OPEN ACCESS
Article

Clustering of green space qualities: evidence from three Australian cities

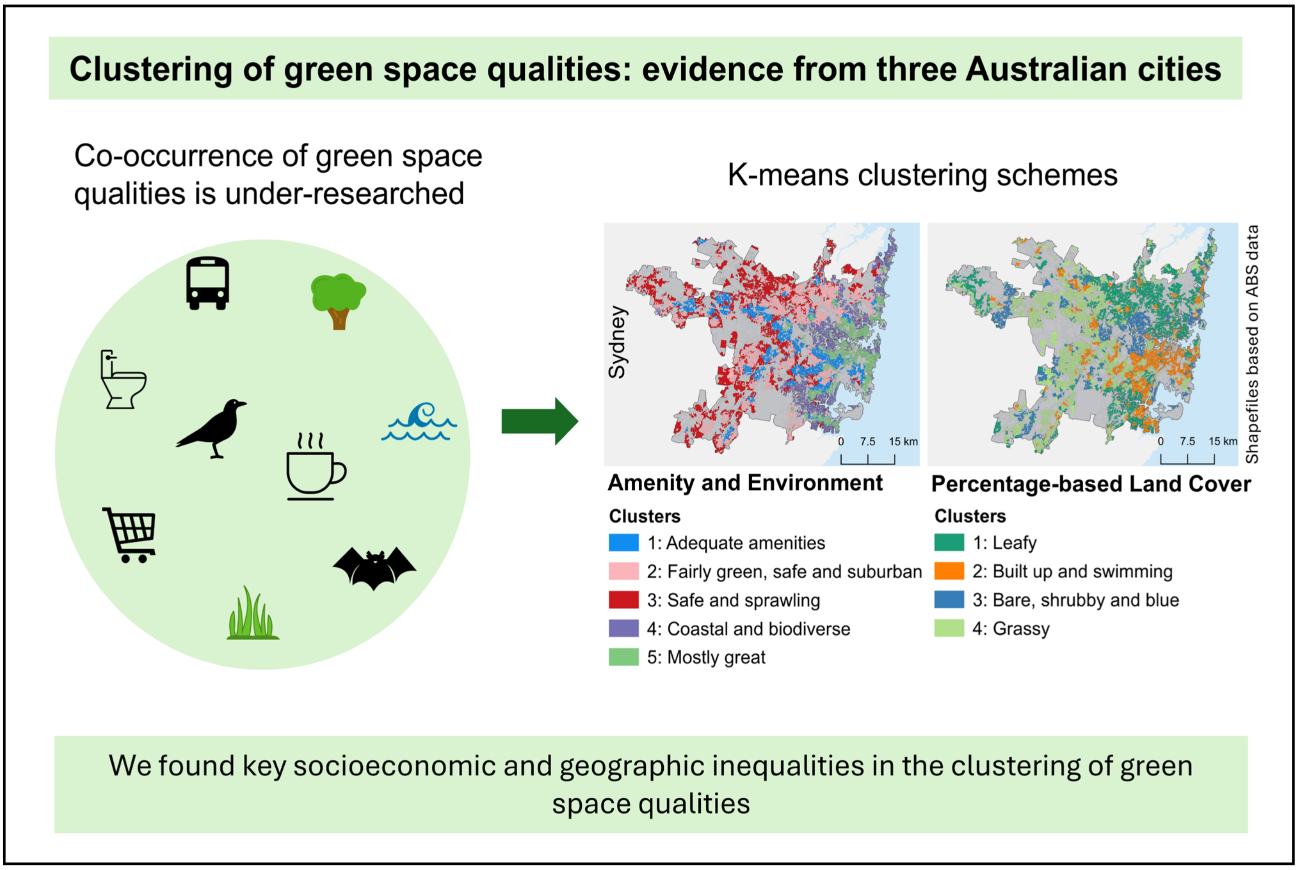
Lauren Del Rosario, Michael Navakatikyan, Xiaoqi Feng and Thomas Astell-Burt*

***E-mail:** thomas.astell-burt@sydney.edu.au

Received 30 March 2025; Revised 8 May 2026; Accepted 18 May 2026; Published xx July 2026

https://doi.org/10.55092/ijee20260011

Graphical Abstract



Highlights:

- Green space qualities were measured to examine cluster typologies.
- Leafy clusters were more likely to be found in areas of lower disadvantage.
- Grassy clusters were more likely to be found in areas of higher disadvantage.

Clustering of green space qualities: evidence from three Australian cities

Lauren Del Rosario[1,2], **Michael Navakatikyan**[1,2], **Xiaoqi Feng**[2,3,4] **and Thomas Astell-Burt**[1,2,5,6,*]

[1] Centre for Flourishing Cities, School of Architecture, Design and Planning, University of Sydney, Sydney, Australia
[2] Population Wellbeing and Environment Research Lab (PowerLab), Sydney, Australia
[3] School of Population Health, Faculty of Medicine and Health, University of New South Wales, Sydney, Australia
[4] The George Institute for Global Health, Sydney, Australia
[5] Westmead Applied Research Centre, Sydney Medical School, Faculty of Medicine and Health, The University of Sydney, Sydney, Australia
[6] Charles Perkins Centre, The University of Sydney, Sydney, Australia

***E-mail:** thomas.astell-burt@sydney.edu.au.

Abstract: Background: Health-relevant features of green space are often referred to under the catch-all term 'quality', but this masks how some qualities may cluster geographically and manifest contrasting patterns with socioeconomic circumstances. We measured multiple green space qualities to derive cluster typologies and their socioeconomic patterning. **Methods:** Novel green space qualities clustering was developed using K-means clustering for "Amenity and Environment" and "Percentage-based Land Cover" schemes. Amenity and Environment qualities spanned access/topography, amenities (positive and negative), biodiversity, safety, surrounding trees, total green space and beaches/coastline. A one-way parametric Analysis of Variance (ANOVA) was conducted to analyse associations between these clusters and Australian Bureau of Statistics Index of Relative Socio-economic Disadvantage (IRSD). A quality check exercise was conducted with team members examining screenshots of cluster results and answering a 4-point Likert scale. **Results:** Leafy clusters were more likely to be found in areas of lower disadvantage (when IRSD was split into quintiles and reversed so that a value close to 1 is low disadvantage and value close to 5 is high, mean = 1.80, SE = 0.011 at 1600 m). Grassy clusters were more likely to be found in areas of higher disadvantage (mean = 3.44, SE = 0.012). Many inland areas of the three cities were typically low in all domains except safety. The quality check exercise showed that the clustering matched lived experiences between 70.6%–88.2% depending on the scheme and scale. **Discussion:** Many green space quality domains are co-located. A typology of clusters was shown to relate to socioeconomic circumstances with potential implications for different health benefits from green space exposure.

Keywords: green space; cluster analysis; park quality; socioeconomic circumstance

1. Introduction

Green spaces have potential to offer a range of health and social benefits [1–4]. However, green spaces differ in their qualities such as those related to accessibility, amenities, biodiversity, land cover, and safety for example [5] which influences health outcomes [6]. Co-occurrence of qualities is likely, though this is under-researched. Co-occurrence of exposures may influence synergistic effects on health [7]. In other words, the combined effect of exposures may have a greater influence on health than the simple addition of the effect of separate exposures [8]. As such, research that does not account for co-occurring exposures may underestimate potential health effects. The aim of this study is to derive cluster typologies from multiple green space qualities and assess their socioeconomic patterning. We hypothesise that clusters with health-promoting qualities will be present in areas of lower socioeconomic disadvantage.

The status quo tends to assess green space quality by analysing individual qualities though there have been some exceptions. For example, a study using latent class analysis found that those with poor/fair memory were less common for green space types that had greater tree canopy (6.2%), compared to types that had similar amounts of tree canopy relative to open grass (8.5%–8.9%) [9]. Apart from studies that have used clustering/typologies for ecosystem services, few studies have employed clustering of green spaces qualities [10–12].

Cluster analysis which groups similar entities according to a number of variables [13] can help to analyse multiple green space dimensions [11]. This is important since various green space qualities do not operate independently of each other. For example, large amounts of tree cover and hilly terrain may both be beneficial for health. Clustering of green spaces can assist with their equitable provision. For example, Kimpton classified green spaces into "amenity rich", "sit or play", "transport" and "amenity poor" and found that wealthier households had more green spaces that were amenity rich [10]. Hepburn *et al.* have similarly found that deprived areas were less likely to have parks classified as "Amenity" and were also more likely to have parks classified as "Sports" [12]. Using a combined principal component analysis and spatial clustering method, Wu *et al.* found the need for heavily populated areas to have more community parks and better visibility of greenery from eye-level [11].

There were approximately 100 qualities in our previous study examining inequalities in green space qualities [5]. These qualities were divided into the following domains: accessibility, amenities and activities, beaches and coastline, biodiversity, safety, land cover and land use. The formation of these domains was mostly guided by the RECITAL tool [14] which was underpinned by a systematic review by Knobel *et al.* [15]. In practical terms, we would like to analyse these qualities with various health outcomes. However, approximately 100 qualities would result in an unprecedented number of models. We first reduced our qualities and created domains of similar qualities: access/topography, amenities (positive and negative), biodiversity, safety, surrounding trees, total green space, beaches/coastline and land cover. To our knowledge, this is the first green space clustering research that includes qualities across each of these domains. The co-occurrence of domains through cluster analysis was then examined, followed by analysis with socioeconomic data (Australian Bureau of Statistics (ABS), Socio-Economic Indexes for Areas (SEIFA), Index of Relative Socio-Economic Disadvantage (IRSD)).

Examples of health benefits that are theoretically linked to domains include: physical activity for access/topography [16,17]; social interaction for positive amenities [17]; physical activity [17] and immune function [18] for biodiversity; physical activity [19], improved quality of life and "reduced psychological distress" [6] for safety; cardiometabolic benefits for surrounding trees and trees within the land cover domain [20]; reduced all-cause mortality for total green space [21]; and general health for beaches/coastline [22].

We tested clusters at three different scales: 400 m; 800 m; and 1600 m. 400 m equates to approximately 5 minutes walking, 800 m (10 minutes) and 1600 m (20 minutes). Due to scale effects and the modifiable areal unit problem [23], we do not expect there to be a relationship between cluster results at differing scales.

2. Study areas

The study areas for this research are the three ABS Urban Centre and Localities (UCLs) of Sydney, Newcastle and Wollongong in NSW, Australia [24] (Figure 1). Sydney is a large metropolitan city (~4.7 M), whereas Newcastle (~350 k) and Wollongong have smaller populations (~300 k) [25]. All three cities are located along the eastern seaboard of the state of New South Wales (NSW) in Australia, and as such, contain a variety of coastal and beachside communities as well as inland suburbs. These three cities were chosen to analyse trends in cities of different sizes as well as for increasing representativeness of the urban form of coastal Australia. As described in Del Rosario *et al.* [5], analysis is at the level of ABS mesh blocks, with mesh blocks containing approximately 30–60 dwellings [26]. 400 m, 800 m and 1600 m network buffers around within-polygon centroids of residential mesh blocks were intersected with green spaces. These buffers were chosen since people may choose to visit multiple green spaces that are close to their places of residence.

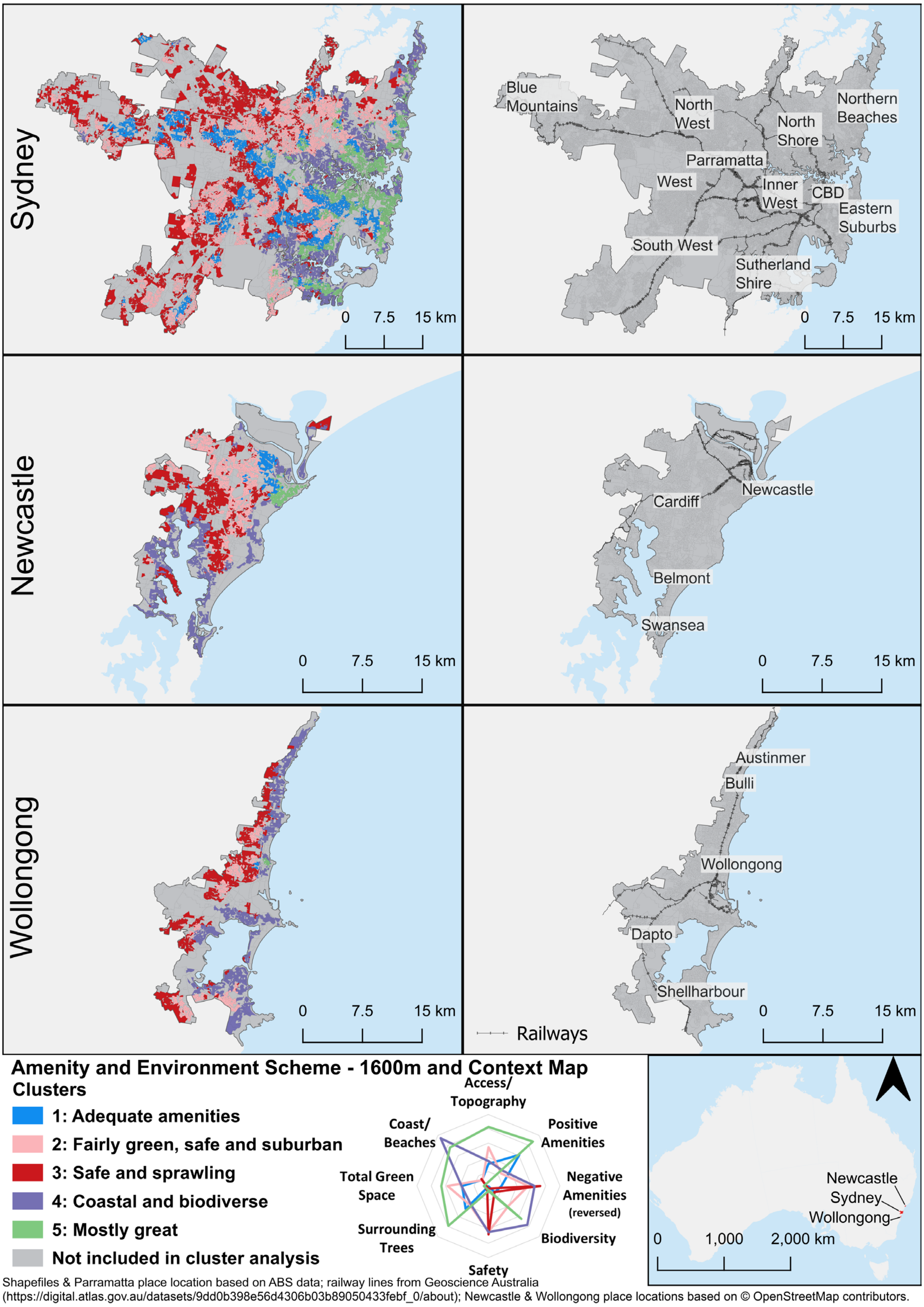


Figure 1. Amenity and Environment Scheme—1600 m and Context Map. Axis for radar graph is between −1.0–1.5. Scores closer to the outer edge of the radar graph are potentially more health promoting, while those closer to the centre are potentially less health promoting. This is why negative amenities (such as takeaway outlets and hotels/bars) was reversed so that a high score for negative amenities indicates relatively low occurrence of negative amenities. Mesh blocks that were not included in the cluster analysis did not have an intersect with a green space, were not residential, or did not have SEIFA data. See Figure S28 for breakdown of mesh blocks not included in cluster analysis.

3. Methods

3.1. Data

As described in Del Rosario *et al.* [5], data for green spaces was derived from NSW Department of Planning, Housing and Infrastructure [27] for Sydney and mesh blocks that were categorised by the ABS as "Parkland" for Newcastle and Wollongong [24].

Data inputs for the accessibility and topography domain included pathways [28], bikeways [29], 5 m Digital Elevation Model [30], public transport stops [31] and green space perimeter [27]. Data inputs for amenities included public toilets [32], eateries and supermarkets/greengrocers [33], and parkrun events [34]. Beaches were sourced from Geofabrik GmbH and OpenStreetMap Contributors [35] and coastlines were derived from ABS States and Territories shapefile data [24]. Biodiversity data was from "Biodiversity Values Map" [36] and Species of National Environmental Significance Distributions [37]. When using this latter dataset, the likely distributions of threatened species or their habitat was used as a proxy for biodiversity in the absence of open-access government records for all species. Calculation of diversity qualities such as the Shannon Index generally rely on individual data. However, we only had data for likely number of species/species habitat, and so we used number of species instead of data for individuals for the categories of birds, fish, flora, frogs, mammals, reptiles and other animals [5]. Data for safety was based upon NSW Bureau of Crime Statistics and Research [38]. As data were for crimes, the results were inverted to align with the other domains, in that larger scores have larger positive effects. It should be noted that the inversion of crime density may not necessarily translate to perceived safety. However, this depends on a range of factors at the individual (e.g. personality traits) and contextual levels (e.g. social networks). The crimes data was based on density and not occurrences, meaning an area with relatively low crime occurrences may have hot spots of crime if the crimes are clustered in a particular location [39] and instances of crime may not necessarily have occurred within green spaces though there was an intersection with a green space [5]. Land cover and surrounding trees data was from GeoVision [40,41]. Intersect area for total green space was based on data from NSW Department of Planning, Housing and Infrastructure [27] and ABS [24].

Green space qualities were aggregated to mesh blocks based on green spaces that intersect with each mesh block network buffer (*i.e.* 400 m, 800 m, 1600 m). For a detailed description of data used and aggregation methods, see Del Rosario *et al.* [5].

Socioeconomic data used was from ABS SEIFA IRSD [42]. The IRSD was chosen for its focus on disadvantage and includes data related to income, employment, education, rent, disabilities, single parent families, separated or divorced, lack of cars and poor English [43].

3.2. Clustering schemes, domains and green space qualities selection

Clustering schemes using Amenity and Environment qualities and Percentage-based Land Cover qualities were developed. Amenity and Environment clustering included the following domains: (1) access/topography; (2) positive amenities e.g. cafes, restaurants, supermarkets, public toilets and parkrun; (3) negative amenities e.g. hotels/bars and takeaway outlets; (4) biodiversity; (5) safety; (6) surrounding trees; (7) total green space; and 8) coast/beaches (Table 1). Percentage-based Land Cover qualities are included in Table 2.

The domains were amended from the seven domains in Del Rosario *et al.* [5]. The access/topography domain was renamed; amenities and activities was split into positive amenities and negative amenities; incivilities was renamed to safety; and the total green space domain was created. Land uses were changed to only include certain surrounding tree cover qualities of green spaces. As the surrounding land use categories derived from the ABS were quite different to each other (e.g. Education, Hospital/Medical, Residential *etc.*), their calculation together would not be practically meaningful as a cluster.

Rather than conduct a cluster analysis on 104 measured qualities, some of which had substantial conceptual and empirical overlap, we opted to reduce this large number to a more practicable amount of 38 (Supplementary Table S4–Supplementary Table S10). This involved excluding qualities firstly based on inter-correlation (close to a 0.80 upper threshold of inter-correlation between qualities as measured by the absolute value of the Spearman's rank correlation coefficient (rho) (Supplementary Table S4–Table S10). Qualities were further excluded for other reasons (e.g. conceptual similarity with other variables, conceptual irrelevance to green space and health) (Supplementary Table S4–Table S10). A few inter-correlated qualities were used in the calculation of land covers separately from the qualities they were correlated with. Due to conceptual similarity, we limited the land cover qualities to those relating to percentage, with the land cover qualities relating to area contained in the Supplementary Material e. g. Supplementary Figures S9–S10.

For qualities that comprise each of the domains and their measurement, see Supplementary Material S1.

Table 1. Amenity and environment qualities.

Domain	Green space quality
Access/ topography	Length of pathways
	Length of bikeways
	Number of Railway Stations within 100 m
	Number of Bus Stops within 100 m
	Number of ferry wharves within 100 m
	Area of green space > 6°
	Ratio of Perimeter (P) to perimeter standardised by area (A)
	Percentage of green space > 6°
Positive amenities	Number of public toilets within 100 m
	Number of cafes within 100 m
	Number of restaurants within 100 m
	Number of supermarkets/greengrocers within 100 m
	Presence of parkrun event
Negative amenities	Number of hotels/bars within 100 m
	Number of takeaway shops within 100 m
Biodiversity	Number of threatened bird species/species habitat likely to occur
	Number of threatened flora species/species habitat likely to occur
	Number of threatened mammal species/species habitat likely to occur
	Area of land with biodiversity value
	Shannon index
Safety	Density of alcohol-related assaults
	Density of malicious damage
	Density of non-domestic violence assaults
	Density of robberies
	Density of steal from persons
Surrounding Trees	Area of tree canopy surrounding green space within 100 m
	Percentage of tree canopy to roads surrounding green space within 100 m
Total green space	Intersect of network buffer with green space area
Beaches/Coastline	Presence of coastline
	Area of beach within 100 m

Table 2. Percentage-based land cover qualities.

Domain	Green space quality
Land cover	Percentage of bare earth
	Percentage of open grass
	Percentage of shrubs ("other vegetative material not included within the Grass or Tree class" e.g. shrub, scrub, agriculture, aquatic plants [44]
	Percentage of tree canopy
	Percentage of built up area
	Percentage of roads/paths
	Percentage of outdoor swimming pools
	Percentage of water in a park that is not a swimming pool e.g. "streams, canals, ponds, lakes, reservoirs, estuaries, and bays" [44]

3.3. Scoring domains

Typically, for domains with more than one quality the qualities were ranked into quintiles from 0 to 4. The qualities in the "Safety" domains were ranked into quartiles, because the underlying data had only 4 categories: zero density, low density, medium density, and high density [5]. Qualities with a binary scoring or low number of categories still underwent the quintile-generating SAS Proc Rank procedure but results were not placed into all five categories from 0–4. The latter was done to equalise the relative weight of all components of a domain.

To produce aggregated scores for domains, we considered whether the green space qualities would generally have a positive or negative influence on health (see Supplementary Tables S1–S2 for qualities scored negatively). The values of the quintiles for the positive scores were summed for the respective domain. The values of the quintiles/quartiles for the negative scores were first inverted, so that the highest quintile would have a positive effect on health, and then summed for the respective domain.

3.4. Data preparation

We included data for mesh blocks if the respective network distance buffers (1600 m, 800 m and 400 m) intersected with green spaces. These mesh blocks also needed to be categorised as residential and have SEIFA data available. If there were data quality issues for any of the green space qualities (Table 3), these values were replaced by mean values calculated from a Box-Cox transformation. Data was prepared based on Office for National Statistics [45], Lu [46] and Bhalla [47].

Table 3. Comparative characteristics of the network buffers.

Size	N MBs		Per MB		Ratio of GS Area *vs.* 400 m buffer	Missing due to data quality issue*	
	With intersect	Without intersect	N intersects	Mean area sq m		Amenity and Environment Scheme Surrounding trees	Percentage-based Land Cover Scheme Each variable
1600 m	47079	459	36.2	211,261	40	2459 (5.2%)	1617 (3.4%)
800 m	46245	1293	9.8	31,118	6	927 (2.0%)	445 (1.0%)
400 m	39570	7968	3.2	5285	1	383 (1.0%)	162 (0.4%)

* All missing values were imputed during Box-Cox transformation, *i.e.*, replaced by the mean value. GS = green space; MB = mesh block. Intersect is the intersection of a green space with mesh block network buffer (note differing scales: 400 m; 800 m; 1600 m). Sample includes mesh blocks that are residential, with SEIFA indices, UCL in Sydney, Newcastle and Wollongong. Total sample = 47,538 mesh blocks at each scale.

We accounted for extreme values where observations exceeded the 99th percentile.

After scores were produced, extreme values were replaced by the value for the 99th percentile.

To reduce skewness, we tested two types of transformation procedures (Box-Cox, inverse hyperbolic sine transformation). The Box-Cox appeared to be the best and was applied uniformly, using SAS Transreg to optimise the parameter lambda. If the distribution of a quality contained values less than 3, the data were shifted toward positive numbers by 3 minus the value of the minimum. The inverse hyperbolic sine transformation produced unsatisfactory results, mostly by decreasing skewness below −1 (in approximately one third of variables), and sometimes increasing above 1.

To make the magnitudes of variables used for clustering comparable, the data were standardised so that the mean = 0 and the standard deviation = 1.

3.5. Clustering

To produce clusters, SAS FASTCLUS for K-means clustering was applied. The K-means clustering is sensitive to the initial seed (*i.e.* initial observation) determined by the algorithm. These in turn depend on the position of the observations in the dataset. To find the best solution, *i.e.* the solution that minimises the within cluster standard deviation (Within STD) pooled over clusters and domains; we generated 500 randomly ordered samples of the same initial dataset. For each K size from 2–10, the best seed and solution was found. Four curves describing absolute and relative change in within cluster variance, the cubic clustering criterion and Pseudo-F statistics were analysed. We relied mostly on the Elbow method applied to the change in within cluster variance which becomes increasing smaller with the increase of K. The first point where the change stabilises is considered the optimal K-value. See Supplementary Material S2 for statistics of the best performing clustering models and for detail on aligning clusters across different network buffers. See Supplementary Material S8 for elbow plots.

3.6. ANOVA

A one-way parametric Analysis of Variance (ANOVA) was conducted to analyse IRSD quintiles (reversed) by clusters of green space qualities. Where the mean is lower, this indicates an association with low disadvantaged areas and where the mean is higher, this indicates an association with more disadvantaged areas.

3.7. Quality check

To determine if the clusters matched lived experiences, we tested the Amenity and Environment and Percentage-based Land Cover clusters with a quality check exercise among five team members for both 800 m and 1600 m scales. Team members lived in the Sydney study area and Geographic Information Systems (GIS) training ranged from none to proficient. Blinding to the cluster classifications did not occur. A quality check exercise was not conducted for the 400 m scale since patterns at 400 m tended to be more dispersed than at 800 m and 1600 m.

Screenshots of cluster results and green spaces for ABS Statistical Area Level 2 (SA2s) [24] that team members were familiar with were provided along with a description of the clusters. They then answered questions such as "The cluster colour type(s) accurately match your experience of parks within Area 1's boundary" using a 4-point Likert scale, with responses ranging from Strongly Agree to Strongly Disagree. If they disagreed or strongly disagreed, they were asked to elaborate, "If you answered 'Disagree' or 'Strongly Disagree', please explain why". A "Description of clusters" that was provided to team members assisting in the quality check is found in Supplementary Material S4. Team members were not required to visit the area but were allowed to use mapping and aerial imagery tools such as Google Maps, Google Imagery, Esri basemaps and imagery, or OpenStreetMap basemaps.

4. Results

4.1. Network buffers

As expected, the larger the network buffer size, the larger the number of mesh blocks that intersected with green spaces (Table 3). There were also more intersects per mesh block and larger mean areas of intersects as buffer size increased. Importantly, at 1600 m, 99% of mesh blocks had an intersect with a green space, suggesting almost universal green space access within 20 minutes' walk. A small portion of data was imputed during the Box-Cox transformation applied to normalise the distribution of qualities. These were mostly at the 1600 m scale and influenced a single domain (Surrounding trees) for the Amenity and Environment Scheme (5.2% of the data), and all single-quality domains for the Percentage-based Land Cover Scheme (3.4% of the data).

4.2. Amenity and environment qualities

At the 1600 m scale for Amenity and Environment qualities, Cluster 1 ("Adequate amenities") scored moderately for positive amenities and surrounding trees (Figure 1). The cluster tended to be located in places along the rail corridor. Cluster 2 ("Fairly green, safe and suburban") scored moderate-highly for safety, and moderately for total green space, access/topography, surrounding trees, and biodiversity. In Sydney, the cluster tended to be located in parts of the Upper North Shore, North West, Blue Mountains, West, Inner West and South West (Figure 1). Cluster 3 ("Safe and sprawling") scored moderately-high for safety, while everything else was low. In Sydney, the cluster tended to be located in parts of the Upper North Shore, North West, Blue Mountains, West, and South West. In Wollongong it tended to be located to the west of the study area (Figure 1). Cluster 4 ("Coastal and biodiverse") scored highly for coast/beaches and moderately-high for biodiversity and safety. This cluster tended to be located along the coast in all three study areas, but was also present in some inland areas such as the Upper North Shore (Figure 1). Cluster 5 ("Mostly great") scored moderately-high/highly for access/topography, positive amenities, biodiversity, surrounding trees, total green space and coast/beaches. Cluster 5 tended to be found in eastern portions of the three study areas and was more geographically limited in Newcastle and Wollongong (Figure 1).

4.3. Percentage-based land cover qualities

At the 1600 m scale for Percentage-based Land Cover qualities, Cluster 1 ("Leafy") scored highly for trees and moderately for shrubs. In Sydney, this cluster was mainly located in the North Shore, North West, Northern Beaches, Blue Mountains and South (Figure 2). Cluster 2 ("Built up and swimming") scored highly for built up areas and swimming pools and moderately for grass. In Sydney, this cluster was predominantly found in the Eastern Suburbs, CBD, Inner West and South but was also found in other parts of Sydney (Figure 2). Cluster 3 ("Bare, shrubby and blue") scored highly for bare earth, shrubs and water. This cluster was present along the coast of the three cities, but in Sydney was also found inland, for example in the centre [In the underlying data used to classify land cover [40], the area in the centre of Sydney appeared to be largely misclassified as bare earth rather than trees or shrubs for example [5].] of Sydney and parts of the West (Figure 2). Cluster 4 ("Grassy") scored highly for grass. In Sydney, the cluster tended to be located in the Western suburbs (Figure 2).

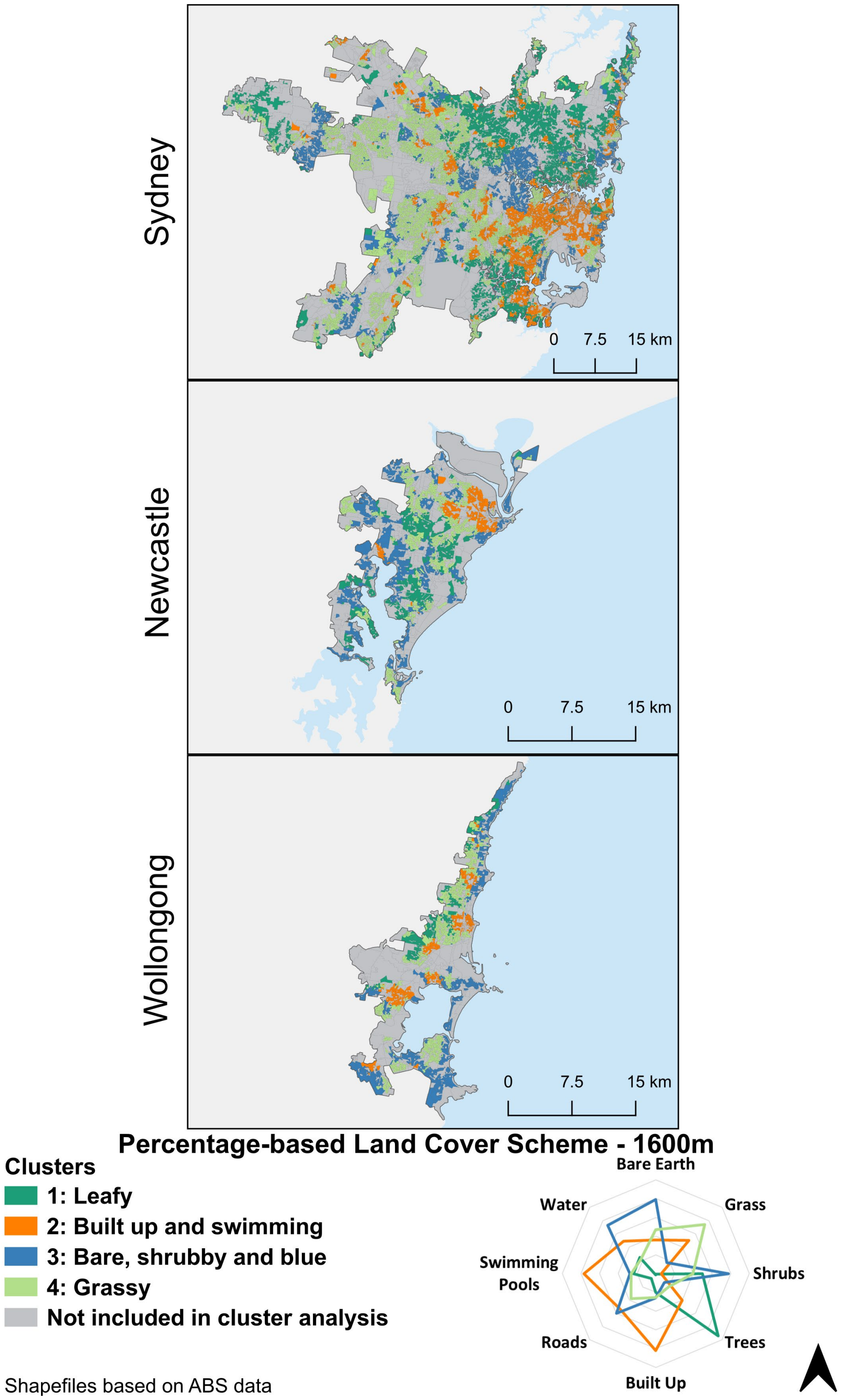


Figure 2. Percentage-based Land Cover Scheme—1600 m. Axis for radar graph is between −1.0–1.5. 'Water' means water in a park that is not a swimming pool e.g. "streams, canals, ponds, lakes, reservoirs, estuaries, and bays" [44]. Mesh blocks that were not included in the cluster analysis did not have an intersect with a green space, were not residential, or did not have SEIFA data. See Figure S28 for breakdown of mesh blocks not included in cluster analysis.

4.4. Associations with socioeconomics

For the Amenity and Environment Scheme at the 1600 m scale, the most disadvantaged areas were more likely to have the "Adequate amenities" cluster. At 800 m and 400 m, the most disadvantaged areas were more likely to have the "Safe and sprawling" cluster. At 1600 m and 800 m, when no green spaces within network buffers were excluded, the least disadvantaged areas where more likely to have the "Mostly great" cluster (Table 4).

For the Percentage-based Land Cover Scheme, when percentages of land cover were considered, the most disadvantaged areas were more likely to have the "Grassy" cluster at each of the three scales. When no green spaces within network buffer were excluded, the least disadvantaged areas where more likely to have the "Leafy" cluster (Table 4).

Table 4. ANOVA analysis of IRSD quintiles (reversed) by clusters of green space qualities.

Clustering scheme/ Cluster name	Cluster Number	Buffer = 1600 m Mean	SE	n	Buffer = 800 m Mean	SE	n	Buffer = 400 m Mean	SE	n
AMENITY AND ENVIRONMENT SCHEME (K = 5)										
No GS within network buffer	0	1.67	0.038	459	2.03	0.030	1293	2.71	0.016	7968
Adequate amenities	1	3.81	0.013	9483	3.13	0.014	11,026	2.95	0.016	8057
Fairly green, safe and suburban	2	2.86	0.014	9556	2.76	0.013	11,477	2.55	0.013	11,918
Safe and sprawling	3	2.93	0.014	9997	3.19	0.012	13,397	3.28	0.012	13,545
Coastal and biodiverse	4	2.28	0.015	8015	1.99	0.017	5400	1.87	0.019	3764
Mostly great	5	2.00	0.011	10,028	1.94	0.016	4945	2.01	0.023	2286
PERCENTAGE-BASED LAND COVER SCHEME (K = 4)										
No GS within network buffer	0	1.67	0.038	459	2.03	0.030	1293	2.71	0.016	7968
Leafy	1	1.80	0.011	10,066	2.03	0.011	12,091	2.15	0.011	12,784
Built up and swimming	2	2.76	0.012	12,961	2.77	0.021	4667	2.68	0.028	2574
Bare, shrubby and blue	3	2.86	0.014	9358	3.07	0.014	10,560	2.61	0.018	5577
Grassy	4	3.44	0.012	14,694	3.14	0.011	18,927	3.29	0.010	18,635

ANOVA = One-way parametric ANOVA. All overall tests are significant with P < 0.001; IRSD = Index of the relative socio-economic disadvantage (reversed, the higher the index, the more disadvantaged). Pair-wise comparison also significant, almost all. SE = standard error of mean. K = number of clusters in scheme excluding "No GS within network buffer" cluster. GS = green space.

There were three non-significant differences between means of IRSD quintiles in the Amenity and Environment Scheme: between clusters 0 and 4, 0 and 5 and 4 and 5 (800 m); and four non-significant differences in the Percentage-based Land Cover Scheme between clusters 0 and 1 (800 m, 1600 m) and 0 and 2, 2 and 3 (400 m) (Table 5).

Table 5. Non-significant differences: pairs of clusters.

Scheme	N non-significant differences	Buffer 1600 m	Buffer 800 m	Buffer 400 m
AMENITY AND ENVIRONMENT SCHEME (K = 5)	3		0 & 4	
			0 & 5	
			4 & 5	
PERCENTAGE-BASED LAND COVER SCHEME (K = 4)	4	0 & 1	0 & 1	0 & 2
				2 & 3

Comparisons using Tukey's Test, two-tailed difference of cluster means with P < 0.05. See Table 4 to match cluster numbers with Cluster names.

4.5. Quality check

Five team members completed the quality check exercise for either three or four SA2s that they were familiar with, resulting in 17 SA2s assessed in total (n = 17 for each scheme and at each scale). Overall, cluster analyses at the 800 m and 1600 m scales tended to match team members' experiences of green spaces. For both schemes and at both scales, instances where team members selected "Strongly Agree" or "Agree" that the cluster analysis accurately matched their experience of green spaces outweighed the instances when they selected "Disagree" (Figure 3). Of 17 areas tested, "Strongly Agree" or "Agree" was chosen 70.6% of the time for the Percentage-based Land Cover scheme at 1600 m and 88.2% for the Amenity and Environment Scheme at 800 m. There were no instances where "Strongly Disagree" was chosen.

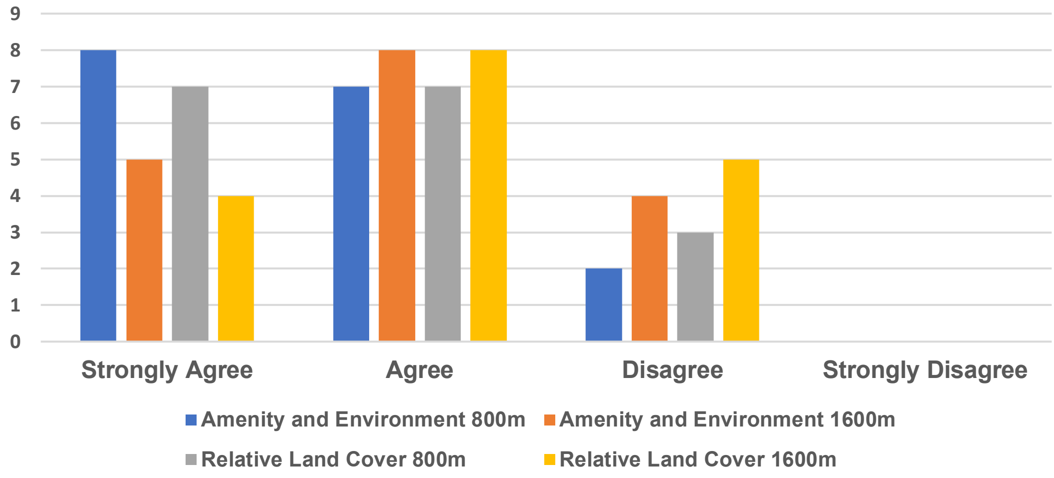


Figure 3. Quality check results for Amenity and Environment Scheme and Percentage-based Land Cover Scheme (800 m, 1600 m). For each scheme and at each scale, n = 17.

An inter-rater reliability metric was unable to be calculated since team members did not assess the same areas. However, we checked that their answers were not random using test statistics (two-tailed one-sample Student test) by comparing each of four rating combinations (buffer size × scheme) by the average mark against an expected average of the random answering = 2.5 (if Strongly Disagree = 1, Disagree = 2, Agree = 3 and Strongly Agree = 4). Means of the four groups ranged from 2.94 to 3.35 (n = 17 in each group), standard errors ranged from 0.17 to 0.18, t-values ranged from 2.43 to 5.01, while p-values ranged from 0.0001 to 0.027 (Supplementary Table S13).

5. Discussion

We have presented clustering schemes that include qualities spanning access/ topography, amenities (positive and negative), biodiversity, safety, surrounding trees, total green space, beaches/coastline and land cover. This has allowed us to examine socioeconomic inequalities in varying domains of green spaces.

Although Wu *et al.* [11] argues in favour of using a spatial clustering method to preserve the geographical influence of similar areas, we wanted the clustering to be principally determined by the green space qualities themselves. Despite the fact that we used a non-spatial clustering method, our results nevertheless showed that there were distinct spatial patterns, especially at the 800 m and 1600 m scales. For example, the "Adequate amenities" cluster (Amenity and Environment Scheme) clearly followed parts of the rail corridor. This was similar to our initial attempt at clustering qualities, with the "Adequate amenities" cluster also following the rail corridor (Figure S9). Future work analysing green space qualities could employ spatial clustering methods and compare them with non-spatial clustering.

While the clusters may share some striking similarities across scales, it is important to note that they are not strictly comparable due to scale effects and the modifiable areal unit problem [23]. For example, for Cluster 2 ("Built up and swimming", Percentage-based Land Cover Scheme), the pattern is dispersed at 400 m, but there is a clearer spatial pattern at 1600 m, particularly in the south-eastern section of the Sydney study area.

Although we find a lack of green spaces within 1600 m network analysis buffers was more common in areas of lower disadvantage (Table 4), it is noted that at this scale, only 459 mesh blocks (< 1%) did not have an intersect with a green space. These mesh blocks were located on the outskirts of Sydney (e.g. Box Hill, Marsden Park) and Wollongong (Figure S28). While we should not draw much from this finding, it is worth noting that a review examining park proximity and socioeconomics found mixed results with studies showing both low socioeconomic groups (six studies) and mid-high socioeconomic groups resided closer to parks (five studies) [48].

The lower disadvantaged areas were more likely to have the "Mostly great" (Amenity and Environment Scheme) cluster which includes a high scoring for positive amenities at 1600 m (Figure 1). While Kimpton [10] defines amenities differently [We define positive amenities as public toilets, cafes, restaurants, and supermarkets within 100 m of a green space and the presence of a parkrun event, and negative amenities as hotels/bars and takeaway outlets within 100 m of a green space. Kimpton [10] defines amenities as "playgrounds, seating, services, and sporting amenities" for example.], they found that higher income households tend to have lots of green spaces that are "amenity rich". More disadvantaged areas were more likely to have the "Adequate amenities" or "Safe and sprawling" clusters (Amenity and Environment Scheme), with the latter scoring low for positive amenities at 1600 m (Figure 1).

Leafy clusters were more likely to be found in areas of lower disadvantage, whereas grassy clusters were more likely to be found in areas of higher disadvantage (Percentage-based Land Cover Scheme). This is in general agreement with previous studies that did not employ clustering [49,50]. These socioeconomic inequalities in tree and grass cover could have implications for health, with tree cover linked with cardiometabolic health benefits [20]. However, in addition to the unidimensional benefits of tree cover, leafy clusters also had moderate levels of shrubs, and low levels of grass cover, bare earth, built up, roads, water and swimming pools at 1600 m. While the low levels of water and swimming pools is unlikely to have a positive effect on health, it is likely that the levels of the other qualities would result in a net positive health effect for this cluster. This could be tested in future with burden of disease data.

Many inland parts of all three cities were characterised by Cluster 3 (“Safe and sprawling”, Amenity and Environment Scheme). While these clusters did not have the highest disadvantage at 1600 m, Cluster 3 was more likely to be in the highest disadvantaged areas at the 400 m and 800 m scales (Amenity and Environment Scheme, Table 4). These clusters had nearby green spaces that were typically low in everything except safety (considering negative amenities was reversed in the Amenity and Environment Scheme). Low biodiversity means that possible protective effects of plants and animals for health and wellbeing [51] may not be realised in these areas, though this warrants future testing. Potential health benefits from blue space [52] would also be largely unrealised away from the coast and beaches, despite potential for some benefits near rivers or streams. There is potential for less health benefits from steep slopes [53], as land in these clusters is also typically flatter.

Generally, we found that positive amenities (public toilets, cafes, restaurants, supermarkets/greengrocers, and parkrun) tended to be found with negative amenities (hotels/bars and takeaway outlets). For example, negative amenities scored the least (tended to have more negative amenities) [It is noted that we reversed the score for negative amenities (radar in Figure 1) so that a positive score indicates a theoretically better outcome for health.] for the “Mostly great” cluster (Amenity and Environment Scheme) which also scored the highest for positive amenities (Figure 1). While some studies have found that fast food outlets and alcohol outlets occur more frequently in low-income/deprived areas [54,55], we found that the “Mostly great” cluster (Amenity and Environment Scheme) tended to be located more in the least disadvantaged areas (800 m, 1600 m) (Table 4). It is noted that we have analysed these negative amenities within a 100 m buffer of green spaces, and perhaps those located in more disadvantaged areas are located far from green spaces due to urban sprawl, however this requires further study.

There were geographic inequalities present in terms of swimming and built up areas. In Sydney, at the 1600 m scale, the “Built up and swimming” cluster (Percentage-based Land Cover Scheme) was predominantly found in the CBD, Eastern Suburbs, Inner West and South. It is noted however, that built up areas refers to artificial structures not including roads, paths and buildings in the GeoVision surface cover data [44]. [Area of Buildings and Area of Built up areas were highly correlated (Rho = 0.83, Supplementary Table S9).] These may not be desirable in a green space setting.

Three of the five clusters in the Amenity and Environment Scheme had relatively high scores for safety. It is noted that the two clusters that had lower levels of safety were located in areas that may see higher visitation and hence potential for crime density, though this requires future testing. For example, train stations overlap to a large degree with the “Adequate Amenities” cluster and Sydney’s CBD and beaches in the eastern suburbs are found in the “Mostly great” cluster.

6. Strengths

Previous research into green space qualities has tended to ignore co-occurrence. This is likely to underestimate the health effects of individual exposures due to possible synergistic effects in which the combined effect of exposures has a greater influence on health than the addition of separate exposures [8]. By clustering green space qualities, this research has laid the foundation for work that would analyse clustering with data for various chronic diseases especially from longitudinal studies to examine if a synergistic effect is present.

7. Limitations

While the access and topography domains were grouped together, topography could be interpreted differently by varying segments of the population. For example, for able-bodied people, steeper topography could be viewed positively for potentially allowing them to access more picturesque views, though this is also influenced by the type of topographic feature present [56]. For people with reduced mobility however, steep topography would represent an access barrier [57].

As described in Del Rosario *et al.*, there were issues with data quality, for example potential land cover misclassification and missing pathways in the roads dataset [5].

There may have also been edge effects related to excluding large areas of bushland/green spaces from analysis due to assumed issues with accessing these spaces as well as effects from excluding motorways [5]. However, people may still choose to access these larger bushland/green spaces which means areas near these natural environments may have recorded lower amounts of certain qualities than in reality. Additionally, the use of buffers around green spaces for calculating certain qualities could also result in double counting for green space polygons that were adjacent or contiguous [5].

The cluster analysis procedures were conducted pragmatically, based on domains of qualities that remained after those with high conceptual and empirical similarity were omitted. We acknowledge that potentially different cluster results may be possible had we retained all of the omitted qualities, or opted to derive clusters based on all qualities individually instead of the domains of qualities approach.

For the quality check, there were conceptual challenges for assessing lived experience for team members not involved in the GIS analysis/statistical analysis. For example, results were based on green space qualities that intersected with the network buffers around mesh blocks. We asked team members to think about the green spaces that fell within an 800 m or 1600 m road distance of mesh blocks, though if one green space had a lot of one type of domain e.g. amenities, but another green space had little amenities, these could cancel each other out. We framed green space size as "Size of green spaces is moderate" for example, despite the size being measured by the intersect area of the network buffer with green spaces which in many cases is different to the size of whole green spaces in an area. Size of green spaces was only included in the Amenity and Environment Scheme (total green space), and it is noted that other domains also influenced the quality check for this scheme. Additionally, the quality check exercise had a small sample size (5 team members, n = 17 areas).

The generalisability of this research beyond coastal NSW to inland areas is unknown. In other contexts, however, clustering could be adapted to inland areas by including riverine environments or other blue spaces rather than coastal areas. Challenges relating to data availability in different contexts may influence the generalisability of this research.

The cross-sectional design of the study means that causation of the socioeconomic patterning of clusters cannot be established. For example, it is unknown from the study design whether tree canopy increases property prices or whether less disadvantaged areas have attracted greater investment (including historical investment) in tree planting.

8. Conclusion

We have presented clustering schemes for analysing green space qualities relating to Amenity and Environment and Percentage-based Land Cover. Rather than analysing disparate green space qualities, we have provided a method to account for co-occurring qualities in potentially influencing health exposures, both positive and negative. By accounting for co-occurrence, potential compounding of disadvantage of green space health exposures can be gleaned in future. We found some key socioeconomic and geographic inequalities in the clustering of green space qualities. We found that leafy clusters are more likely to be in areas of lower disadvantage, whereas grassy clusters are more likely to be in areas of higher disadvantage. Many inland areas of the three cities were typically low in all domains except safety (Cluster 3, Amenity and Environment Scheme). Socioeconomic inequalities in these clusters indicate that a spatial targeting of combined green space qualities is needed. Future research directions could involve applying these clusters in longitudinal health studies.

Acknowledgments

This project was supported by an NHMRC-UKRI grant application (#GNT1192691). TAB is supported by an Australian Research Council Future Fellowship (#FT220100857). We thank Nicole Evangelidis and Annie Lacoba for assisting in the quality check exercise. This research used the NCRIS-enabled Australian Urban Research Infrastructure Network (AURIN) e-Infrastructure to access Australian Government Department of Social Services "DSS—National Public Toilets (Point) 2017" on 5 February 2024. The research used OpenStreetMap for map data (copyrighted OpenStreetMap contributors and available from https://www.openstreetmap.org).

Supplementary data

Supplementary data relating to domains and qualities, clustering methods and models, maps, quality check exercise, cluster results, radar graphs and quality check results are provided with this article.

Data availability statement

The data that support the findings of this study are not publicly available as the authors do not have permission to share data.

Declaration of generative AI and AI-assisted technologies

During the preparation of this manuscript, the authors used generative AI tools only to improve language and readability. Specifically, the authors used Copilot for language polishing only in limited sections. The authors take full responsibility for the content of the manuscript.

Authors' contribution

Conceptualization, T.A.B. and X.F.; methodology, M.N., L.D.R. and T.A.B.; software, M.N. and L.D.R.; formal analysis, M.N. and L.D.R.; investigation, M.N. and L.D.R.; resources, T.A.B. and X.F.; data curation, M.N. and L.D.R.; writing—original draft preparation, L.D.R. and M.N.; writing—review and editing, L.D.R., M.N., T.A.B. and X.F.; visualization, L.D.R. and M.N.; supervision, T.A.B. and X.F.; project administration, T.A.B. and X.F.; funding acquisition, T.A.B. and X.F. All authors have read and agreed to the published version of the manuscript.

Conflicts of interest

The authors declare no conflicts of interest.

References

[1] Nutsford D, Pearson AL, Kingham S. An ecological study investigating the association between access to urban green space and mental health. *Public Health* 2013, 127:1005–1011. DOI: 10.1016/j.puhe.2013.08.016

[2] Twohig-Bennett C, Jones A. The health benefits of the great outdoors: a systematic review and meta-analysis of greenspace exposure and health outcomes. *Environ. Res.* 2018, 166:628–637. DOI: 10.1016/j.envres.2018.06.030

[3] Astell-Burt T, Hartig T, Putra I, Walsan R, Dendup T, *et al*. Green space and loneliness: a systematic review with theoretical and methodological guidance for future research. *Sci. Total Environ.* 2022, 847:157521. DOI: 10.1016/j.scitotenv.2022.157521

[4] Huang W, Lin G. The relationship between urban green space and social health of individuals: a scoping review. *Urban For. Urban Green.* 2023, 85:127951. DOI: 10.1016/j.ufug.2023.127969

[5] Del Rosario L, Astell-Burt T, Navakatikyan M, Olsen JR, Caryl F, *et al*. Socioeconomic inequities within and between cities in objectively measured green space qualities at small geographical scales: evidence from Australia. *medRxiv* 2025. DOI: 10.1101/2025.04.09.25325554

[6] Nguyen PY, Astell-Burt T, Rahimi-Ardabili H, Feng X. Green space quality and health: a systematic review. *Int. J. Environ. Res. Public Health* 2021, 18(21):11028. DOI: 10.3390/ijerph182111028

[7] Bourke M, Wang HFW, McNaughton SA, Thomas G, Firth J, *et al*. Clusters of healthy lifestyle behaviours are associated with symptoms of depression, anxiety, and psychological distress: a systematic review and meta-analysis of observational studies. *Clin. Psychol. Rev.* 2025, 118:102585. DOI: 10.1016/j.cpr.2025.102585

[8] Blot WJ, Day NE. Synergism and interaction: are they equivalent? *Am. J. Epidemiol.* 1979, 110(1):99–100. DOI: 10.1093/oxfordjournals.aje.a112793

[9] Astell-Burt T, Feng X. Greener neighbourhoods, better memory? A longitudinal study. *Health Place* 2020, 65:102393. DOI: 10.1016/j.healthplace.2020.102393

[10] Kimpton A. A spatial analytic approach for classifying greenspace and comparing greenspace social equity. *Appl. Geogr.* 2017, 82:129–142. DOI: 10.1016/j.apgeog.2017.03.016

[11] Wu C, Yang S, Ma Y, Liu P, Ye X. Urban green space assessment: Spatial clustering method based on multisource data to facilitate

zoning planning. *J. Urban Plan. Dev.* 2024, 150(4):04024032. DOI: 10.1061/JUPDDM.UPENG-4865

[12] Hepburn J, Olsen JR, Mitchell R, Astell-Burt T, Feng X, *et al*. Park types and equity: who has access to what? A national assessment of multidimensional park qualities and socioeconomic disparities across Local Authorities in Great Britain. *Urban For. Urban Green*. 2025, 112:128930. DOI: 10.1016/j.ufug.2025.128930

[13] de Souza DC, Taconeli CA. Spatial and non-spatial clustering algorithms in the analysis of Brazilian educational data. *Commun. Stat. Case Stud. Data Anal. Appl.* 2022, 8(4):588–606. DOI: 10.1080/23737484.2022.2117744

[14] Knobel P, Dadvand P, Alonso L, Costa L, Español M, *et al*. Development of the urban green space quality assessment tool (RECITAL). *Urban For. Urban Green*. 2021, 57:126895. DOI: 10.1016/j.ufug.2020.126895

[15] Knobel P, Dadvand P, Maneja-Zaragoza R. A systematic review of multi-dimensional quality assessment tools for urban green spaces. *Health Place* 2019, 59:102198. DOI: 10.1016/j.healthplace.2019.102198

[16] Szabo B, Gollner E, Schnabel F. Topographic conditions and physical activity behaviour of young adults in Austria. *J. Hum. Sport Exerc.* 2014, 9(1):S210–S222. DOI: 10.14198/jhse.2014.9.Proc1.01

[17] Veitch J, Ball K, Rivera E, Loh V, Deforche B, *et al*. What entices older adults to parks? Identification of park features that encourage park visitation, physical activity, and social interaction. *Landsc. Urban Plan.* 2022, 217:104254. DOI: 10.1016/j.landurbplan.2021.104254

[18] Haahtela T, Alenius H, Lehtimäki J, Sinkkonen A, Fyhrquist N, *et al*. Immunological resilience and biodiversity for prevention of allergic diseases and asthma. *Allergy* 2021, 76(12):3613–3626. DOI: 10.1111/all.14895

[19] Weimann H, Rylander L, van den Bosch MA, Albin M, Skärbäck E, *et al*. Perception of safety is a prerequisite for the association between neighbourhood green qualities and physical activity: Results from a cross-sectional study in Sweden. *Health Place* 2017, 45:124–130. DOI: 10.1016/j.healthplace.2017.03.011

[20] Astell-Burt T, Feng X. Urban green space, tree canopy and prevention of cardiometabolic diseases: a multilevel longitudinal study of 46 786 Australians. *Int. J. Epidemiol.* 2020b, 49(3):926–933. DOI: 10.1093/ije/dyz239

[21] Wang H, Tassinary LG. Effects of greenspace morphology on mortality at the neighbourhood level: a cross-sectional ecological study. *Lancet Planet. Health* 2019, 3(11):e460–e468. DOI: 10.1016/S2542-5196(19)30217-7

[22] Ballesteros-Olza M, Gracia-de-Rentería P, Pérez-Zabaleta A. Effects on general health associated with beach proximity in Barcelona (Spain). *Health Promot. Int.* 2020, 35(6):1406–1414. DOI: 10.1093/heapro/daaa013

[23] Openshaw S. *The modifiable areal unit problem. Concepts and techniques in modern geography*. Norwich: Geo Books, 1984.

[24] Australian Bureau of Statistics. Digital boundary files: Australian Statistical Geography Standard (ASGS) Edition 3. 2021. Available: https://www.abs.gov.au/statistics/standards/australian-statistical-geography-standard-asgs-edition-3/jul2021-jun2026/access-and-downloads/digital-boundary-files (accessed on 18 September2023.)

[25] Australian Bureau of Statistics. General Community Profile for Urban Centres and Localities (UCL): New South Wales. 2021. Available: https://www.abs.gov.au/census/find-census-data/datapacks?release=2021&product=GCP&geography=UCL&header=S (accessed on 26 January 2026)

[26] Australian Bureau of Statistics. Mesh Blocks: Australian Statistical Geography Standard (ASGS) Edition 3. 2021. Available: https://www.abs.gov.au/statistics/standards/australian-statistical-geography-standard-asgs-edition-3/jul2021-jun2026/main-structure-and-greater-capital-city-statistical-areas/mesh-blocks (accessed on 14 August 2023)

[27] NSW Department of Planning, Housing and Infrastructure. Existing Green Assets. 2019. Available: https://datasets.seed.nsw.gov.au/dataset/c79c0f66-845b-4926-b750-1752f245f62b/resource/f98874cd-5d86-445d-bf0c-72944dd6740c/download/planning_green_assets.zip (accessed on 24 January 2024)

[28] Geoscape Australia. *Roads February 2020, NSW_STREET_LINE_shp.shp.* Canberra: Geoscape Australia Pty Ltd., 2020.

[29] Transport for NSW. *NSWBicycleNetwork5March2018.shp*. Sydney: Transport for NSW, 2018. Available: https://opendata.transport.nsw.gov.au/dataset/cycleway-data/resource/c29501ba-9b7f-4857-8d7f-8bdf2265201f (accessed on 19 February 2019)

[30] Geoscience Australia. Digital Elevation Model (DEM) of Australia derived from LiDAR 5 Metre Grid. 2015. Available: https://doi.org/10.26186/89644 (accessed on 25 January 2024).

[31] Transport for NSW. *Timetables Complete GTFS*. Sydney: Transport for NSW, 2023. Available: https://opendata.transport.nsw.gov.au/dataset/timetables-complete-gtfs (accessed on 25 January 2024).

[32] Australian Government Department of Social Services. *DSS—National Public Toilets (Point) 2017*. Melbourne: Australian Urban Research Infrastructure Network (AURIN), 2017. Available: https://data.aurin.org.au/dataset/au-govt-dss-national-public-toilets-2017-na (accessed on 5 February 2024)

[33] Sensis/Thryv Australia. *University of Wollongong_Australia Extract_11032020.xlsx*. Melbourne: Sensis/Thryv Australia, 2020.

[34] parkrun. Online map. 2024. Available: http://www.parkrun.com.au/events/ (accessed on 18 March 2024).

[35] Geofabrik GmbH, OpenStreetMap contributors. *Australia.shp.zip [australia-latest-free.shp.zip]*. Karlsruhe: Geofabrik GmbH, 2022. Available: https://download.geofabrik.de/australia-oceania/australia.html (accessed on 12 December 2022).

[36] NSW Department of Climate Change, Energy, the Environment and Water. *Biodiversity Values Map*. Parramatta: NSW Department of Climate Change, Energy, the Environment and Water, 2024. Available: https://datasets.seed.nsw.gov.au/dataset/biodiversity-values-map (accessed on 7 February 2024.)

[37] Australian Government Department of Climate Change, Energy, the Environment and Water. *Australia—Species of National Environmental Significance Distributions (public grids)*. Canberra: Australian Government Department of Climate Change, Energy, the Environment and Water, 2024. Available: https://fed.dcceew.gov.au/datasets/erin::australia-species-of-national-environmental-significance-distributions-public-grids/about (accessed on 7 February 2024).

[38] NSW Bureau of Crime Statistics and Research. *Crime Tool Hotspots*. Sydney: NSW Bureau of Crime Statistics and Research (BOCSAR), 2023. Available: https://bocsar.nsw.gov.au/statistics-dashboards/open-datasets/hotspot-maps---data.html (accessed on 5 February 2024).

[39] Burgess M. Understanding crime hotspot maps. 2011. Available: https://bocsar.nsw.gov.au/content/dam/dcj/bocsar/documents/publications/bb/bb01-100/bb60.pdf (accessed on 9 April 2025).

[40] Geoscape Australia. *GeoVision Surface Cover, 2m, December 2020*. Canberra: Geoscape Australia, 2020.

[41] Geoscape Australia. *GeoVision Trees, November 2021*. Canberra: Geoscape Australia, 2021.

[42] Australian Bureau of Statistics. Socio-Economic Indexes for Areas (SEIFA), Australia. Index data cubes. 2023. Available: https://www.abs.gov.au/statistics/people/people-and-communities/socio-economic-indexes-areas-seifa-australia/latest-release (accessed on 31 July 2023.)

[43] Australian Bureau of Statistics. Construction of the indexes: Socio-Economic Indexes for Areas (SEIFA): technical paper. 2023. Available: https://www.abs.gov.au/statistics/detailed-methodology-information/concepts-sources-methods/socio-economic-indexes-areas-seifa-technical-paper/2021/construction-indexes#technical-details-of-each-index-variables-and-loadings (accessed on 30 October 2025.)

[44] Geoscape Australia. Surface cover: product description: version 1.6. 2020.

[45] Office for National Statistics. Methodology and variables: details of the methodology and 2011 Census statistics used for the 2011 area classifications. 2018. Available: https://www.ons.gov.uk/methodology/geography/geographicalproducts/areaclassifications/

2011areaclassifications/methodologyandvariables (accessed on 8 October 2024).

[46] Lu X. An introduction to clustering techniques. 2018. Available: https://support.sas.com/resources/papers/proceedings18/2615-2018.pdf (accessed on 8 October 2024).

[47] Bhalla D. Cluster Analysis using SAS. 2024. Available: https://www.listendata.com/2014/10/cluster-analysis-using-sas.html (accessed on 8 October 2024).

[48] Rigolon A. A complex landscape of inequity in access to urban parks: a literature review. *Landsc. Urban Plan.* 2016, 153:160–169. DOI: 10.1016/j.landurbplan.2016.05.017

[49] Gerrish E, Watkins SL. The relationship between urban forests and income: a meta-analysis. *Landsc. Urban Plan.* 2018, 170:293–308. DOI: 10.1016/j.landurbplan.2017.09.005

[50] Fossa AJ, Zelner J, Bergmans R, Zivin K, Adar SD. Sociodemographic correlates of greenness within public parks in three US cities. *Wellbeing Space Soc.* 2023, 5:100157. DOI: 10.1016/j.wss.2023.100157

[51] Aerts R, Honnay O, Van Nieuwenhuyse A. Biodiversity and human health: mechanisms and evidence of the positive health effects of diversity in nature and green spaces. *Br. Med. Bull.* 2018, 127(1):5–22. DOI: 10.1093/bmb/ldy021

[52] White MP, Elliott LR, Gascon M, Roberts B, Fleming LE. Blue space, health and well-being: a narrative overview and synthesis of potential benefits. *Environ. Res.* 2020, 191:110169. DOI: 10.1016/j.envres.2020.110169

[53] Wolf ID, Wohlfart T. Walking, hiking and running in parks: a multidisciplinary assessment of health and well-being benefits. *Landsc. Urban Plan.* 2014, 130:89–103. DOI: 10.1016/j.landurbplan.2014.06.006

[54] Fleischhacker SE, Evenson KR, Rodriguez DA, Ammerman AS. A systematic review of fast food access studies. *Obes. Rev.* 2011, 12(5):e460–e471. DOI: 10.1111/j.1467-789X.2010.00715.x

[55] Hay GC, Whigham PA, Kypri K, Langley JD. Neighbourhood deprivation and access to alcohol outlets: a national study. *Health Place* 2009, 15(4):1086–1093. DOI: 10.1016/j.healthplace.2009.05.008

[56] Misthos LM, Nakos B, Krassanakis V, Menegaki M. The effect of topography and elevation on viewsheds in mountain landscapes using geovisualization. *Int. J. Cartogr.* 2019, 5(1):44–66. DOI: 10.1080/23729333.2018.1477569

[57] Lepoglavec K, Papeš O, Lovrić V, Raspudić A, Nevečerel H. Accessibility of urban forests and parks for people with disabilities in wheelchairs, considering the surface and longitudinal slope of the trails. *Sustainability.* 2023, 15(10):7741. DOI: 10.3390/su15107741