
Identifying urban functional zones by analysing the spatial distribution of amenities

Environment and Planning B: Urban Analytics and City Science
2023, Vol. XX(X) 1–16
© The Author(s) 2023
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/1ToBeAssigned
journals.sagepub.com/home/epb



Wei Chien Benny Chin

Singapore University of Technology and Design, Singapore

National University of Singapore, Singapore

Yuming Fu

National University of Singapore, Singapore

Kwan Hui Lim

Singapore University of Technology and Design, Singapore

Thomas Schroepfer

Singapore University of Technology and Design, Singapore

Lynette Cheah

Singapore University of Technology and Design, Singapore

Abstract

Comprehending urban form and function is crucial for effective urban planning and redevelopment. However, delineating urban functional zones and understanding the diverse activities within a city poses challenges. To address this, our study proposes a comprehensive analysis framework utilising point-of-interest (POI) data to identify and characterise urban functional zones and associated amenities. Leveraging geocoded user-generated content enables an effective capture of the spatial structure of human activities. We applied our framework in case studies in two Singapore locations, demonstrating its effectiveness in identifying urban functional zone shapes and examining associated amenities. The findings reveal a spatial configuration of amenities across the urban area, highlighting a diversity often differing from planned land-uses. This underscores the complex and non-uniform nature of the urban living experience, as observed variations in mixed uses challenge the homogeneity of intended land-use types. Our study and analysis framework provide a foundation for further investigations, including exploring travel behaviour patterns and assessing the vibrancy and vitality of urban areas.

Corresponding author:

Lynette Cheah, Engineering Systems and Design Pillar, Singapore University of Technology and Design, 8 Somapah Road, 487372, Singapore.

Email: lynette@sutd.edu.sg

Keywords

Point-of-interest, urban planning, mixed use, zoning, spatial partitioning

Introduction

Urban space is complex (Batty, 2013; Gallotti et al., 2021; Zhong et al., 2014). From the urban planner's view, each land parcel of a city can be assigned a specific land-use. From the residents' view, the urban space is composed of various types of amenities (Niu and Silva, 2021). A modern city is usually planned with the mixed development concept – to design a compact place where people can get access to most of their daily needs within a short travelling distance (Carlton, 2009; Shen and Wu, 2020). Thus, it would be difficult to categorise an urban space into a specific functional type from the experience of living in the space. For example, a square can contain a garden (leisure), some cafés (food and beverage, F&B), a sitting area (social gathering), and is located between a train station (transportation), a shopping mall (commercial), and a library (education and public services). These amenities give the space – the square – a series of functions and a mixed user experience for people. These situations indicate that the 'function' of an urban space is usually composed of mixed amenities (Niu and Silva, 2021). Indeed, some amenities tend to be co-located (Hidalgo et al., 2020; Gao et al., 2017). While urban space and amenities play a key role in shaping the identity of an urban area (Ye et al., 2021), the mixed nature of urban spaces makes it challenging to identify and delineate the 'urban functional zones' (Bettencourt and West, 2010; Batty, 2013).

Urban functional zones are subsections of an urban area that provide a specific service or a set of services (Niu and Silva, 2021; Zhong et al., 2014). The understanding of urban functional zones can be split into two parts: urban form and urban activity (Živkovic, 2019). The analysis of urban form focuses on the physical space and structure of the urban environment, such as buildings, streets (Lynch and Rodwin, 1958; Kropf, 2009) – the physical environment that provide space for establishing amenities. Various amenities located within the urban space provide different services and functions. The existence of these services and functions at a location enables specific urban activities (Jiang et al., 2015). Thus, analysing the spatial distribution of amenities is an effective way to reveal the spatial structure of urban functional zones (Hu et al., 2021; Hong and Yao, 2019).

As a long-standing challenge in urban studies, delineating urban functional zones is essential in the planning of an urban area (Hong and Yao, 2019; Zhong et al., 2014). To identify zones in a city, previous studies used data from user-generated sources to identify the daily activity space of people. Examples of data sources include location-based check-ins (Gao et al., 2017), mobile phone records (Gao et al., 2013), migration records (Hemmasi, 1980), smart-card data (Zhong et al., 2014), and bike-sharing trips (Niu and Jin, 2021). These studies analysed the routine and regularity of human movement between spaces, that is, where they usually go, and how the spatial units are connected by the frequent visits of some common strangers (Zhang et al., 2022). In urban geography literature, this type of spatial region is also defined as a 'functional region' – an area that covers both the supply of a service and the location where the demand is (Klapka et al., 2013). Other studies have made use of static data to identify the location of amenities. For instance, points-of-interest (POI) data (Gao et al., 2017), crowdsourced mapping data (Jiang et al., 2015), and satellite images (Chen et al., 2020) have been used to understand how a city can be split into different zones or 'formal regions' that provide specific sets of services. A formal region describes an area that has a consistent cultural or physical attribute and which properties are uniform within the zone.

Among previous studies on the delineation of urban functional zones, three types of data were frequently discussed: (1) POI datasets that capture different types of human activities, (2) remote satellite images that shows the land-use and land-cover of the urban surface, and (3) spatial interaction networks within an urban area that represent the strength of social interactions. Based on

Open Street Map data, Hong and Yao (2019); Wang et al. (2021) analysed the urban functional zones for each road blocks from POI data. Gao et al. (2017); Calafiore et al. (2021) integrated POI data and location-based social media data to analyse the context of human activities at different places within a city and delineate urban functional zones. Satellite images were used to analyse land-use/land-cover for identifying urban-rural structure. To further understand the detail functions and activities of each area, previous studies had fused several types of datasets, including POI data and social media data with the satellite images (Zhang et al., 2019; Niu and Jin, 2021; Yang et al., 2022). The human movement pattern within a city is also frequently used to identify neighbourhoods of strong internal flows, then the major functions of the places were identified using some semantic or topic modelling techniques (Wei and Wang, 2021; Zhang et al., 2022; Hu et al., 2021). Previous studies focused on recognising the urban function of areas with preset boundaries, including administrative areas (Wang et al., 2023), street blocks (Wang et al., 2021), and neighbourhoods (Hong and Yao, 2019). Urban functional zones can take any shape, regardless of urban planning or administrative boundaries. For example, Orchard Road in Singapore, Queen's Road in Hong Kong, Ximending in Taipei City, and Akihabara Electric Town in Tokyo developed along roads and intersections, with density decreasing with distance. This raises the question of 'where to draw boundaries for these zones', and highlights the importance of identifying their types and shapes directly from the amenities and their mix (Hidalgo et al., 2020). Recent studies have emphasised the level of mix in urban functional zones, with some zones identified as mixed-use and others as single-use (Wang et al., 2021; Wei and Wang, 2021). This stresses the need to retrieve zone functions from actual data extracted from human activities, such as POI and social media data, rather than relying on urban planning data, which often does not capture the detailed spatial variations within the same type of planning zone.

In dense cities, the urban landscape tends to offer a diversity of functions – a mixed use development (Schröpfer, 2020; Carlton, 2009). This increases the difficulty of delineating urban functional zones as the spatial configuration is complex with a mixed composition of functions. Singapore is a high-density and high-diversity country. Previous studies had used remote sensing (Sidhu et al., 2018) and household surveys (Zhong et al., 2014) to understand the land-use/land-cover and urban structure in Singapore. Singapore's limited land area forces a mixed and diverse spatial structure of urban functions, making planning and management difficult and complex. Most areas in Singapore contain multiple functions, while land use is typically assigned to a single category. From a planning perspective, assigning land use type in small spatial units is not useful for urban management and development. However, assessing the spatial structure of urban functions is important for managing and redeveloping the city after organic urban development. Therefore, the research question becomes: how can we partition urban space into zones and identify the mix of functions that are available in each zone?

Using two study sites in Singapore as case studies, the objective of this study is to develop a novel, analytical framework to map urban functional zones using POI data and discuss the composition of amenities in each zone. As a type of user-generated content, POI data can reveal the location of various services, thus providing a way to determine the spatial distribution of amenities, which reflects the functional use of each area (Hong and Yao 2019; Gao et al., 2017). Moreover, POI data reflects the current state of amenities, which can evolve over time, making it a valuable tool for analysing the distribution of amenities and delineating the boundaries of urban functional zones. Following this approach, we can identify the major urban functions and compare it with the planned land-uses. These analyses reveal the urban structure and relationships between planned land-use and actual human activities. The empirical results of identified urban functional zones will reveal the underlying structure of a city, and since POI data is generally available from various online map services, the proposed study framework is applicable to any other city.

Study framework and study sites

Analytical framework

The study contains three parts (Figure 1): (1) data and data pre-processing, (2) spatial partitioning of urban functional zones, and (3) analysis of the urban functional zones.

Study sites

To illustrate the universality of the framework, we chose two study sites with different area size, POI density, and development plans. Queenstown is a densely populated research hub with an area of about 21.7 km^2 . Punggol is a newly developed residential district with an area of 9.4 km^2 . In this study, land-use types are placed into 10 categories (Suppl. Figure S1): business, commercial, residential, mixed residential and commercial (mixed R&C), civic-education-medical combination (Civ./Edu./Med.), transportation, leisure-recreational combination (Lei./Rec.), waterbody, white, and other.

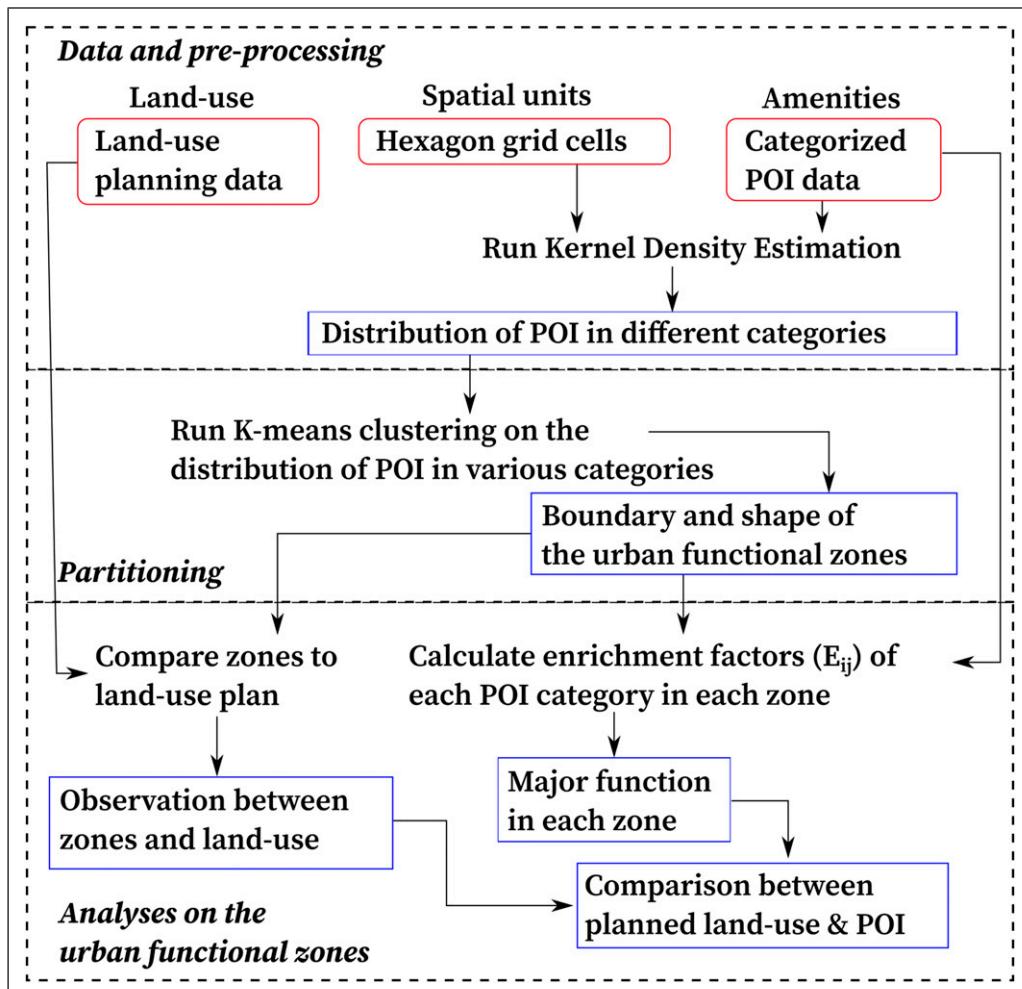


Figure 1. The analysis flowchart of this study.

Queenstown is located next to the Central Region of Singapore, that is, the Central Business District (CBD) of Singapore. It is not only a satellite residential town, but also serves as an education, innovation, research, and technology development core of Singapore. Punggol is a recent development consists of residential, commercial, and business parks at the North-East Region of Singapore. As a newly planned and developed town, the street network in Punggol forms a regular lattice structure. Each rectangular block is composed mainly of high-rise residential buildings and community facilities such as schools, community clubs, restaurants, and shopping centres.

The points-of-interest dataset

In this study, POI data is used as a surrogate for the location of amenities. The POI data was downloaded in 2021 (July to August) from five map sources – two commercial online mapping platforms (Google Map and HERE Map), an open-source project (OpenStreetMap), the official map source maintained by the Singapore government (OneMap), and the licenced Street Directory data set from the Singapore Land Authority (SLA). We used as many available data sources as possible to capture the most comprehensive spatial distribution of amenities in the two study sites. These POI datasets from the five sources were conflated using a framework described in [Low et al. \(2021\)](#) to merge the data and remove any duplicates. The conflated data was subsequently reviewed and validated ([Yeow et al., 2021](#)).

The POI data was categorised into eight categories: business, food and beverage (F&B), commercial, education, transportation, outdoor, health, and others ([Suppl. Table S1–S2](#)). A summary of the POI dataset is shown in [Suppl. Table S3](#). Nearly half of the POIs in the conflated data were labelled as ‘premises’, ‘point-of-interest’, or ‘establishment’, which could not be categorised, so they were excluded from the analysis. It is worth noting that Queenstown has more POIs than Punggol, likely due to its larger size, more developed state, and proximity to the Central Region.

The spatial distribution of various amenities

In this study, a hexagonal grid with 50 m spacing was generated and each cell was used as the spatial unit. Because POI data consists of groups of point locations scattered across the study sites ([Supp. Figure S2\(c\)](#)), the actual location of POI and its relative location (aggregate by cells) may not be useful in presenting the spatial pattern. To uncover the spatial distribution of the different categories POI, the Kernel Density Estimation (KDE) model was used ([Kloog et al., 2009](#); [Adolphson, 2010](#)). Each amenity point has an accessible area that could be captured by an accessible distance parameter. The calculation of KDE (Scikit-Learn, v0.22.1) estimates the density of a specific group of points for each spatial unit (hexagon cell). In the KDE calculation, each spatial unit is used to search and measure the distance from the cell centre to the POIs. The weight for each point for the spatial unit was calculated using the distance and a kernel function. A bandwidth parameter is set to determine the shape of the kernel function, which indicates a decreasing trend from the centre of the bell shape to outward area. If a point is nearer to the centre of the kernel, the weight is higher. The density of the cell is calculated based on the weights of the points. In the resulting KDE map, the more points of a specific POI are closer to a cell, the higher the density, which indicated that the higher the accessibility ([Supp. Figure S2\(b\)](#)). We use the KDE of a specific type of amenity to represent the spatial accessibility to that amenity ([Adolphson, 2010](#)).

The bandwidth parameter was set to be 200 m, and the Gaussian function was used as the kernel function. 400 m is considered as a walkable distance ([Hou et al., 2020](#)). Here we used half of the walkable distance as the bandwidth to determine the shape of the Gaussian kernel function, which indicated that the POIs within 200 m would get a relatively high weight, and the weights decrease with a steep slope beyond 200 m. In other words, the POIs within 200 m can be considered as accessible; the accessibility beyond 200 m decreased quickly with distance ([Suppl. Figure S3–\(s\)4](#) and [Table S4](#)).

Partitioning the study sites into functional zones using POI data

Clustering analysis process

For each hexagon cell in the study area, density values were calculated for each of the seven POI categories. For each of the two study sites, the clustering analysis was done on the hexagon cells using the seven density values as vectors. While the seven POI categories indicate different types of human activities, the clustering analysis can capture how these amenities and human activities are mixed and distributed across space. The resulting clusters indicate that cells in the same cluster would have similar amenity density mix – they serve a similar urban function within the study site. For example, a shopping mall may have a high density of commercial and F&B points and a low density of outdoor and transportation facilities. The identification of clusters based on amenity density could reveal the composition of amenities and boundaries of these spatial clusters. These spatial clusters represent the urban functional zones in this study. The clustering analysis is performed separately for each study site. This is because the two independent sites were designed with different objectives, thus, the composition of amenities can have a different patterns.

The process of KDE calculation has considered the spatial neighbouring structure, that is, cells within a proximity range would find a similar set of neighbouring POIs and have similar distances to those points; thus, the density has a smoothing effect for neighbouring cells (Lum, 2008). Therefore, the spatial continuity can be neglected in the clustering analysis process. The spatial units in the same cluster would be adjacent to each other and not be isolated because of the smoothing effect from the KDE process. In addition, it is possible for places in different parts of the study site to serve the same urban function. These places would be shown as separated groups of hexagon cells belonging to the same cluster – multipart polygons.

In this study, the k -means clustering approach (Scikit-Learn, v0.22.1) is used to identify the urban functional zones (Gao et al., 2017). The k -means clustering can be used to identify clusters of data points (hexagon cells) with multiple vectors (the seven density values) using an iterative process. The number of clusters/zones (k) was decided based on a pretest using the elbow method (Suppl. Figure S5), and set to seven and five for Queenstown and Punggol, respectively. In the pretest, the changes of the k parameter from low to high presented a hierarchical structure that the added zones tend to split from the previous zones. In other words, the shapes of zones were similar between different k .

The boundary of urban functional zones

Seven and five functional zones were identified for the two study sites (Figure 2). The zones were sorted by POI density, which differed across the sites since the clustering analysis was performed separately. Table 1 summarises the identified functional zones, including their area size, number of POIs, and number of parts. The spatial continuity feature was not considered in the k -means clustering process, so separated areas could have similar spatial compositions of amenities (the same function). The ‘No. Parts’ column counts the number of separated parts belonging to the same urban functional zone. The last few zones in Table 1 had a larger area size or more separated parts, indicating that they served as buffer zones between the core regions and were thus scattered throughout the study sites. In contrast, the first few zones had a smaller number of parts, indicating that their function was usually unique within the study site and occupied a smaller portion of the area. The number of POI in each zone was shown in Suppl Table S5–S6.

Figure 2 shows the spatial distribution of the clustering zones. Despite the differences in the composition, the results showed that the functional zones could capture the spatial patterns of the urban structure. Some core regions with more intense concentration of amenities could be observed. Some peripheral regions or buffer zones that filled up the spaces between core regions were also

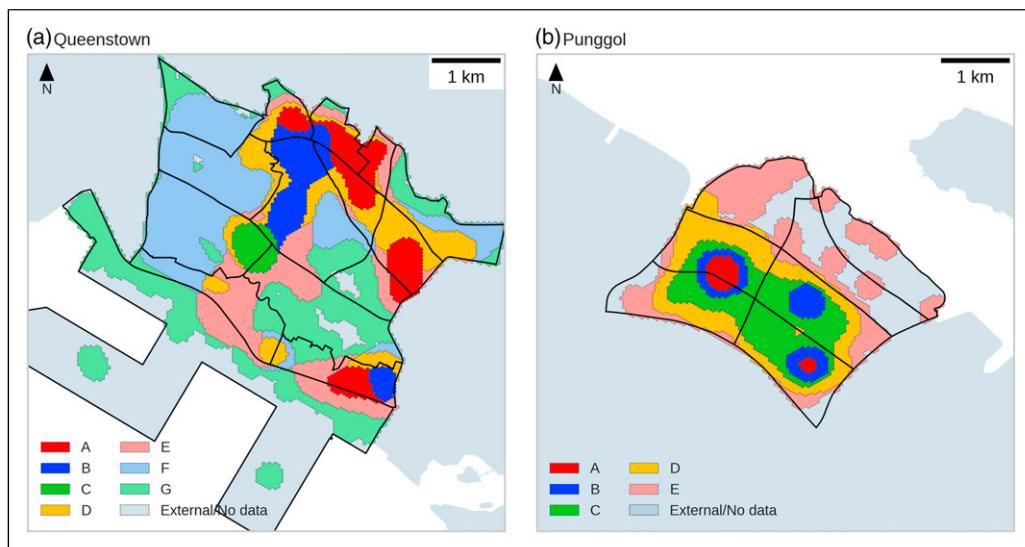


Figure 2. The spatial distribution of urban functional zones in (a) Queenstown, and (b) Punggol. The amenity compositions of the zones in the two sites are different.

Table 1. A summary of the sizes of the 12 zones (sorted by POI density) in the two sites: area size (of hexagons cells proximate to at least one POI, in km^2), the number of individual parts, and density (total number of POI/ km^2) for each zone.

Sites	Queenstown				Punggol				
	Zone	No. Parts	POI	Area	Density	No. Parts	POI	Area	Density
A	3	1,651	1.440	1,146	1,146	2	357	0.247	1,446
B	2	1,154	1.117	1,032	1,032	3	263	0.641	410
C	1	307	0.383	801	801	1	689	1.868	368
D	6	911	2.089	436	436	2	490	1.949	251
E	12	826	2.637	313	313	11	131	2.786	47
F	5	963	4.291	224	224	-	-	-	-
G	23	340	5.019	67	67	-	-	-	-
Total	52	6,152	16.976	362	362	19	1,930	7.491	258

revealed. For instance, in Queenstown (Figure 2(a)), there were three core regions: (1) the eastern region that contained Zone-A, (2) the northern area that emerged as Zone-A and -B, and (3) the centre of the map that was mainly Zone-C and which connected to the second core through Zone-B. The eastern side of the third core (Zone-C) was composed of Zone-E and -G, and the western side was composed of Zone F – these places had lower densities.

Analysing the functional zones in the two study sites

The major functions of the zones

Enrichment factor (Verburg et al., 2004) is used as an indicator of the main types of amenities within each urban functional zone. Enrichment factor was used in previous studies to identify the main

types of POIs that play significant roles within a region (Chen et al., 2017; Niu and Silva, 2021). The calculation for the enrichment of j POIs at zone i is shown in equation (1)

$$E_{ij} = \frac{n_{ij}/n_i}{N_j/N_T} = N_T \times \frac{n_{ij}}{n_i \times N_j} \quad (1)$$

where n_{ij} is the number of j POI located in zone i ; n_i is the total number of POI within zone- i ; N_j is the total number of j POI in the study area; and N_T is the total number of all POI in the study area. N_T is a constant for each study site. The enrichment factor can be compared across the same category (that N_j would be fixed) or across the same zone (n_i would be fixed). When comparing across the same category, the enrichment factors would show the proportion of POI of a specific category in each zone (n_{ij}/n_i) – a higher proportion indicates that this POI category is the main amenity within this zone (i). Even when n_{aj} is larger than n_{bj} , if n_a is significantly larger than n_b , this would lead to $E_{aj} < E_{bj}$. This situation indicates that the j POI has stronger effect in zone- b than its effect in zone- a because zone- a has a lot more POI in other categories. For example, the F&B POI in Zone-A (641, Supp. Table S5) was slightly more than Zone-B (566), but the total count of POI in Zone-A is +497 more than Zone-B, making its F&B POI enrichment factor (0.96, Figure 3) to be lower than Zone-B's value (1.21). This suggested that although the two zones contained similar number of F&B POI, the F&B played more important role in Zone-B since it covered 49% of the total POI in the zone; F&B covered only 39% of POI in Zone-A, making it less significant in the area. Similarly, when comparing across the same zone, a higher enrichment factor (E_{ij}) indicates that compared to other categories (non- j), most of the j POI in the study site is located within zone- i . For example, the F&B and commercial POI in Zone-A were both around 640, but the total F&B POI were higher (2,487) than commercial POI (1,516) in Queenstown, leading to a higher commercial enrichment factor for Zone-A (1.57) compared to its moderate F&B value (0.96).

The POI count in different categories varies from hundreds to thousands (Suppl Table S3) – the amenity categories with smaller total numbers would be underestimated in all zones if we compare the count of POI directly. This implies that the metrics based on the proportion may be problematic

(a) Queenstown								(b) Punggol								
	A	1.57	0.70	0.79	0.45	0.80	0.44		A	1.08	1.41	1.80	1.62	0.30	0.45	0.20
A	0.96	1.57	0.70	0.79	0.45	0.80	0.44	B	1.21	0.78	1.29	1.28	0.87	0.39	0.56	
B	1.21	0.78	1.29	1.28	0.87	0.39	0.56	C	1.27	0.45	0.91	2.99	0.39	0.74	0.05	
C	1.27	0.45	0.91	2.99	0.39	0.74	0.05	D	0.86	0.90	0.98	1.17	1.26	1.18	1.68	
D	0.86	0.90	0.98	1.17	1.26	1.18	1.68	E	0.93	1.31	1.35	0.87	0.39	1.18	0.46	
E	0.93	1.31	1.35	0.87	0.39	1.18	0.46	F	1.01	0.37	0.91	0.43	2.67	1.72	1.67	
F	1.01	0.37	0.91	0.43	2.67	1.72	1.67	G	0.74	0.79	1.01	0.78	0.71	1.30	3.70	
G	0.74	0.79	1.01	0.78	0.71	1.30	3.70	F&B								
	F&B	Commercial	Business	Health	Education	Transportation	Outdoor	F&B	Commercial	Business	Health	Education	Transportation	Outdoor		

Figure 3. The enrichment factor (E_{ij}) matrix for the POI types in each zone for the two sites.

given this situation. Therefore, in our study, the enrichment factor is appropriate for cross-amenities comparison and cross-zonal comparison because it includes both the total number of POI in each category (N_j) and in each zone (n_i). We used the enrichment factor metric for the subsequent zone function identification and similarity analysis across study sites.

The matrix of enrichment factors for all combinations between amenities and zones is presented in [Figure 3](#). The enrichment factors could be compared and discussed in both directions: across different zones for an amenity (columns in [Figure 3](#)), and across various amenities for a zone (rows in [Figure 3](#)).

The enrichment factors of Queenstown were mostly in light colours and only one to three cells in each row had higher values. This means that most of the zones in Queenstown served some specific purposes. For example, Zone-A was mainly commercial, whereas the health function was mainly in Zone-C and education in Zone-F. The enrichment factor of F&B in most zones was relatively high compared to other functions, indicating that most area had some level of F&B function in Queenstown. The values for Punggol were generally higher, indicating that most zones contain multiple functions. Punggol was planned and built as a highly integrated residential town. Because houses and apartments would usually not be recorded as a POI, this led to a lower number of POIs in the area. Thus, other types of amenities became the major functions, which indicated a diverse urban function pattern. In other words, besides residential, other types of amenities were mixed and highly integrated in Punggol.

Based on the enrichment factors, the major functions in each zone are summarised in [Suppl Table S7](#). In Queenstown ([Figure 4](#)), Zone-A – a commercial area – is at the west, neighbouring the Central Region of Singapore. Located beside Zone-A and Zone-C, Zone-B is a business park region ([Suppl. Figure S1\(a\)](#)). The major functions here were F&B, business, and health. Zone-C is a hospital, which also contains a slightly higher F&B enrichment factor. This might be because these types of amenities could usually attract more users. Zone-F covered some campuses of high education and research centres, leading to a high education enrichment factor. The other zones are the buffer regions or the outer regions of the core area that contain multiple functions. The 3D map for Punggol is shown in [Suppl. Figure S8](#).

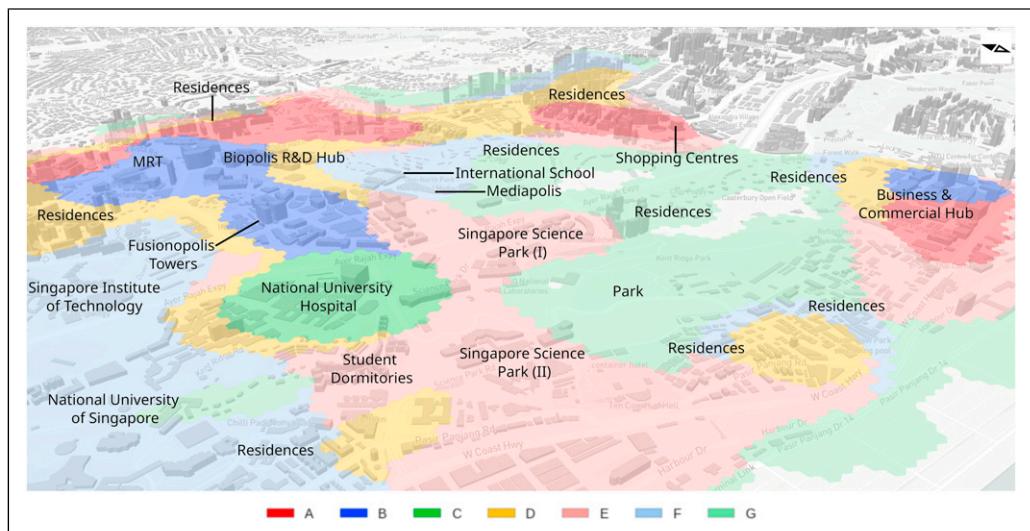


Figure 4. Urban functional zones in Queenstown. Some key locations are labelled on the 3D map. The map was generated using Mapbox and OpenStreetMap.

The planned land-use type for each zone

While functional zones were delineated based on the spatial distribution of amenity categories, the relationship between the planned land-use type and the zones remains unclear. We calculated the proportion of area for various land-use types in each zone and the enrichment factors (Figure 5). In both study sites, transportation land-use had occupied a relatively large proportion of area in almost all of the zones (about 0.2 to 0.3). Streets and MRT/LRT stations were uniformly distributed within the study sites. Other than transportation, each zone tended to be dominated by one or two land-use types. In Queenstown, Zone-B and -E were mainly business areas; Zone-A and -D contained more

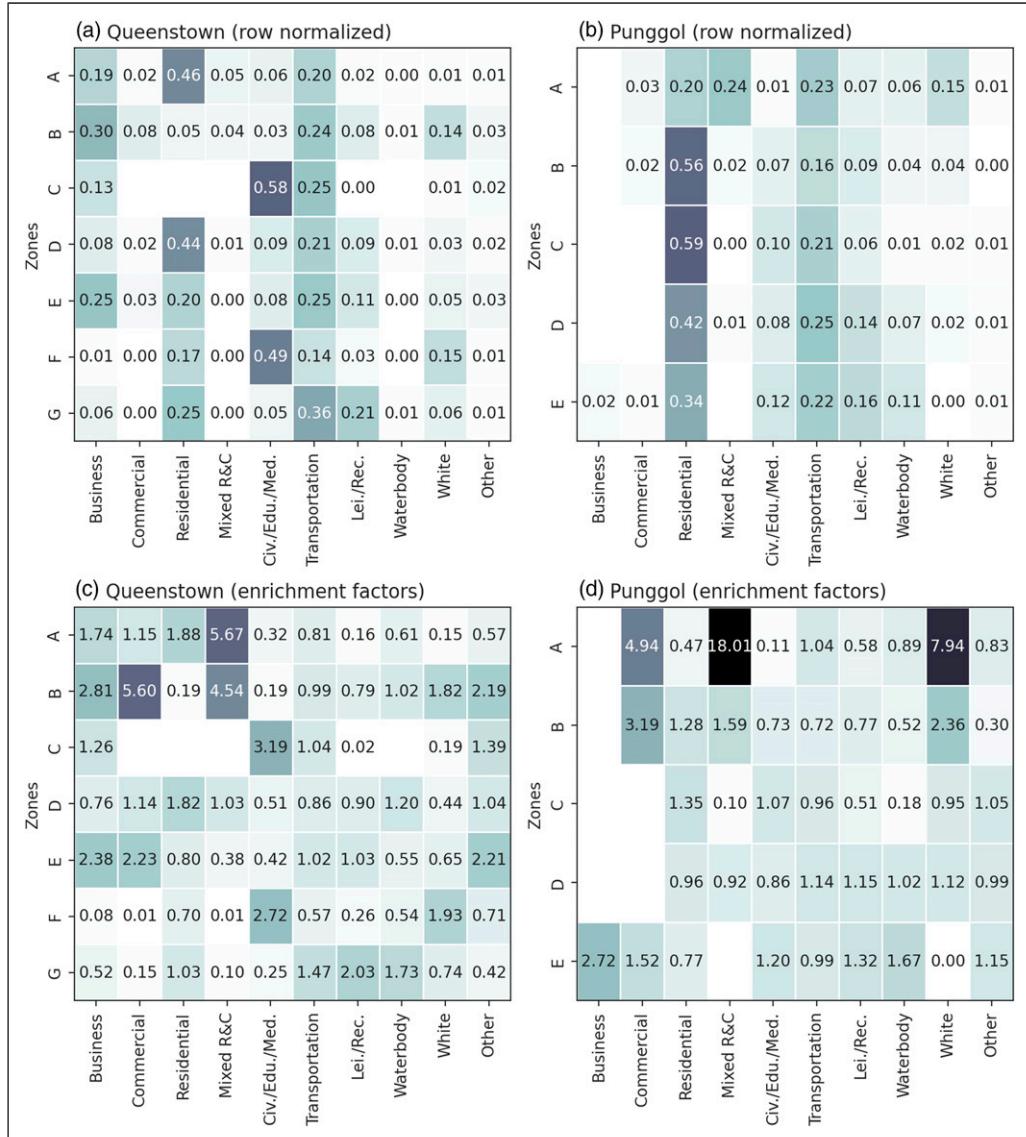


Figure 5. The matrix of identified clusters (rows) and the planned land-use types (columns) for the two sites. The top row shows the row standardised value in each cell indicates the proportion of individual cluster area that falls within the 10 land-use types and the bottom row shows the enrichment factor of land-uses' areas.

residential areas; Zone-C and -F were civic/education/medical areas. In Punggol, most of the zones contained large proportion of residential land-use, except Zone-A (a mixed land-use of residential and commercial).

Enrichment factors ([Figure 5\(c\) and 5\(d\)](#)) revealed slightly different results than proportions. High enrichment factors were mainly located in the zones with denser activities, while transportation had low enrichment factors in both sites. In Queenstown, mixed R&C land-use had a stronger concentration within the Zone-A and Zone-D, while commercial land-use had high enrichment factors in Zones-B and -E, indicating that these zones covered most of the commercial land-use type within the study site. In Punggol, residential land-use was no longer the major land-use type within the zones, but business, commercial, mixed R&C, and white land-uses stood out, indicating that some of the zones included most of these land-use types within the site and formed some clustering pattern. The maps of the land-use and POI enrichment factors are presented in [Suppl. Figures S9–\(s\)10](#).

Discussions and conclusion

This study presents a framework with two objectives: (1) identifying urban functional zone boundaries and (2) determining zone functions. Two case studies in Singapore were conducted to delineate urban functional zones based on the spatial distribution of different amenity categories. By leveraging social sensing ([Liu et al., 2015](#)) and Volunteered Geographic Information ([Jiang and Thill, 2015](#)) techniques, real-life experiences from crowdsourced data were captured from geocoded user-generated content, specifically POIs. POIs in seven categories served as a proxy for diverse amenities representing various human activities. KDE were applied to assess the accessibility to various types of amenities. Then, k -means clustering analysis was performed to identify urban functional zone boundaries, which denote areas with similar amenity compositions. The dominant amenity type and mixture within each zone were explored using enrichment factor matrices. As the primary input data, POI data is widely available for most cities worldwide, making the proposed framework applicable beyond the study area.

KDE was used to model accessibility as a smooth surface with higher density concentrated at locations with more points of a specific POI. This resulted in a blurred boundary for the amenity service area and opened up the possibilities of identifying different boundaries at the periphery areas, namely, the low-density areas or buffer zones. Therefore, the choice of the k value in the k -means clustering algorithm is critical in practice. A higher k may generate some secondary urban functional zones that surround the high-density area, pushing the boundary outwards from the core area. For example, Zones-E, -F, and -G in Queenstown were identified as buffer zones. The elbow method was used in this study to determine the optimal k . However, there are other methods that can be used, depending on the specific purpose of the zoning. While the resulting patterns may be similar ([Suppl. Figure S6–\(s\)7](#)), the modifiable boundaries may cause the modifiable areal unit problem ([Goodchild, 2022](#)).

Mixed development is the de facto trend in highly urbanised cities, emerging dynamically and organically. Different urban functions overlap to form composite urban functional zones ([Hidalgo et al., 2020; Wang et al., 2021](#)). This study proposes a framework to reveal the spatial structure of urban functional zones with mixed functions. The framework assumes urban functions are mixed from the outset and calculates accessibility to different types of amenities using KDE, followed by clustering analysis to delineate the shapes of urban functional zones with various compositions. Key strengths of this framework include no preset boundaries and the ability to capture differences between places with similar functions/programmes but different intensity/composition.

Human activity in urban space is complex and dynamic ([Batty, 2013; Gallotti et al., 2021](#)), and urban functional zones are the outcome of the organic urban development process. While urban

land-use planning is the starting point, the identified urban functional zones often differ from planned land-use distribution. In our results, several zones were detected within the same planned land-use, with similar programs but different activity levels. This can be caused by the existence or lack of important landmarks. While modelling and predicting the dynamic development process from planned land-use to functional zones is difficult due to the complexity, this study demonstrates that retrieving urban functional zones from actual activities is possible.

Understanding the spatial structure of urban functions would have benefits for long-term and short-term urban planning. For long-term urban planning (urban revitalisation and urban redevelopment), understanding the current situation could facilitate the simulation of scenarios of future designs and plans (Li et al., 2021). Analysts can explore different possible combinations of POI and assess the resulting urban functional zone patterns. For short-term planning, the questions of what is provided at different location, who are the users, and where these users are coming from, become critical in the management of communities and relevant resource allocation. This study provides a partitioning approach according to the services that indicates the spatial distribution of resources and supports the allocation of essential resources. For example, during an epidemic, the spatial partitioning for movement controls and the allocation of medical resources can benefit from this perspective. Regarding the spatial partitioning process, decision makers may need to know the boundaries of urban functional zones that correspond to the spatial structure of amenities, adopting a zonal approach to restrict human movement. In addition, places where people gather or access their daily essential goods are more ideal for distributing goods and disseminating information.

Urban functional zones and land-use are two distinct concepts and both are pivotal in urban planning. Land-use planning optimises resource allocation, promoting sustainability and efficiency in urban areas. Urban functional zones classify dynamic urban functions and activities, offering insights into functional diversity and supporting efficient planning and resource allocation (Liu et al., 2020; Niu and Silva, 2021). Land-use zoning guides city layout, while urban functional zones, shaped by spatial distribution of amenities and services, support urban life. Land-use zoning categorises land by purpose, while urban functional zoning focuses on urban dynamic. Both concepts are vital for creating liveable cities, optimising land use, and supporting diverse urban activities (Niu and Jin, 2021).

By identifying urban functional zones and amenity composition in the two study sites in Singapore, we observed that the urban functional zones are not always correlated with planned land-use. For example, Zone-A in Queenstown is mainly occupied by commercial amenities, but almost none of the area is planned for commercial land-use; instead, most of the area is planned for business, residential, and transportation, for which enrichment factors are relatively low. Indeed, detailed information about urban space that provide higher resolution than the land-use data can bring benefits for the urban revitalisation and redevelopment (Niu and Jin, 2021; Hong and Yao, 2019). Planned and actual land-uses can differ, indicating a mismatch between top-down (planning and governance) and bottom-up (user experiences) perspectives. Planned land use only captures the main function of a place, while places are always mixed with different functions, especially in high-density urban areas (Hidalgo et al., 2020). This highlights the value of using user-generated data, such as POI data, to analyse urban spatial structure and understand the evolving use and functions of urban spaces.

The presented analysis has several limitations. We did not compare the results with other socioeconomic-demographic, and human activity due to the study scope, which focuses on identifying urban functional zones. Future research can integrate urban statistics to enhance urban process modelling. As there was no alternative validation data available, we couldn't provide accuracy/precision analysis for the classification. Future studies can include data sources, such as social media (Gao et al., 2017) and movement data (Niu and Jin, 2021) for validation. Although POI data served as a surrogate for spatial distribution of amenities, some points like residential units and industrial facilities were missing. Future research can include additional data types to address this

limitation. Moreover, this study did not consider the capacity of each amenity, treating every POI equally. This approach may underestimate amenities with a large footprint. Future studies should incorporate capacity and service distance variations into the data. The edge effect near boundaries may underestimate accessibility to amenities located outside the study area. Future efforts can expand the study area to cover surrounding spaces and focus on the designated region to mitigate the edge effect. Lastly, population dynamics, an important aspect of urban function, were not considered due to the difficulty in collecting human activity intensity data. Future research can incorporate population dynamics into the partitioning of urban functional zones.

Acknowledgements

The authors would like to thank Raymond Low for preparing and processing the raw data of POI in the three study sites.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by a grant from the Singapore University of Technology and Design (Cities: Urban Science and Design for Density, PIE-SGP-CTRS-1803).

ORCID iDs

Wei Chien Benny Chin  <https://orcid.org/0000-0001-7215-3303>

Lynette Cheah  <https://orcid.org/0000-0001-6312-0331>

Supplemental Material

Supplemental material for this article is available online.

References

- Adolphson M (2010) Kernel densities and mixed functionality in a multicentred urban region. *Environment and Planning B: Planning and Design* 37(3): 550–566. DOI: [10.1068/b35061](https://doi.org/10.1068/b35061).
- Batty M (2013) *The New Science of Cities*. Cambridge, MA: MIT press. ISBN 978-0262534567. DOI: [10.7551/mitpress/9399.001.0001](https://doi.org/10.7551/mitpress/9399.001.0001).
- Bettencourt L and West G (2010) A unified theory of urban living. *Nature* 467(7318): 912–913. DOI: [10.1038/467912a](https://doi.org/10.1038/467912a).
- Calafiore A, Palmer G, Comber S, et al. (2021) A geographic data science framework for the functional and contextual analysis of human dynamics within global cities. *Computers, Environment and Urban Systems* 85: 101539. DOI: [10.1016/j.compenvurbsys.2020.101539](https://doi.org/10.1016/j.compenvurbsys.2020.101539).
- Carlton I (2009) *Histories of transit-oriented development: Perspectives on the development of the TOD concept. Technical report*. Berkeley, CA: Working Paper. <http://hdl.handle.net/10419/59412>
- Chen Y, Liu X, Li X, et al. (2017) Delineating urban functional areas with building-level social media data: a dynamic time warping (DTW) distance based k-medoids method. *Landscape and Urban Planning* 160: 48–60. DOI: [10.1016/j.landurbplan.2016.12.001](https://doi.org/10.1016/j.landurbplan.2016.12.001).
- Chen THK, Qiu C, Schmitt M, et al. (2020) Mapping horizontal and vertical urban densification in Denmark with Landsat time-series from 1985 to 2018: a semantic segmentation solution. *Remote Sensing of Environment* 251: 112096. DOI: [10.1016/j.rse.2020.112096](https://doi.org/10.1016/j.rse.2020.112096).

- Gallotti R, Sacco P and De Domenico M (2021) Complex urban systems: challenges and integrated solutions for the sustainability and resilience of cities. *Complexity* 2021: 1–15. DOI: [10.1155/2021/1782354](https://doi.org/10.1155/2021/1782354).
- Gao S, Liu Y, Wang Y, et al. (2013) Discovering spatial interaction communities from mobile phone data. *Transactions in GIS* 17(3): 463–481. DOI: [10.1111/tgis.12042](https://doi.org/10.1111/tgis.12042).
- Gao S, Janowicz K and Couclelis H (2017) Extracting urban functional regions from points of interest and human activities on location-based social networks. *Transactions in GIS* 21(3): 446–467. DOI: [10.1111/tgis.12289](https://doi.org/10.1111/tgis.12289).
- Goodchild MF (2022) The openshaw effect. *International Journal of Geographical Information Science* 36(9): 1697–1698. DOI: [10.1080/13658816.2022.2102637](https://doi.org/10.1080/13658816.2022.2102637).
- Hemmasi M (1980) The identification of functional regions based on lifetime migration data: a case study of Iran. *Economic Geography* 56(3): 223–233.
- Hidalgo CA, Castañer E and Sevtsuk A (2020) The amenity mix of urban neighborhoods. *Habitat International* 106: 102205. DOI: [10.1016/j.habitatint.2020.102205](https://doi.org/10.1016/j.habitatint.2020.102205).
- Hong Y and Yao Y (2019) Hierarchical community detection and functional area identification with OSM roads and complex graph theory. *International Journal of Geographical Information Science* 33(8): 1569–1587. DOI: [10.1080/13658816.2019.1584806](https://doi.org/10.1080/13658816.2019.1584806).
- Hou Y, Yap W, Chua R, et al. (2020) The associations between older adults' daily travel pattern and objective and perceived built environment: a study of three neighbourhoods in Singapore. *Transport Policy* 99: 314–328. DOI: [10.1016/j.tranpol.2020.06.017](https://doi.org/10.1016/j.tranpol.2020.06.017).
- Hu S, Gao S, Wu L, et al. (2021) Urban function classification at road segment level using taxi trajectory data: a graph convolutional neural network approach. *Computers, Environment and Urban Systems* 87: 101619. DOI: [10.1016/j.compenvurbsys.2021.101619](https://doi.org/10.1016/j.compenvurbsys.2021.101619).
- Jiang B and Thill JC (2015) Volunteered geographic information: towards the establishment of a new paradigm. *Computers, Environment and Urban Systems* 53: 1–3. DOI: [10.1016/j.compenvurbsys.2015.09.011](https://doi.org/10.1016/j.compenvurbsys.2015.09.011).
- Jiang S, Alves A, Rodrigues F, et al. (2015) Mining point-of-interest data from social networks for urban land use classification and disaggregation. *Computers, Environment and Urban Systems* 53: 36–46. DOI: [10.1016/j.compenvurbsys.2014.12.001](https://doi.org/10.1016/j.compenvurbsys.2014.12.001).
- Klapka P, Halás M and Tonev P (2013) Functional regions: concept and types. In: 16th international colloquium on regional sciences. Klímová V and Žítek V (Eds.), *Valtice, Czech Republic, 1994–21101 June 2013*, pp, 94–101. Brno: Masaryk University Press. DOI: [10.5817/CZ.MUNI.P210-6257-2013-11](https://doi.org/10.5817/CZ.MUNI.P210-6257-2013-11)
- Kloog I, Haim A and Portnov BA (2009) Using kernel density function as an urban analysis tool: Investigating the association between nightlight exposure and the incidence of breast cancer in Haifa, Israel. *Computers, Environment and Urban Systems* 33(1): 55–63. DOI: [10.1016/j.compenvurbsys.2008.09.006](https://doi.org/10.1016/j.compenvurbsys.2008.09.006).
- Kropf K (2009) Aspects of urban form. *Urban Morphology* 13(2): 105–120. DOI: [10.51347/jum.v13i2.3949](https://doi.org/10.51347/jum.v13i2.3949).
- Li M, Liu J, Lin Y, et al. (2021) Revitalizing historic districts: identifying built environment predictors for street vibrancy based on urban sensor data. *Cities* 117: 103305. DOI: [10.1016/j.cities.2021.103305](https://doi.org/10.1016/j.cities.2021.103305).
- Liu Y, Liu X, Gao S, et al. (2015) Social sensing: a new approach to understanding our socioeconomic environments. *Annals of the Association of American Geographers* 105(3): 512–530. DOI: [10.1080/00045608.2015.1018773](https://doi.org/10.1080/00045608.2015.1018773).
- Liu H, Xu Y, Tang J, et al. (2020) Recognizing urban functional zones by a hierarchical fusion method considering landscape features and human activities. *Transactions in GIS* 24(5): 1359–1381. DOI: [10.1111/tgis.12642](https://doi.org/10.1111/tgis.12642).
- Low R, Tekler ZD and Cheah L (2021) An end-to-end point of interest (POI) conflation framework. *ISPRS International Journal of Geo-Information* 10(11): 779. DOI: [10.3390/ijgi10110779](https://doi.org/10.3390/ijgi10110779).
- Lum C (2008) The geography of drug activity and violence: Analyzing spatial relationships of non-homogenous crime event types. *Substance Use & Misuse* 43(2): 179–201. DOI: [10.1080/10826080701690573](https://doi.org/10.1080/10826080701690573).
- Lynch K and Rodwin L (1958) A theory of urban form. *Journal of the American Institute of Planners* 24(4): 201–214. DOI: [10.1080/01944365808978281](https://doi.org/10.1080/01944365808978281).

- Niu N and Jin H (2021) Integrating multiple data to identify building functions in China's urban villages. *Environment and Planning B: Urban Analytics and City Science* 48(6): 1527–1542. DOI: [10.1177/2399808320938796](https://doi.org/10.1177/2399808320938796).
- Niu H and Silva EA (2021) Delineating urban functional use from points of interest data with neural network embedding: a case study in Greater London. *Computers, Environment and Urban Systems* 88: 101651. DOI: [10.1016/j.compenvurbsys.2021.101651](https://doi.org/10.1016/j.compenvurbsys.2021.101651).
- Schröpfer T (2020) *Dense + Green Cities*. Basel, Switzerland: Birkhäuser. ISBN 978-3035615319.
- Shen J and Wu F (2020) Paving the way to growth: transit-oriented development as a financing instrument for Shanghai's post-suburbanization. *Urban Geography* 41(7): 1010–1032. DOI: [10.1080/02723638.2019.1630209](https://doi.org/10.1080/02723638.2019.1630209).
- Sidhu N, Pebesma E and Câmara G (2018) Using google earth engine to detect land cover change: Singapore as a use case. *European Journal of Remote Sensing* 51(1): 486–500. DOI: [10.1080/22797254.2018.1451782](https://doi.org/10.1080/22797254.2018.1451782).
- Verburg PH, de Nijs TC, Ritsema van Eck J, et al. (2004) A method to analyse neighbourhood characteristics of land use patterns. *Computers, Environment and Urban Systems* 28(6): 667–690. DOI: [10.1016/j.compenvurbsys.2003.07.001](https://doi.org/10.1016/j.compenvurbsys.2003.07.001).
- Wang Z, Ma D, Sun D, et al. (2021) Identification and analysis of urban functional area in Hangzhou based on OSM and POI data. *PLoS One* 16(5): e0251988. DOI: [10.1371/journal.pone.0251988](https://doi.org/10.1371/journal.pone.0251988).
- Wang J, Feng CC and Guo Z (2023) A novel graph-based framework for classifying urban functional zones with multisource data and human mobility patterns. *Remote Sensing* 15(3): 730. DOI: [10.3390/rs15030730](https://doi.org/10.3390/rs15030730).
- Wei S and Wang L (2021) Community detection, road importance assessment, and urban function pattern recognition: a big data approach. *Journal of Spatial Science* 68(1): 23–43. DOI: [10.1080/14498596.2021.1936669](https://doi.org/10.1080/14498596.2021.1936669)
- Yang M, Kong B, Dang R, et al. (2022) Classifying urban functional regions by integrating buildings and points-of-interest using a stacking ensemble method. *International Journal of Applied Earth Observation and Geoinformation* 108: 102753. DOI: [10.1016/j.jag.2022.102753](https://doi.org/10.1016/j.jag.2022.102753).
- Ye C, Zhang F, Mu L, et al. (2021) Urban function recognition by integrating social media and street-level imagery. *Environment and Planning B: Urban Analytics and City Science* 48(6): 1430–1444. DOI: [10.1177/2399808320935467](https://doi.org/10.1177/2399808320935467).
- Yeow LW, Low R, Tan YX, et al. (2021) Point-of-interest (poi) data validation methods: an urban case study. *ISPRS International Journal of Geo-Information* 10(11): 735. DOI: [10.3390/ijgi10110735](https://doi.org/10.3390/ijgi10110735).
- Zhang Y, Li Q, Tu W, et al. (2019) Functional urban land use recognition integrating multisource geospatial data and cross-correlations. *Computers, Environment and Urban Systems* 78: 101374. DOI: [10.1016/j.compenvurbsys.2019.101374](https://doi.org/10.1016/j.compenvurbsys.2019.101374).
- Zhang B, Zhong C, Gao Q, et al. (2022) Delineating urban functional zones using mobile phone data: a case study of cross-boundary integration in Shenzhen-Dongguan-Huizhou area. *Computers, Environment and Urban Systems* 98: 101872. DOI: [10.1016/j.compenvurbsys.2022.101872](https://doi.org/10.1016/j.compenvurbsys.2022.101872).
- Zhong C, Huang X, Müller Arisona S, et al. (2014) Inferring building functions from a probabilistic model using public transportation data. *Computers, Environment and Urban Systems* 48: 124–137. DOI: [10.1016/j.compenvurbsys.2014.07.004](https://doi.org/10.1016/j.compenvurbsys.2014.07.004).
- Živković J (2019) Urban form and function. In: Leal Filho W, Azeiteiro U, Azul AM, et al. (eds) *Climate Action, Encyclopedia of the UN Sustainable Development Goals*. Cham: Springer International Publishing, pp. 862–871. DOI: [10.1007/978-3-319-95885-9_78](https://doi.org/10.1007/978-3-319-95885-9_78).

Wei Chien Benny Chin is a Postdoctoral Researcher at the Singapore University of Technology and Design and National University of Singapore. He is a geographer and a Geographical Information Scientist. His research interests include GIScience, analysis of spatial structures, and complex

network analysis. He is participating in projects associated with complex human movement and urban spatial structures. More information can be found at <https://wcchin.github.io/>.

Yuming Fu is an Associate Professor with the Department of Real Estate, NUS Business School, and the Department of Architecture, College of Design and Engineering, National University of Singapore (NUS). His research focuses on real estate and urban economics. More information is available at <https://bizfaculty.nus.edu.sg/faculty-details/?profId=661>.

Kwan Hui Lim is an Assistant Professor at the Information Systems Technology and Design Pillar, Singapore University of Technology and Design, and leads the Social and Urban Analytics Lab. His research interests are in Data Mining, Machine Learning, Artificial Intelligence, Social Network Analysis, and Social Computing. For more information, please refer to <https://sites.google.com/site/limkwanhui/>.

Thomas Schroepfer is Full Professor of Architecture and Sustainable Design at the Singapore University of Technology and Design where he leads the University's multidisciplinary Cities: Urban Science and Design for Density research thrust. He also serves as Co-Director of the Singapore-ETH Centre Future Cities Laboratory (FCL) Global. For more information, please visit <https://asd.sutd.edu.sg/people/faculty/thomas-schroepfer>.

Lynette Cheah is an Associate Professor of Engineering Systems at the Singapore University of Technology and Design. She directs the Sustainable Urban Mobility research laboratory, which develops data-driven models and tools to reduce the environmental impacts of passenger and urban freight transport.