functions were plotted using RStudio's "FuzzyNumbers" package, where α represents the degree of membership, in other words the suitability value, and x represents the input value, in other words the quantitative measure of the environmental factors.

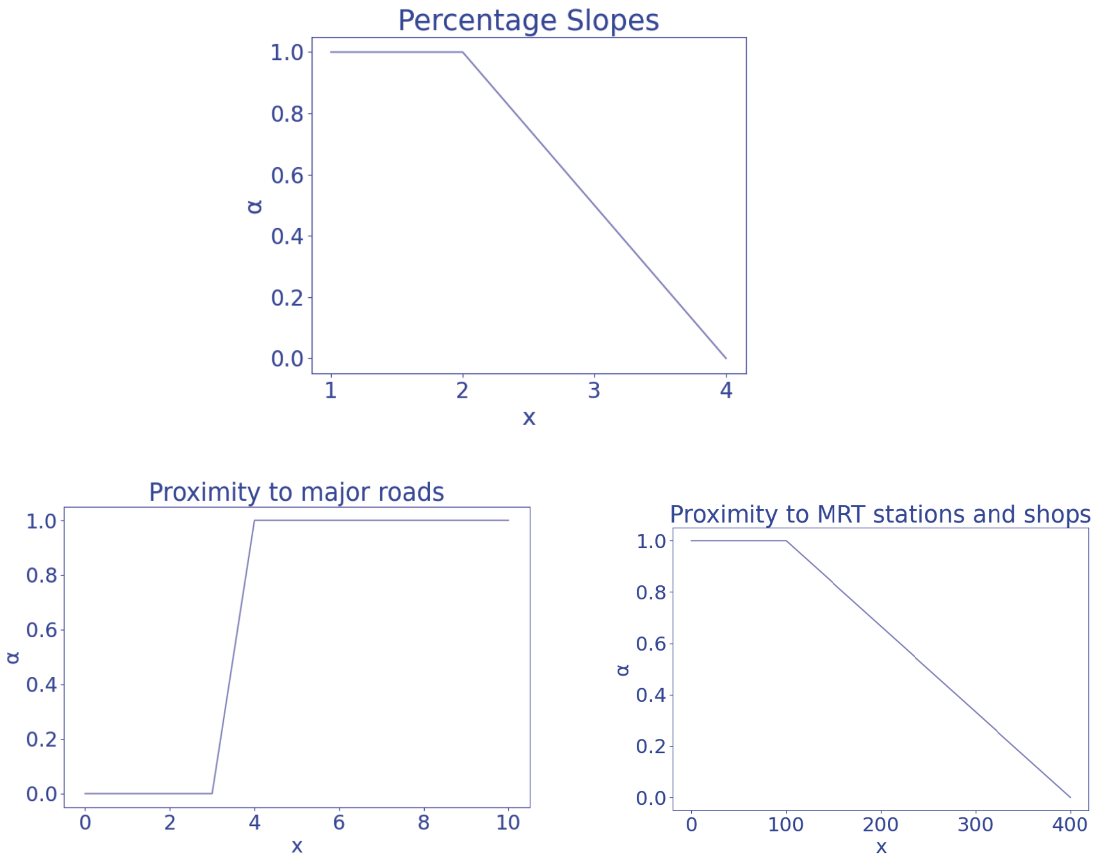


Figure 6. Fuzzy Membership Functions for the four environmental factors. Made with R.

Appendix B3: Suitability Maps for Factors

For the following descriptions, some terms are defined as follows. Footpath refers to the entire footpath dataset as a whole, a footpath segment refers to the separate linestrings making up the dataset, and points on the segment refer to the data points that make up each linestring.

Data for locations of shops and locations of MRTs and LRTs, were imported. The shortest distance between each point of the footpath and any one of the shops was calculated, before the mean value was taken for all the points on each segment to represent the mean proximity of the footpath segment to a shop. The mean values were passed through the fuzzy membership function for shops and the respective suitability values were calculated. This was also done for proximity to shops. With the suitability values for each segment of the footpath on a range of values from zero to one, suitability maps were plotted for each of the two factors.

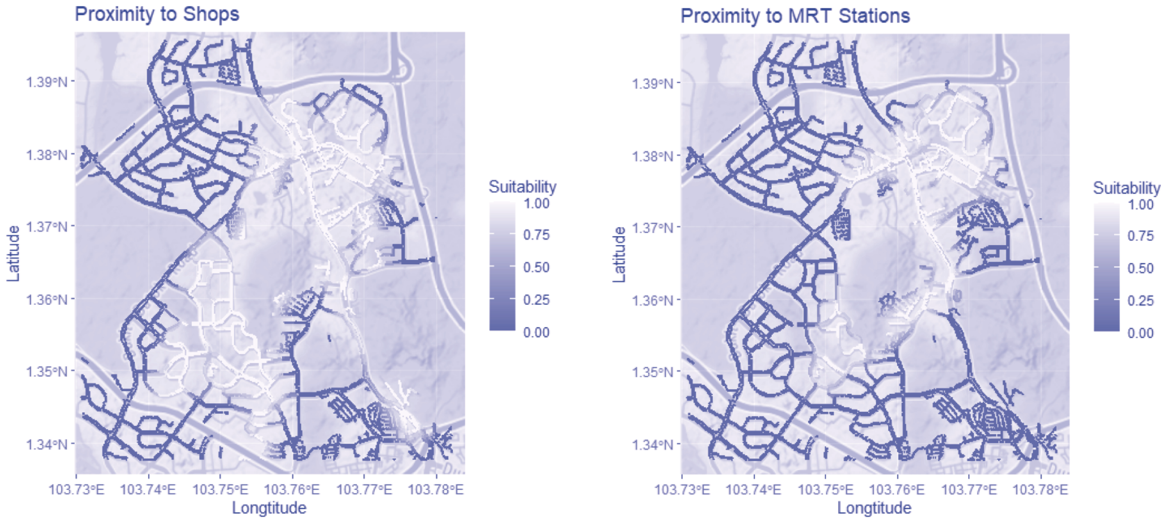


Figure 7. Suitability maps for Shops and MRT Stations. Made with R.

For the other two factors, Proximity to Major Roads and Percentage Slope, Python was used instead to calculate the suitability values for each segment of the footpath, as it was deemed to be better at handling data calculation of large datasets than R. The shortest distance between the major roads and the footpath was taken, but instead of taking the mean value of all the points on each footpath segment, the point with the shortest proximity to a major road was considered and taken as the proximity of the footpath segment, because a major road is considered a constraint and if one point on the footpath segment does not fulfill the proximity criteria, the entire segment should be affected. The gradient of the slope was converted into percentage slope, and the slope value for each footpath segment is taken to be the mean slope values of the points on the segment. These values were passed through the respective fuzzy membership functions for Major Roads and Slope, to generate the suitability values needed to plot each suitability map.

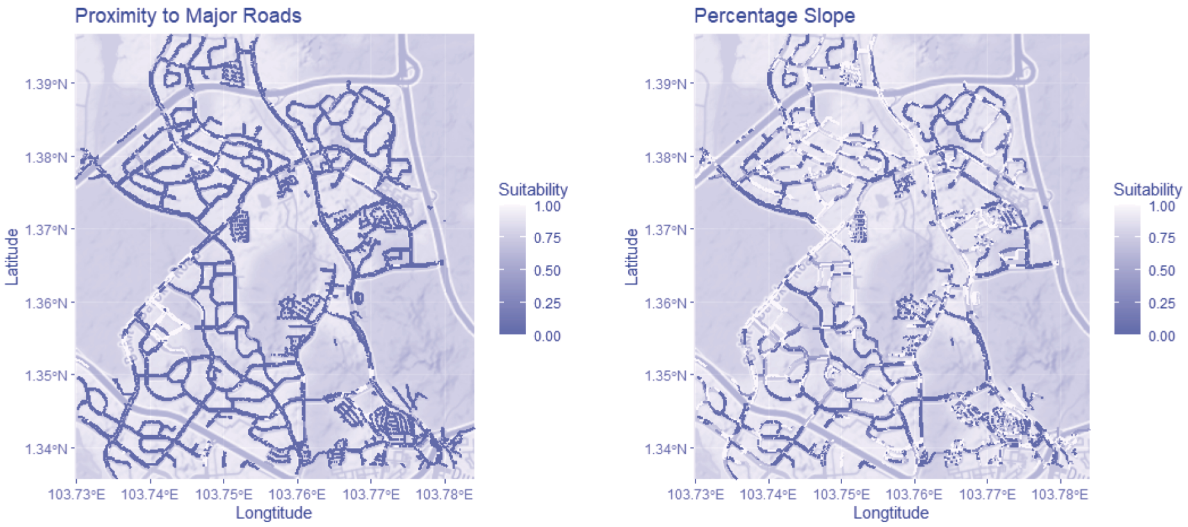


Figure 8. Suitability maps for Major Roads and Percentage Slope. Made with R.

Appendix C: Results for RQ2

Appendix C1: Visualisation of Footpath Network

For graph theory to be applied to answer this question, Python was first used to plot the edges of the graph, which are the footpath segments, as well as the nodes of the graph, which are the intersection points of the footpaths.

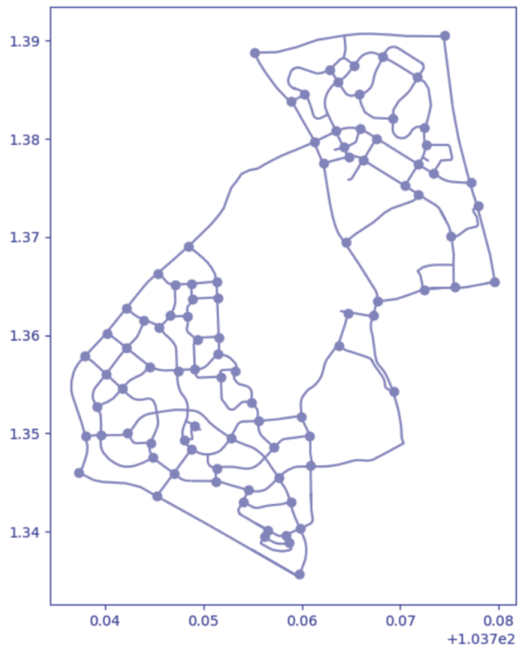


Figure 9. Map of footpath network. Made with Python.

The graph of the footpath network was overlaid with the combined suitability map produced by the Weighted Linear Combination method. Points were then given to each node on the graph based on the overall suitability value of the footpath segment the node is located at.

The datasets of housing areas and of facilities, namely Community Institutions, Healthcare facilities, and Sports and Recreational facilities were obtained and filtered to only include those in the areas of focus. Both datasets were overlaid onto the graph of the footpath network.

Appendix C2: Visualisation of Most Efficient Placement of Cycling Paths

For the housing dataset, a nearest node was chosen for each housing area where a score was given respective to the geometrical area of the housing. Similarly, for the facilities dataset, a nearest node was chosen for each of the facilities.

Then, a shortest path was generated from each of M housing areas to each of N facilities using Dijkstra's algorithm in Python' networkx module. As a result, $M \times N$ paths were generated. They were then layered together to obtain a frequency graph indicating the frequency of each edge being used as the shortest paths.

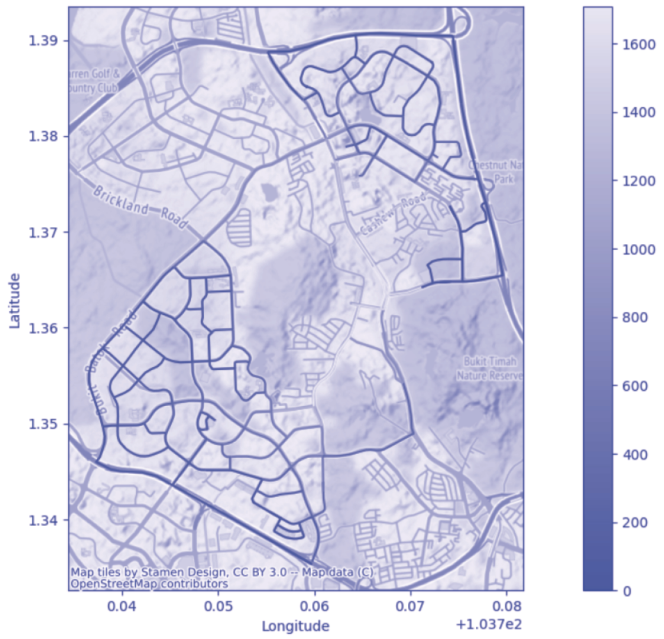


Figure 10. Map of frequency of each edge being used as shortest path. Made with Python.

The value of each node and edge was then multiplied by their corresponding values in the combined suitability map to form a new graph. Then, a map of the most efficient placement of cycling paths was obtained by filtering the graph with the median value of all the values in the graph, where those paths with values higher than the median will be selected.

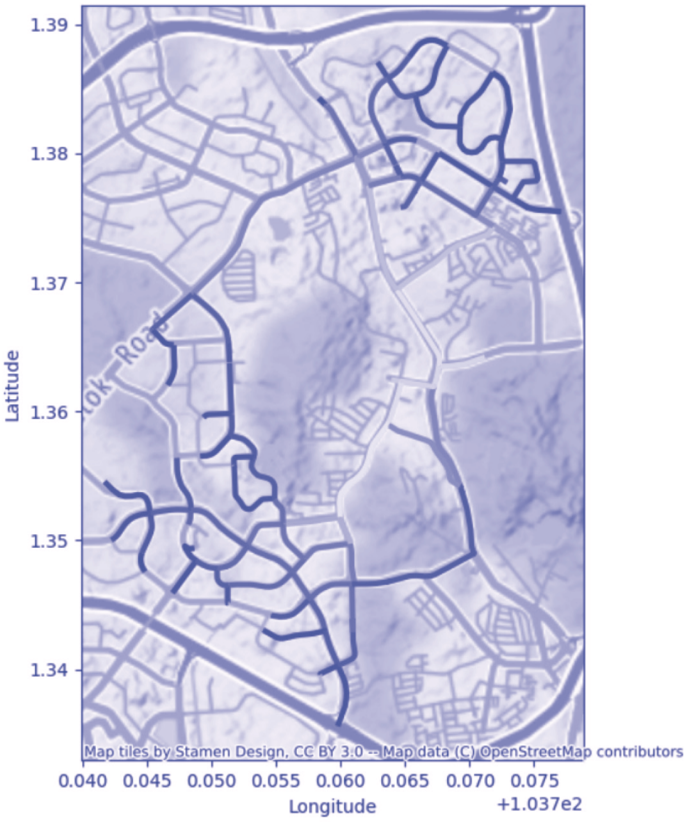


Figure 11. Map of the optimal cycling path. Made with Python.

Appendix D: Results for RQ3

Appendix D1: Visualisation of Yellow Boxes

The bicycle parking facilities we focused on were the yellow boxes within the planning area because they are increasingly being introduced to more locations islandwide to address indiscriminate parking caused by bike - sharing schemes (Urban Redevelopment Authority, 2018).

The dataset of bicycle parking facilities was first taken from Land Transport DataMall. The coordinates of the yellow boxes were then filtered out from the other bicycle parking facilities as a csv file, which was plotted and analysed.

Figure 12. Map of yellow boxes. Made with R.

Appendix D2: Suitability Maps for Proximity to Facilities

The suitability of placement of yellow boxes in the two planning areas was analysed in relation to surrounding facilities. The facilities taken into consideration were the same as those used in graph theory analysis for RQ2, and the shortest distance between each facility and any yellow box was calculated, before being represented as a suitability map based on shortest distance to a parking lot.

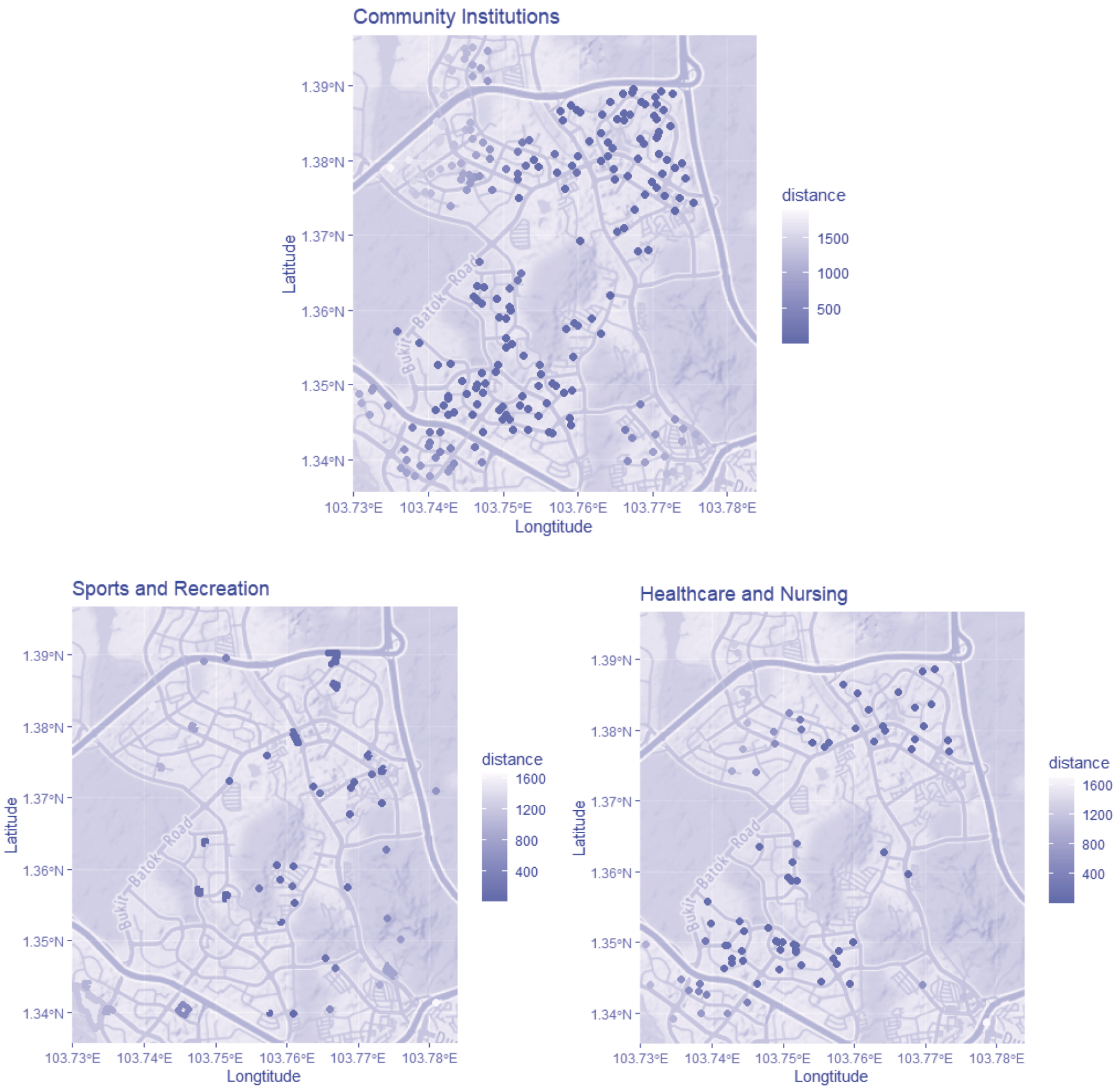


Figure 13. Suitability maps of shortest distance of facilities to yellow boxes. Made with R.

The shortest distance between each yellow and three nearest facilities is also calculated so as to analyse the ease of travelling from any yellow box to a facility.

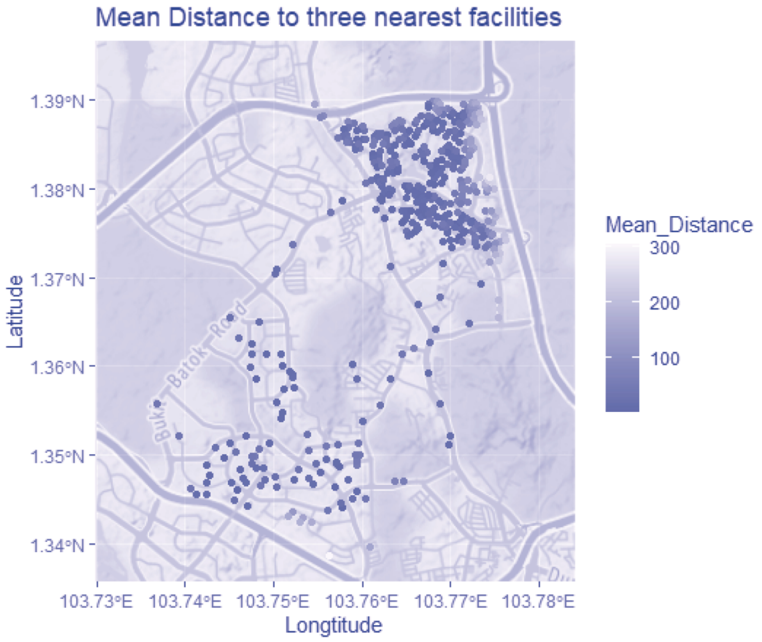


Figure 14. Map of mean distance of yellow boxes to three nearest facilities. Made with R.

From the analysis of accessibility of facilities to yellow boxes, almost all facilities in Bukit Batok and Bukit Panjang are within 500 metres of a yellow box, proving the positioning of yellow boxes convenient in proximity to the facilities in the area. From the analysis of mean distance of yellow boxes to the three nearest facilities, it can also be concluded that apart from the parking lots in the outskirts of the planning areas, almost all the yellow boxes are positioned close to a facility.

Appendix D3: Application of Graph Theory

To further analyse the connectivity of yellow boxes in Bukit Batok, graph theory was applied, by taking the yellow boxes as the nodes, and connecting each yellow box to the four nearest neighbours to form a weighted graph. Four nearest neighbours are chosen because these represent the potential destinations of cyclists travelling from a yellow box, and forming a network with such nodes would allow for accessibility of one yellow box from another to be measured. The actual distance of the shortest route between the yellow boxes was measured from geojson.io, and taken to be the weight of the edges in the graph.

Figure 15. Map of network of yellow boxes in Bukit Batok. Made with R.

Appendix E1: Dijkstra's Algorithm

Computer Program (using Python):

```
import json
import networkx as nx
import geopandas as gpd
import matplotlib.pyplot as plt

d2 = json.loads(open('graphnodes.geojson').read())
d3 = json.loads(open('graphedges.geojson').read())

gdf1 = gpd.read_file('combinedsuitability.geojson') #edges data
gdf2 = gpd.read_file('graphnodes.geojson') #nodes
gdf3 = gpd.read_file('graphedges.geojson') #edges

temp=[]
for x in range(len(gdf2)):
    temp.append(d2['features'][x]['properties']['edges'])

gdf2 = gdf2.assign(edges = temp)

temp=[]
for x in range(len(gdf3)):
    temp.append(d3['features'][x]['properties']['nodes'])

gdf3 = gdf3.assign(nodes = temp)
gdf1 = gdf1.assign(nodes = temp)

G = nx.MultiGraph()

for x in range(len(gdf1.nodes)):
    if len(gdf1.nodes[x]) == 2:
        G.add_edge(gdf1.nodes[x][0],gdf1.nodes[x][1], x, Combined = gdf1.Combined[x], Length =
gdf1.geometry[x].length)


gdfpoints = gpd.read_file('combineddataset.geojson')
```

```
gdfarea = gpd.read_file('housingdataset.geojson')
gdfbase = gpd.read_file('base.geojson')
geombase = gdfbase.geometry[0]

bounds = list(gdf2.total_bounds)
gdfpoints = gdfpoints.cx[bounds[0]:bounds[2], bounds[1]: bounds[3]]
gdfarea = gdfarea.cx[bounds[0]:bounds[2], bounds[1]: bounds[3]]
gdfpoints = gdfpoints.reset_index(drop=True)
gdfarea = gdfarea.reset_index(drop=True)
gdfpoints0 =gdfpoints
gdfarea0=gdfarea



l = []
for i, x in enumerate(gdfpoints.geometry):
    check = True
    if geombase.intersects(x):
        check = False
    if check:
        l.append(i)
gdfpoints = gdfpoints.drop(l)
gdfpoints = gdfpoints.reset_index(drop=True)



l = []
for i, x in enumerate(gdfarea.geometry):
    check = True
    if geombase.intersects(x):
        check = False
    if check:
        l.append(i)
gdfarea = gdfarea.drop(l)
gdfarea = gdfarea.reset_index(drop=True)



la = []
```

```
for i, x in enumerate(gdfarea.geometry):
    la.append(-1)
    for j, y in enumerate(gdf2.geometry):
        dist = x.distance(y)
        if la[i] == -1:
            la[i] = j
        else:
            prevdist = x.distance(gdf2.geometry[la[i]])
            if dist < prevdist:
                la[i] = j

lp = []
for i, x in enumerate(gdfpoints.geometry):
    lp.append(-1)
    for j, y in enumerate(gdf2.geometry):
        if j == 91:
            continue
        dist = x.distance(y)
        if lp[i] == -1:
            lp[i] = j
        else:
            prevdist = x.distance(gdf2.geometry[lp[i]])
            if dist < prevdist:
                lp[i] = j


templ = []
for x in la:
    for y in lp:
        if G.has_node(x) and G.has_node(y):
            if nx.has_path(G,x,y):
                templ += nx.shortest_path(G,x,y,weight='Length')
        else:
            print(x,y)

t=[]
```

```
for x in range(len(gdf2)):
    t.append(0)

for x in templ:
    t[x] += 1



n_num = []
for x in t:
    n_num.append(x)
n = len(n_num)
n_num.sort()

if n % 2 == 0:
    median1 = n_num[n//2]
    median2 = n_num[n//2 - 1]
    median = (median1 + median2)/2
else:
    median = n_num[n//2]



gdf2 = gdf2.assign(graph=t)

foo = []
for x in gdf3.nodes:
    cnt = 0
    count = 0
    for y in x:
        cnt = (cnt*count+t[y])/(count+1)
        count += 1
    foo.append(cnt)
gdf3 = gdf3.assign(graph = foo)

templ = gpd.GeoDataFrame()
for x in range(len(gdf3)):
```

```
    if gdf3.graph[x] > median:
        templ = templ.append(gdf3[x:x+1])
fig, ax = plt.subplots(sharex = True, sharey = True)

gdf4 = gdf3.append(gdf2)
gdf3.plot(ax=ax, column = 'graph')
templ.plot(column = 'graph')

plt.show()
nx.draw(G)
```

Appendix E2: Weighted Linear Combination

The suitability maps for the four factors obtained in RQ1 were combined into a complete map using the Weighted Linear Combination method in Python.

$$C = \sum_{i=1}^{n} F_i W_i$$

Where C refers to the final combined value, n refers to the number of factors considered, F_i refers to the suitability values of factors and W_i the weights of the factors respectively.

The Weighted Linear Combination was chosen to be the ideal method to combine suitability values for the four factors as a weighted average, because the factors were ensured to be mutually preference independent of each other as they address different areas of concerns of cyclists, and thus the suitability value of one of the factors would not affect that of another factor. The difference in t statistics from the paired t-tests conducted by Koh and Wong (2013) had also shown that the factors influence a cyclist's choice to use a cycling path to different degrees. Thus, this met the additivity assumption underlying the use of a Weighted Linear Combination as pointed out by Malczewski (2000), allowing the Weighted Linear Combination method to give valid results.

Appendix E3: Local and Weighted Clustering Coefficients

Appendix E3.1: Results

The local clustering coefficient measures local group cohesiveness as the fraction of connected neighbours of each vertex i, out of the total possible number of links that can exist between the neighbours of i.

$$c_i = \sum_{j,h} a_{ij} a_{jh} a_{ih} \frac{1}{k_i(k_i-1)}$$

(Watts and Strogatz, 1998)

Where i represents the vertex being assessed, j and h represent the neighbouring vertices, $a_{ij}a_{jh}a_{ih}$ represents the elements of the adjacency matrix, and k_i represents the vertex degree. $k_i(k_i-1)$ acts as a normalizing factor that normalizes the local clustering coefficient to a value between 0 and 1, for comparison with the weighted clustering coefficient. The average local clustering coefficient of the entire network is calculated by summing up the local clustering coefficient for all nodes in the network and dividing it by the total number of nodes.

$$\begin{aligned} C &= \frac{1}{N} \sum_i c_i \\ &= \frac{1}{96}(58.15714) \\ &= 0.606 \end{aligned}$$

Next, the weighted clustering coefficient, which is a measure of the local cohesiveness that takes into account the importance of the structure on the basis of edge weights, was calculated.

$$c_i^w = \frac{1}{s_i(k_i-1)} \sum_{j,h} \frac{(w_{ij}+w_{ih})}{2} a_{ij} a_{ih} a_{jh}$$

(Barrat, Barthelemy, Pastor-Satorras and Vespignani, 2004)

Where s_i represents the strength of the vertex, k_i represents the vertex degree, w_{ij} and w_{ih} represent the weights of the 2 adjacent edges of vertex i, and $a_{ij}a_{ih}a_{jh}$ represents the elements of the adjacency matrix. Similarly, the average weighted clustering coefficient was found by summing up the weighted clustering coefficient values for all nodes, and dividing it by the total number of nodes.

$$\begin{aligned} C^w &= \tfrac{1}{N} \sum_i^w c_i^w \\ &= \tfrac{1}{96}(56.66841) \\ &= 0.590 \end{aligned}$$

Upon comparing the two coefficient values, it has been found that the average weighted clustering coefficient is less than the average local clustering coefficient of the network of bicycle parking lots. According to Barrat, Barthelemy, Pastor-Satorras and Vespignani (2004), this result would indicate a network in which the topological clustering is generated by edges with lower weight, and in the case of bicycle parking lots, a network in which there is minimal distance between the bicycle parking lots.

Appendix E3.2: Computer Program (using R)

```
library("igraph")
library("tidyverse")
library("ncdf4")
library("ggplot2")
library("sf")

parking<-read.csv("Filtered bicycle parking 2.csv",TRUE,",")
parking2<-parking %>% filter(Longitude>103.7307 & Longitude<103.7822 & Latitude>1.3376 &
Latitude<1.3957)
parking<-as.data.frame(parking)
parking2<-st_as_sf(parking,coords=c("Longitude","Latitude"))
parking_viz<-as.data.frame(parking2)
parking_viz<-merge(parking_viz,links2,by.x=c("Index"),by.y=c("to"))

parking_viz <-parking_viz %>%
  mutate(long = unlist(map(parking_viz$geometry,1)),
         lat = unlist(map(parking_viz$geometry,2)))

links<-read.csv("Graph_Theory_Final.csv",stringsAsFactors=FALSE)

links2<-links[!duplicated(links[,c("to")]),]
nodes<-merge(parking_graph,links2,by.x=c("Index"),by.y=c("to"))
nodes<-nodes%>%subset(select=c(Index))

net<-graph_from_data_frame(d=links,vertices=new_nodes,directed=F)
simple_net<-simplify(net)
plot(net,edge.arrow.size=.4,vertex.label=NA)

trans<-transitivity(simple_net,type="weighted")
trans<-as.data.frame(trans)
sum_trans<-summarise(trans,trans=sum(trans))
average_trans<-summarise(trans,trans=mean(trans))

trans_lo<-transitivity(simple_net,type="local")
trans_lo<-as.data.frame(trans_lo)
```

```
sum_translo<-summarise(trans_lo,trans_lo=sum(trans_lo))
avg_translo<-summarise(trans_lo,trans_lo=mean(trans_lo))
```