Journal of Contemporary Urban Affairs

2026, Volume 10, Number 1, pages 153–170

Original scientific paper

Introducing A Novel Compositional Data Analysis Framework for Measuring Land Use Mix in Station Areas

*[1] **Bhagya Kahatagahawatte**, [2] **Ryoji Matsunaka**, [3] **Nobuhiro Uno**, [4] **Tomoki Nishigaki**

[1, 2, 3, &4] Department of Urban Management, Graduate School of Engineering, Kyoto University, Japan

[1] E-mail: yakupitiyage.kahatagahawatte.42v@st.kyoto-u.ac.jp, [2] E-mail: matsunaka.ryoji.3v@kyoto-u.ac.jp, [3] E-mail: uno.nobuhiro.2v@kyoto-u.ac.jp,
[4] E-mail: nishigaki.tomoki.6d@kyoto-u.ac.jp

ARTICLE INFO:

Article History:
Received: 14 March 2026
Revised 1: 28 April 2026
Revised 2: 26 May 2026
Accepted: 16 June 2026
Available online: 25 June 2026

Keywords:
Land-Use Mix;
Isometric Log Ratio Balances;
Compositional Data Analysis;
Shannon Entropy Index;
Station Areas;

ABSTRACT

This study presents isometric log-ratio (ILR) balances derived from compositional data analysis (CoDA) as a novel framework for measuring the land-use mix, addressing the category insensitivity of the Shannon entropy index (SEI). Six ILR balances were constructed using a theory-driven Sequential Binary Partition of seven residential and commercial functional uses in station buffers within the Osaka Densely Inhabited Districts in Japan. The findings demonstrated that ILR balances successfully clustered three functionally distinct station-area typologies that align with urban outcomes. Comparative analysis of ILR balances and the Shannon entropy index revealed an inverted U-shaped ILR–SEI relationship, demonstrating that ILRs can distinguish compositionally distinct buffers with identical entropy scores. ILR balances further elucidated the sub-component-level interactions, which are masked by the single index of Shannon entropy. Overall, the findings suggest that ILR balances effectively complement the Shannon entropy index by providing category-sensitive directional measurement of functional mix and hierarchical sub-compositional decomposition, offering a more complete and planning-relevant characterization of functional land-use composition around transit stations.

JOURNAL OF CONTEMPORARY URBAN AFFAIRS (2026), ***10***(1), 153–170.
https://doi.org/10.25034/ijcua.2026.v10n1-b7

www.ijcua.com

Highlights:	Contribution to the field statement:
- ILR balances are introduced as a novel land use mix measurement framework. - Six ILR balances represent the functional trade-offs in station areas. - K means clustering on ILR balances yielded three statistically validated and spatially coherent station-area typologies. - ILR balances complement Shannon entropy by revealing category-specific functional dominance and sub compositional trade-offs.	The study offers a novel isometric log-ratio (ILR) framework for measuring land use mix, that simultaneously captures functional dominance and sub-compositional balances while generating statistically valid indices for statistical modelling.

* **Corresponding Author:** Bhagya Kahatagahawatte
Department of Urban Management, Graduate School of Engineering, Kyoto University, Japan
Email address: yakupitiyage.kahatagahawatte.42v@st.kyoto-u.ac.jp

How to cite this article? (APA Style)
Kahatagahawatte, B., Matsunaka, R., Uno, N., & Nishigaki, T. (2026). Introducing a novel compositional data analysis framework for measuring land use mix in station areas. *Journal of Contemporary Urban Affairs*, 10(1), 153–170. https://doi.org/10.25034/ijcua.2026.v10n1-b7

1. Introduction

In urban contexts, railway station areas are commonly recognized as economically active areas, characterized by concentrated commercial functions, high accessibility, and dense residential development. However, the high development around station areas generates urban challenges such as rising property prices, reduced housing affordability, and car dependence (Cervero & Kockelman, 1997; Raj et al., 2024). These dynamics underscore the importance of understanding how residential and commercial functions are distributed and balanced within station areas, thereby making land-use mix a central concept in transit-oriented urban planning. Mixed use, as understood through Dovey and Pafka (2017), is the co-presence and functional integration of diverse activities within a defined spatial unit. Mixed-use environments enhance urban vitality and social sustainability, highlighting the importance of mixed-use computation and monitoring in transit-oriented planning practice (Gu et al., 2024; Ibraeva et al., 2020; Yang et al., 2021).

Numerous methods have been proposed to quantify mixed use, with the Shannon entropy index (SEI) being the most widely used due to its computational simplicity. SEI measures the degree of evenness of the distribution of proportions across land-use categories (Cervero & Kockelman, 1997; Mohabey et al., 2023). It provides a single scalar value for a defined spatial unit, with higher values indicating a better mix of land uses and stronger functional diversity. While SEI has been extensively applied in urban land-use research, the literature has identified several limitations and proposed various methods to overcome them (Jiao et al., 2021; Zhuo et al., 2022). For instance, Zhao et al. (2023) proposed a 3-dimensional, multi-aspect mixed-degree index that considers the diversity, accessibility, and compatibility of land uses. Motieyan and Azmoodeh (2021) proposed a mixed-use distribution index that incorporates quantity, importance, distance, and balance in land-use mixture to assess ideal mixedness. Further, Chen and Song (2020) proposed a weighted entropy index by incorporating surrounding land-use attributes and regional land-use context to mitigate the modifiable area unit problem (MAUP) and the equal composition issue.

More critically, SEI's category insensitivity has long been recognized, yet improvements to the gap remain limited. SEI lacks information on which functions are mixed and their contribution to the overall mix. Two areas with identical entropy values may exhibit fundamentally different functional compositions, yet SEI cannot distinguish between them. Pedreira et al. (2026) proposed a Bidirectional Global-centered Balance Index (BGBI) as a signed index of land-use composition to measure directionality. However, their index only distinguishes between residential and non-residential discrimination, without capturing the heterogeneity within each category. In some cases, SEI is employed alongside other metrics to address the gaps in SEI (Akyol et al., 2024; Kashef, 2023). Some other studies employed functional-use proportions to represent the land use profile in regression models (Gu et al., 2024; Wu et al., 2023). However, using a set of functional proportions can induce statistical challenges due to the compositional nature of land-use data. Land-use proportions exhibit a constant-sum constraint, in which the sum of all land-use proportions equals 1, inducing spurious correlations among variables (Filzmoser et al., 2018; Pearson, 1897). Under this constraint, an increase in one land-use category necessarily reduces the others, making it difficult to interpret their independent effects. Therefore, from a planning perspective, it is important to have both an overall mixed-use score and detailed information about underlying mixing patterns.

Based on the above arguments, this study proposes a compositional data analysis (CoDA) based isometric log-ratio (ILR) balances as a possible remedy for the category insensitivity of SEI. CoDA methods have been used in ecology and geology (Grunsky et al., 2024; Kida et al., 2026) yet remain considerably underutilized in urban planning. Compositional data represent parts of a whole that sum to a unit, carrying information about the relative relationships among parts (Egozcue & Pawlowsky-Glahn, 2005; Thomas-Agnan et al., 2021). Compositional data can be transformed into a set of isometric log ratios (ILRs), thereby overcoming the unit-sum constraint. ILR operates within the Aitchison simplex and defines the composition as a set of balances between user-defined category groupings, thereby indicating trade-offs among land-use components (Aitchison, 1982).

Very few previous studies incorporated CoDA methods in land use planning. Thomas-Agnan et al. (2021) applied CoDA to econometric land-use modelling, transforming broader land-use categories into ILR coordinates. Yoshida and Tsutsumi (2018) applied CoDA to land-use planning in the Japanese context, using ILR coordinates in multivariate conditionally autoregressive models to examine how the spatial weight matrix affects prediction accuracy. However, both studies use ILR in statistical models rather than as a standalone indicator of land use mix. Consequently, the potential of ILR balances to characterize complex urban land-use compositions was not thoroughly assessed.

ILR balances offer several merits addressing the above-discussed limitations. First, ILR balances preserve category identity and represent pairwise or group-level trade-offs between specific functional types. Second, the planners are informed which contrasts are substantively meaningful, thereby making analytical choices transparent and replicable rather than implicit. Third, ILR balances are subcompositionally coherent, meaning that results remain stable when categories are added or removed from the analysis.

Based on the identified gap of category insensitivity in the SEI, this study aims to pursue three objectives, as outlined in the research framework in Figure 1. First, to construct a set of theory-driven ILR balances. Second, to derive functionally interpretable station area typologies through ILR-based clustering. Finally, systematically compare ILR balances and SEI across the identified typologies, evaluating the extent to which the two measures capture distinct dimensions of land-use composition. The study hypothesizes that the ILR balances provide category-sensitive and hierarchical sub-compositional information.

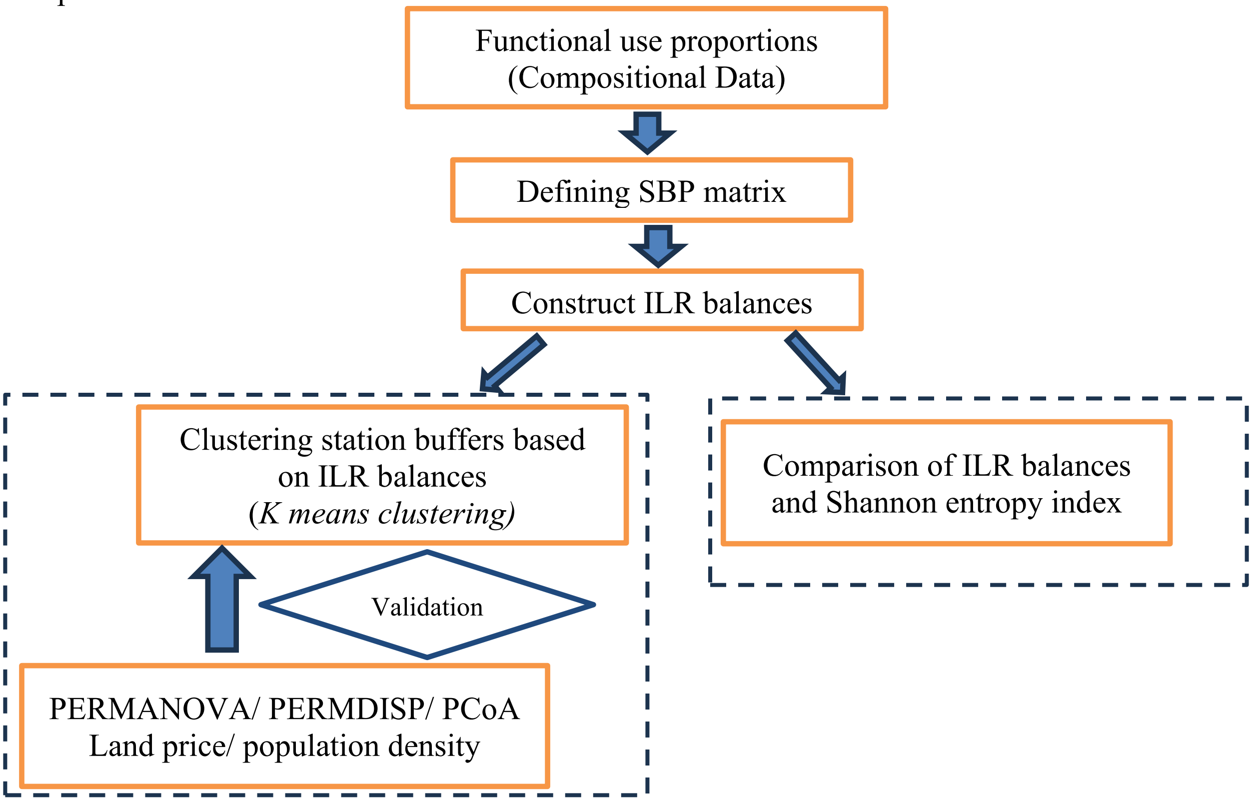


Figure 1. Conceptual research framework.

2. Materials and Methods

2.1 Study area

The Densely Inhabited Districts (DIDs) of Osaka City, Japan, were selected as the study area (Figure 2). DIDs are compact urban zones, characterized by high population density, diverse land uses, and extensive rail networks, as defined by the Statistics Bureau of the Ministry of Internal Affairs and Communications. Focusing on Osaka DID allows the analysis to concentrate on compact, fully urbanized environments where land-use interactions and station-area dynamics are most pronounced.

Osaka is one of Japan's largest cities and the leading economic hub in western Japan. It has a population of 2,754,742 (as of the 2020 national population census) and covers an area of 225.34 km^2. Multiple private and public railway operators serve the area, forming a dense network of local, rapid, and terminal stations.

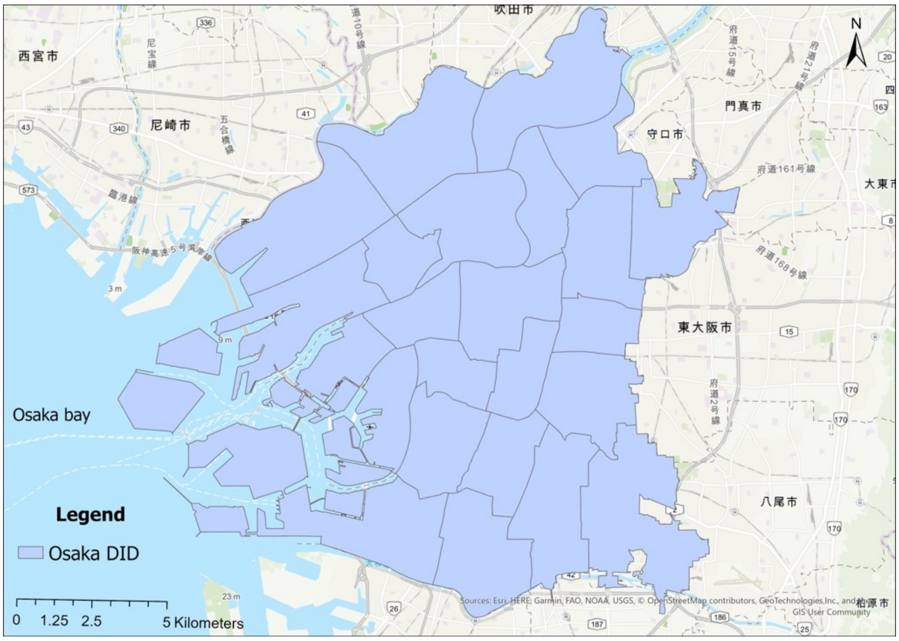


Figure 2. Map of the study area: Osaka DID area including 24 wards.

2.2 Data sources and processing

The study primarily used train network data and functional land-use data. Train network data were obtained from the Ministry of Land, Infrastructure, Transport and Tourism (MLIT) for the base year 2021. Bullet trains were excluded from the study because they operate on specialized or long-distance routes. The tram network was also excluded because it primarily serves short-distance, neighbourhood-scale mobility through street-running operations and closely spaced stops.

Japanese urban areas are categorized into 13 zonings spanning from Category I exclusively low-rise residential zone to Exclusive industrial zone by the City Planning Act. Beyond the strictly residential and industrial zones, other zones permit a broad range of functional uses. Therefore, the stations located in or near exclusively residential or industrial zones were excluded from the study to ensure an unbiased mixed-use analysis.

In Osaka's train network, some stations are shared by multiple railway operators, with facilities located in close proximity and connected to facilitate passenger transfers. Therefore, nodes representing the same location, or nodes within 100 m of each other, were merged to avoid spatial double-counting, which can lead to similarity between observations (Yang et al., 2021). As shown by Gao et al. (2024), station facilities are more dominant within a 100 m radius, so the merging will have a marginal effect on compositional independence.

TOD-related studies commonly adopt a catchment radius of 400 m to 800 m, with 400m widely accepted as a core pedestrian walkability threshold around transit stations (Gao et al., 2024; Ibraeva et al., 2020). In this study, a 400 m radius buffer was selected because the Osaka DID is characterized by an exceptionally dense railway network, with stations located close to one another.

The study uses non-aggregated building-use data for 2021 obtained from the Osaka City Planning Basic Survey. The dataset includes detailed information for each building, overcoming the limitations of mesh or parcel-based datasets and POI data in capturing fine-grained functional structures.

Population data were sourced from the 2020 national population census, and land price data for 2021 were obtained from the Osaka city office.

2.3 Land use categories and functional use types

This study focused exclusively on residential (RS) and commercial (CM) land use categories, and the term “mixed use” refers to the mixing of RS and CM functional uses. Competition for space around station areas occurs primarily between residential and commercial uses because of the location’s high accessibility (Bertolini, 1999; Yang et al., 2022). Therefore, other major land-use categories, such as industrial and public facilities, were excluded to maintain analytical focus on RS-CM functional mixing. Furthermore, because the application of ILR balances in land-use analysis remains relatively novel, limiting the analysis to the principal functional contrast enhances methodological validation and reduces computational complexity.

Seven functional uses belonging to the RS and CM were selected for the analysis. Single family housing, multi family housing, mixed use apartments and housing with stores were selected as 4 residential (RS) functional uses while retail, office and business and food and hospitality functional uses were selected as 3 commercial (CM) functional uses to derive ILR balances. All data processing and analysis were performed using ArcGIS Pro 3.4.0 and Python.

2.4 Isometric Log Ratio (ILR) transformation

Seven functional use proportions were transformed into a set of isometric log-ratios (ILRs) to represent station-area functional composition. ILRs were derived from an orthonormal basis of the Aitchison simplex, transforming the constrained simplex into standard euclidean space (Egozcue et al., 2003). Before the ILR transformation, the zeros in the functional-use composition were handled using the multiplicative replacement method (Martín-Fernández et al., 2003).

First, an orthonormal basis was constructed using a Sequential Binary Partitioning (SBP) of the D functional use proportions, with each basis axis representing a specific functional trade-off. In SBP, the composition is hierarchically divided into two groups, given the + group (numerator) and the − group (denominator), and the functional uses that do not belong to the balances were indicated as zero (Table 1). The SBP developed D-1 non-overlapping binary contrasts (Egozcue & Pawlowsky-Glahn, 2005).

The partition sequence was designed to reflect the theoretical residential–commercial functional hierarchy of station areas. The first balance (ILR1) contrasts all residential (RS) components collectively against all commercial (CM) components, capturing the primary functional trade-off in station areas as grounded in bid rent theory and the node place model (Alonso, 1964; Bertolini, 1999). The second balance (ILR2) decomposed the RS components into low-density and high-density housing because it captures the residential densification (Cervero & Kockelman, 1997). Third (ILR3) and fourth (ILR4) balances, decomposed the low-density housing components and high-density housing components into single-use and mixed-use housing, respectively. ILR3 distinguishes between purely residential uses and coexisting residential and small-scale commercial uses. ILR4 captures vertical residential-commercial functional integration, thereby informing an important aspect of modern planning (Khan et al., 2022). Fifth (ILR5) balance contrasts structured commercial uses with food and hospitality uses because it captures the trade-offs between stable daytime economic activity and social interaction activities beyond commercial core hours (Kretzer et al., 2024). The last balance (ILR6) decomposes the structured commercial use into retail and offices and business uses to capture trade-offs between the local consumer population and an employment hub attracting workers from neighbouring areas (Bertolini, 1999; Chatman & Noland, 2011)

Table 1: Sequential Binary Partition (SBP) and interpretation of Isometric Log-Ratio (ILR) Balances for station area functional uses

Balance	Single family housing	Multi family housing	Mixed use apartment	Housing with stores	Retail	Offices and business	Food and hospitality	Interpretation
ILR1	+	+	+	+	-	-	-	Residential districts (RS) vs. commercial districts (CM)
ILR2	+	-	-	+	0	0	0	Low density housing (LD) vs. high density housing (HD)
ILR3	+	0	0	-	0	0	0	Low density monofunctional housings (LS) vs. low density mixed-use housings (LM)
ILR4	0	+	-	0	0	0	0	High density mono functional housings (HS) vs mixed use buildings (HM)
ILR5	0	0	0	0	+	+	-	Structured commercial uses (SC) vs food and hospitality (FH)
ILR6	0	0	0	0	+	-	0	Shopping districts (RT) vs. office districts (OB)

ILR values were calculated using Equation 1 based on the above orthonormal basis. $ILR_{i,n}$ is the isometric log-ratio for the i^th^ binary partition in the n^th^ buffer. $\prod_{+} x_j$ is the product of the r parts coded positively in the i^th^ binary partition (numerator), while $\prod_{-} x_k$ is the product of the s parts coded negatively (denominator). Here, r is the number of parts in x_{+}, and s is the number of parts in x_{-}. Compared to raw functional-use proportions, ILR balances provide interpretable measures of functional dominance and internal balance, avoid spurious correlations common in compositional data, and help identify context-specific mixed-use patterns around train stations.

$$ILR_{i,n} = \sqrt{\frac{rs}{r+s}} \ln\left(\frac{(\prod_{+} x_j)^{\frac{1}{r}}}{(\prod_{-} x_k)^{\frac{1}{s}}}\right) \quad \text{for } i = 1,2,\dots,D-1 \qquad j = 1,2,\dots,r\text{: } k = 12,\dots,s$$

Equation 1. General Isometric Log-Ratio (ILR) transformation.

ILR can yield both positive and negative values, depending on which part is dominant. When the numerator group dominates, ILR is positive; when the denominator group dominates, it is negative. A zero value indicates a perfect balance between the two groups.

2.5 ILR-based station area clustering and multivariate validation

2.5.1 Station area typology derivation using K-means clustering

Cluster analysis was performed to analyze station typologies using ILR balances and to comparatively evaluate ILR balances and the entropy of functional land-use composition across station-area typologies. Stations were clustered using the K-means clustering method in Python's scikit-learn library. K-means clustering partitions observations into k clusters by minimizing within-cluster variance measured as the Sum of the Squared Errors (SSE) (Equation 2) and maximizing between-cluster differences. Buffers were clustered using 6 ILR balances (ILR1to ILR6) without prior standardization, as ILR balances are dimensionless units on a common scale. Centroid initialization

followed the k-means++ procedure, and the algorithm was run with 10 independent initializations, retaining the solution with the lowest SSE as the final result. A fixed random seed (Random state = 42) was specified to ensure reproducibility. The maximum number of iterations per run was set to 300 (Ketchen & Shook, 1996). The optimal number of clusters was determined using the elbow method followed by the silhouette score. The elbow plot was obtained for k = 2 to 10. The elbow method identified k = 3 as the point of diminishing variance reduction, while the silhouette score reached its maximum at k = 2. However, only 2 clusters would reproduce the primary RS-CM contrast without revealing the finer typological differentiation. Therefore, the next-highest silhouette score (0.321 for k=3) was selected, yielding 3 clusters for the analysis.
Equation 2 calculates the SSE, where k is the number of clusters, x_i is a set of values in cluster i, and C_i is the mean of cluster i.

$$\mathrm{SSE} = \sum_{i=1}^{k} \sum_{x \in C_i} (x_i - C_i)^2$$

Equation 2. Sum of the squared error (SSE)

2.5.2 Multivariate validation of ILR-based cluster separation

Given that ILR-based clustering is methodologically novel, clustering results were statistically validated by PERMANOVA, PERMDISP, and Principal Coordinates Analysis (PCoA). A PERMANOVA test was conducted to assess statistically significant differences in the composition of ILR balance among clusters (Anderson, 2001). A significant result indicates that group centroids are meaningfully separated, evidencing the distinct characteristics among clusters. A Permutational Analysis of Multivariate Dispersion (PERMDISP) was conducted after the PERMANOVA test (Anderson et al., 2006) to confirm differences in clusters, as PERMANOVA can be influenced by differences in within-group dispersion. A non-significant result confirms the genuine cluster separation. Additionally, Principal Coordinates Analysis (PCoA) was performed on the euclidean distance matrix derived from the ILR coordinates, enabling visual interpretation of compositional gradients and cluster separation across station buffers (Legendre, 2018).

2.6 Shannon Entropy Index (SEI)

To assess whether ILR balances provide insights beyond traditional land-use mix metrics, the Shannon Entropy Index (SEI) was computed for each station buffer. The SEI is a common measure of land-use diversity, ranging from 0 to 1, with 0 indicating complete homogeneity and 1 indicating an even distribution across all land-use types (Cervero & Kockelman, 1997). In this analysis, the SEI was calculated using seven functional use categories, consistent with the ILR transformation structure to enable direct comparison. Equation 3 presents the method for calculating the Shannon Entropy Index, where SEI_i denotes the entropy in buffer i, P_j is the proportion of land-use type j, and J is the total number of land-use categories.

$$SEI_i = \sum_{j=1}^{J} \frac{P_j \ln P_j}{\ln J}$$

Equation 3. Calculation of the Shannon entropy index (SEI)

3. Results

3.1 Descriptive statistics of ILR balances

The six ILR balances exhibited different distributions across station buffers. ILR1, ILR5, and ILR6 showed wide ranges, from −5.373 to 5.851, −2.374 to 5.093, and −4.974 to 1.665, respectively, along with large standard deviations. In contrast, ILR2, ILR3, and ILR4 showed narrower ranges (−3.219 to 1.026, −1.267 to 2.405, and −0.822 to 2.633, respectively), indicating less variability and more consistent structural relationships within residential subcategories across the study area. Outliers were found in ILR1, ILR3, ILR4, ILR5, and ILR6, indicating station buffers with highly specialized functional uses. These outliers are not anomalies but rather reflect the near saturation of single-function use within a station buffer.

The mean ILR1 value of 1.305 indicates that residential uses predominate across the network's station areas as a whole. The mean ILR2 of -1.251 supports Osaka's dense structure, indicating a dominance of high-density housing. Similarly, the mean ILR6 was reported as -1.338, indicating that the station areas are dominated by offices and business. However, the means of ILR3, ILR4, and ILR5 were closer to 0, indicating a balanced structure of residential and commercial subcategories. It is evident that Osaka's urban structure has a central tendency toward high-density housing, along with office and business functions.

3.2 ILR-based station area clustering analysis

Three clusters were obtained from K-means clustering (k = 3), named Cluster 0 (C0), Cluster 1 (C1), and Cluster 2 (C2) with 15, 60, and 36 buffers, respectively. Figure 3 shows the mean ILR values for each cluster. C0 is characterized by the strong dominance of commercial uses, as the ILR1 value was −2.610. ILR2 was -2.280, indicating a dominance of high-density housing over low-density housing. ILR2, ILR3, and ILR5 values were closer to zero, indicating a balanced prevalence of single-use and mixed-use housing and structured commercial activities and food and hospitality functions. However, C0 showed a high dominance of offices and business functions relative to retail activities (ILR6 = −2.669), indicating a strong employment-oriented structure.

C1 is characterized by housing-oriented functional uses, with a cluster mean ILR1 of 2.936. ILR2, ILR3, ILR4, and ILR6 values were -0.633, 0.892, 0.920, and -0.967, respectively, indicating a very slight dominance of each group in the balance. Structured commercial uses dominate over food and hospitality functions, with a mean ILR5 of 1.643, indicating a low presence of restaurants or hotels in the residential-oriented cluster.

C2 is characterized by balanced residential–commercial land-use structures, as the cluster mean ILR1 is close to zero (0.324), indicating a relative balance between residential and commercial uses. However, high-density housing clearly dominates low-density housing (ILR2 = −1.811), indicating that these mixed-use areas are primarily composed of compact, high-density developments. Similar to the C0, ILR3 (0.280), ILR4 (0.285) and ILR5 (0.233) showed a balanced single use and mixed use housings as well as structured commercial uses and food and hospitality functions. Offices and business functions were dominant in C2, with an ILR6 value of -1.4.

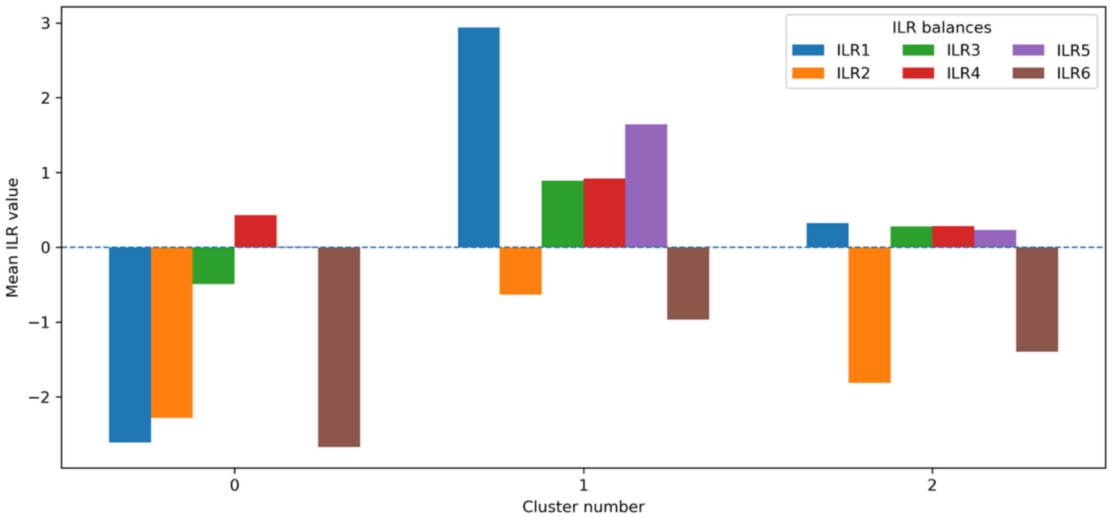


Figure 3. Mean ILR balance profiles of the three clusters.

Across all 3 clusters, single-use and mixed-use housing types (ILR3 and ILR4) remained relatively balanced, indicating that they are consistently balanced regardless of residential density or cluster type.

Mean land prices differed significantly across clusters (Kruskal-Wallis H = 54.86, $p < 0.05$), with C0 recording the highest mean of 1,805,000 yen/m², followed by C2 (389,000 yen/m²) and C1 (162,000 yen/m²). Population density similarly differed significantly across clusters (Kruskal-Wallis H = 27.2, $p < 0.05$), with C1 recording the highest mean of 9426 persons/buffer, followed by C2 with 8872 persons/buffer and C0 with 4222 persons/buffer.

3.2.1 Multivariate statistical validation of cluster separation

The PERMANOVA results indicated a statistically significant difference among clusters (pseudo-F = 72.69, p = 0.001), confirming that the clusters represent distinct compositional structures. The large pseudo-F value indicates that between-cluster variation substantially exceeds within-cluster variation. The PERMDISP result was not statistically significant at the 5% level (F = 3.07, p = 0.053), supporting the assumption of homogeneous dispersion. These results suggest that ILR genuinely clusters the station buffers in the Osaka DID.

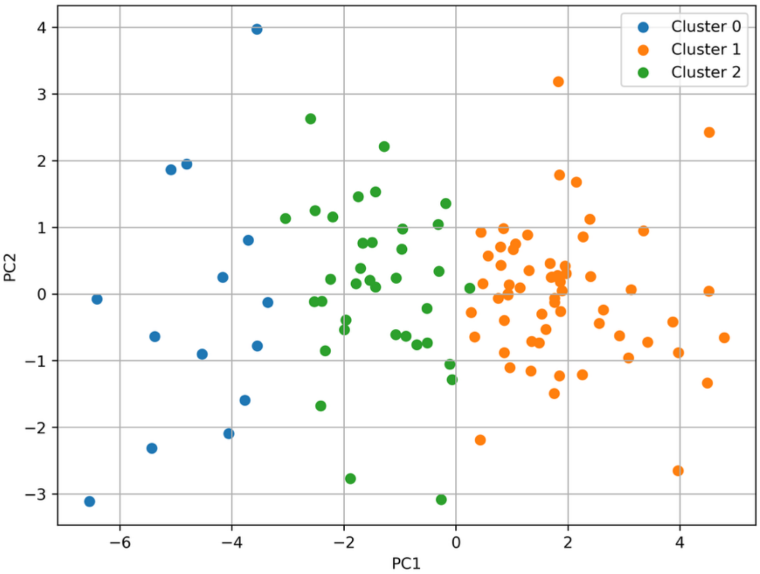


Figure 4. Principal coordinate analysis (PCoA) graph of ILR balances.

The PCoA visualization (Figure 4) further supports this finding by showing clear separation of clusters along the first principal coordinate (PC1). PC1 captured the largest variation in the data, accounting for 67.5%. It was strongly associated with ILR1 (r = 0.98), indicating that buffers were more likely to cluster by residential-commercial balance. The second principal coordinate (PC2) captured additional variation orthogonal to PC1, accounting for 15.2% of the total variation. PC2 was strongly associated with ILR6 (r = −0.650) and ILR5 (r = −0.633), indicating that this axis captures variation in the internal structure of commercial uses.

3.2.2 Spatial distribution of station clusters within the Osaka DID

The spatial distribution of stations within each cluster is shown in Figure 5. Stations in the C0 cluster are concentrated in the central areas of the Osaka DID, particularly around major stations such as Osaka Station and Namba Station. A few stations, including Shin-Osaka Station and Osaka Business Park, lie on the periphery yet remain part of the same cluster. In contrast, C1 cluster stations are mainly located outside the city center and are identified as suburban stations. C2 cluster stations are primarily located between the C0 and C1 clusters, forming a middle-ring zone. Notably, some C2 stations are also found in more peripheral parts of the DID, suggesting a presence of relatively mixed buffers in suburban areas. These results indicate a clear spatial differentiation of clusters across the Osaka train network.

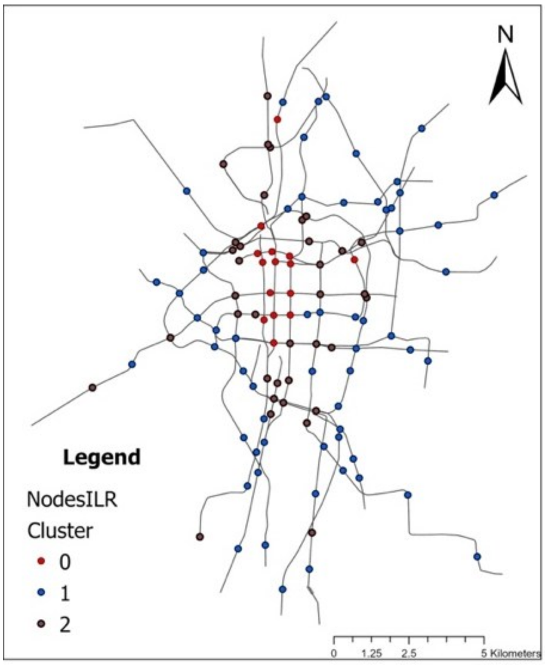


Figure 5. Spatial distribution of station clusters derived from ILR balances across the railway network in the Osaka DID area.

3.3 Comparative analysis of ILR balances and Shannon Entropy Index (SEI)

3.3.1 Shannon Entropy Index (SEI) distribution across ILR clusters

Figure 6 box plots present the SEI for each cluster. The mean SEI value of C0 differed substantially from those of C1 and C2, whereas the distributions of C1 and C2 were more similar, with considerable overlap between their entropy ranges.

C0 exhibited the lowest mean entropy of 0.496 and the highest standard deviation of 0.196. Entropy values ranged from 0.191 to 0.844, indicating substantial internal heterogeneity in functional diversity within the cluster. C1 showed a moderate mean entropy of 0.723 and a low standard deviation of 0.084. The internal entropy distribution was more compact, with fewer extreme values ranging from 0.498 to 0.886, indicating a more consistent degree of mixing across its stations. C2 displayed the highest mean entropy of 0.778 and the lowest standard deviation of 0.08. The SEI values range from 0.612 to 0.927, indicating that stations in this cluster consistently exhibit strong functional diversity.

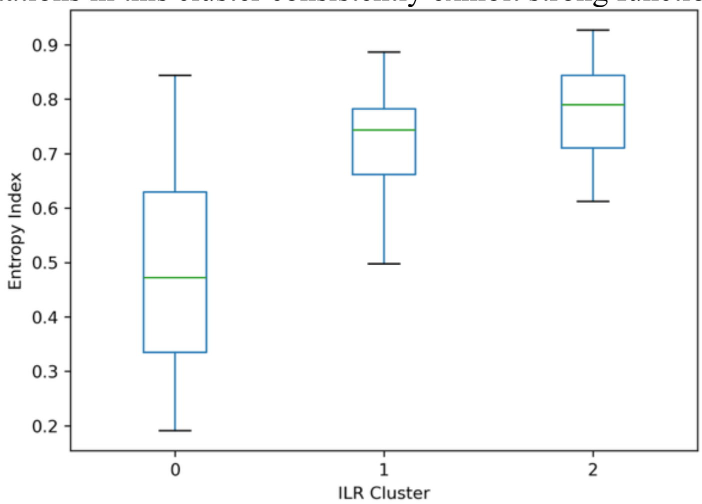


Figure 6. Shannon entropy index distribution in three clusters.

The mean entropy of each cluster increases monotonically from C0 to C1 to C2, describing a gradient from functionally specialized to functionally diverse station areas. A Kruskal–Wallis test confirmed that differences in entropy distributions across the three clusters were statistically significant (χ^2 = 27.72, $p < 0.05$). However, the overlapping entropy ranges across clusters highlight the need for further analysis of SEI and functional composition.

3.3.2 Relationship between Shannon Entropy Index (SEI) and ILR balances

The scatter plots of SEI versus each of the six ILR balances, with colors denoting cluster membership, are shown in Figure 7. Because SEI is a unidimensional index that aggregates all functional uses into a single value, the same entropy value is plotted against each of the six ILR balances, enabling independent evaluation of SEI and each compositional contrast.

Among the six balances, ILR1 exhibited the closest structural alignment with SEI because both measures operate across the full residential-commercial compositional spectrum. The ILR1 vs SEI scatter plot revealed an inverted U-shaped curve, indicating that stations with substantially different compositional structures yield similar entropy scores. ILR1 spanned a wide range from strongly negative (commercial-dominated) to strongly positive (residential-dominated) values, whereas SEI spans nearly its full range of 0 to 1. The three clusters were clearly separated along the ILR axis, with C0 showing the most negative values, C1 showing positive ILR values, and C2 showing values near zero. In contrast, SEI overlapped considerably in C1 and C2, and C0 was clearly separated from C1 and C2 along the entropy axis. SEIs were predominantly concentrated between 0.6 and 0.9 in C1 and C2, whereas the C0 cluster showed a wider spread along the SEI axis from extremely low to high values.

SEI versus ILR2, ILR3, ILR4, and ILR5 showed similar patterns. ILR values were distributed within a relatively narrow range closer to zero, while SEI values were scattered through the full entropy range. C0 showed predominantly low SEI values, whereas C1 and C2 SEI values overlapped and concentrated in the upper SEI ranges.

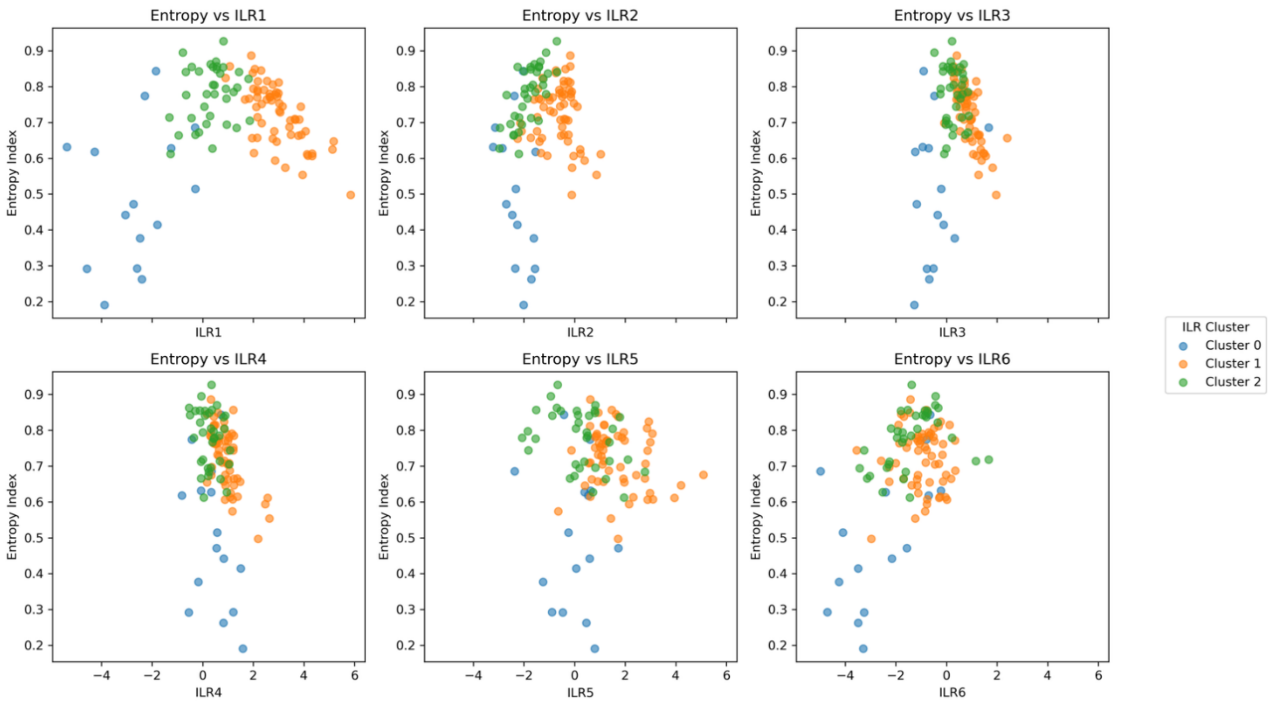


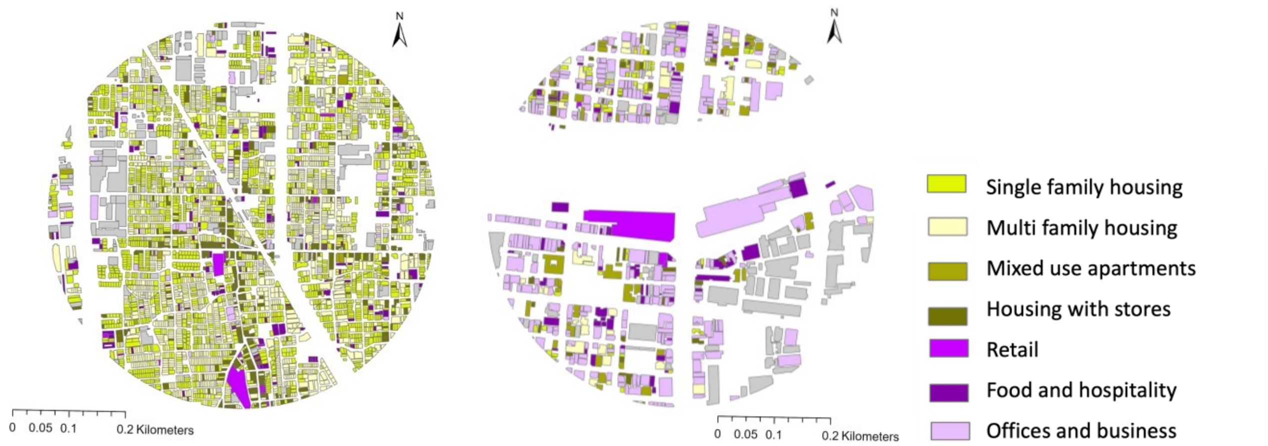
Figure 7. Scatter plots of Shannon Entropy Index (SEI) against each of the six ILR balances, with cluster membership indicated by colour.

Interestingly, ILR6 exhibited a pattern structurally similar to ILR1, with a tendency toward an inverted U-shaped curve along the ILR axis. ILR6 was predominantly in the negative range, closer to zero, indicating dominance of office and business uses, or a balance between office and business uses and retail functional uses. The C0 cluster showed ILR6 values ranging from extremely negative to zero and SEI values ranging from extremely low to moderate. In contrast, the C1 and C2 clusters overlapped on both the ILR and SEI axes, with slightly negative or near-zero ILR values and moderate or high SEI values.

To better understand the relationship between ILR balances and SEI, a pairwise comparison of stations with identical SEI values but different ILR cluster memberships was conducted. Raw functional-use proportions of each buffer were analyzed to discuss the additional input of ILR balances to land use mix studies.

For instance, both the Nishitengachaya and Temmabashi buffers in Osaka DID have a SEI value of 0.61, yet they were assigned to C1 and C2 clusters, respectively. As shown in Figure 8, Nishitengachaya is predominantly residential, with multi-family housing accounting for 43.1% and single-family housing for 41.2%. Commercial uses (including retail, offices and business, and food and hospitality) are minimal, comprising only 5.2% of all considered functional uses. Mixed-use apartments account for just 1.2%, indicating limited mixed-use development despite the area's location within the DID. This distribution was reflected in the ILR balances with ILR1 of 4.124, ILR2 of 1.028, ILR3 of 1.046, and ILR4 of 2.562. These positive values of residential-commercial contrast and the residential sub-component contrasts indicate a specialization toward single-use, low-density housing rather than toward commercial, high-rise, or mixed-use buildings in the Nishitengachaya buffer.

In contrast, the Temmabashi buffer exhibited 63.5% office and business functional uses, 22.7% high-density housing, and 11.5% of mixed-use apartments. ILR1 was −1.267, indicating slight commercial dominance, while ILR2 was −2.206, reflecting a greater presence of high-density housing relative to low-density housing.



Nishi Tengachaya station buffer
SEI = 0.61, ILR cluster = C1

Temmabashi station buffer
SEI = 0.61, ILR cluster = C2

Figure 8. Pairwise comparison of station buffers with identical SEI values but different ILR cluster memberships (Nishi Tengachaya buffer and Temmabashi buffer).

4. Discussion

4.1 Interpretation of ILR Balances as a representation of functional use mixing

ILR transformation was originally developed to ensure statistical robustness in Euclidean-based methods. However, this study applies ILR balances in a deliberately interpretive role, as a novel framework for expressing the mixing of functional uses in station areas. In this role, each ILR balance is not merely a mathematical coordinate but a meaningful functional contrast, with a quantitative measure that quantifies differences in functional use.

Six ILR balances capture independent contrasts for each functional group and explain the station area's functional character in 6 different ways. Most station buffers consist of residential rather than commercial uses, but extreme commercial buffers were identified as those in the urban core with a high concentration of offices and other business uses. However, the narrower distributional range and negative mean of ILR2 indicate that most station areas tend toward high-density housing, regardless of the dominance of broader categories. This finding reflects the higher presence of compact apartments and mixed-use housing within station catchments as identified in the literature (Gao et al., 2024). The near-zero mean values of ILR3 and ILR4, along with their narrower distributional ranges, reveal that single-use and mixed-use housing are similarly present in station buffers regardless of residential density, indicating that mixed-use housing is not confined to high-density compact environments but is consistently present across both low-density and high-density station areas.

4.2 Derivation of station area typologies from ILR balances

The ILR-based clustering aligned with a clear functional typology, distinguishing three station-area types within the Osaka DID. Cluster 0, characterized by commercial predominance, was concentrated in central locations with high accessibility. Cluster 1, with residential dominance, was distributed across peripheral zones. Cluster 2, defined by mixed- functional uses, occupied an intermediate spatial ring. These spatial patterns align well with the bid-rent theory prediction of declining commercial intensity with increasing distance from the most accessible transit nodes (Alonso, 1964).

C0 suggested a commercially intensive employment hub with limited residential presence. It is aligned with Bertolini (1999) node-place model, where highly accessible areas tend to dominate commercial functions over residential uses. The specialization of office and business functions aligns with Kumakoshi et al. (2021) observation in the Tokyo area. The low presence of retail in C0 reflects land-price competition, in which retail activities generate lower revenue per unit area than office uses and are therefore outbid for central locations, as explained by bid-rent theory (Alonso, 1964). Relatively low population density and very high land prices in C0 suggested intense competition for land, indicating pressure on housing affordability (Bardaka, 2024).

C1, with the highest population density and the lowest land price, suggested a residential environment in peripheral station areas in Osaka. High-density housing was more prevalent than low-density housing, reflecting residential densification near rail infrastructure (Kim and Li, 2021). However, the dominance of single-use housing and the non-dominance of food and hospitality uses suggested that C1 is primarily sustained by limited neighbourhood services rather than consumer-related services.

C2, with approximately balanced residential and commercial uses, was closely aligned with TOD principles (Cervero & Kockelman, 1997). Negative ILR2 and ILR6 values in C2 indicated that this balance is built on compact, high-density housing and a noticeable presence of office activity. Furthermore, rather than relying solely on the ILR1 balance, the sub-residential and sub-commercial functional uses also demonstrate the cluster's mixed-use character, highlighting the informative nature of ILR balances.

4.3 Interpretation of SEI distribution in relation to ILR cluster profiles

The statistically significant variation in SEI across the three ILR-derived clusters confirmed that the ILR balances correspond to meaningfully distinct levels of categorical land use diversity. C0's low mean entropy (0.496) aligned with the cluster's strong specialization of commercial functions, which SEI identified as low diversity. Similarly, C2's high mean entropy (0.778) aligned with the cluster's balanced distribution of functional use categories. However, the wider range of SEI values (0.191 to 0.844) in the C0 reflected the prevalence of mixed-use buffers in the commercially specialized cluster, suggesting that entropy masks the real functional diversity due to the equal representation of commercial subtypes. ILR balances, by contrast, consistently identify all C0 stations as commercially dominant regardless of commercial subtypes. More critically, the substantial overlap between C1 and C2 entropy distributions revealed a categorical insensitivity of SEI, as demonstrated by Im & Choi (2019), in which two station areas scored the same SEI while differing in functional composition. For instance, the pairwise comparison between Nishitengachaya and Temmabashi, both with an SEI value of 0.61, provided a compelling demonstration of this limitation. Despite identical entropy scores, these two station areas represented structurally opposite land use environments.

The scatter plot analysis demonstrated that SEI and ILR balances capture fundamentally different yet complementary dimensions of functional land-use composition. While SEI aggregates all functional categories into a single evenness score, ILR decomposes the composition into directional contrasts, revealing the specific functional character that entropy conceals (Zhuo et al., 2022). Due to the inverted U-shaped relationship observed between ILR1 and SEI, a strongly commercial-dominated station and a strongly residential-dominated station can produce identical or near-identical entropy scores due to category insensitivity, as discussed by Jiao et al. (2021). ILR1, by contrast, preserves this directional information, clearly separating commercial-dominated stations (negative values) from residential-dominated ones (positive values). The clear cluster separation along the ILR axis, but the substantial C1 and C2 overlap along the SEI axis, demonstrated that entropy-based indices are effective at identifying extreme specialization, but lose discriminatory power precisely in the intermediate compositional range.

ILR2 through ILR5 patterns indicated residential and commercial sub-compositional contrasts, demonstrating a balanced nature, while SEI was scattered across the full range. This indicates that relationships between subcomponents in a functional use mixture cannot be inferred from an entropy score alone. The ILR framework's hierarchical SBP decomposition captures these sub-compositional distinctions, which are inaccessible to any unidimensional index (Egozcue & Pawlowsky-Glahn, 2005).

These findings support the argument that the Shannon entropy index (SEI) provides a useful but incomplete characterization of functional land-use composition, and that ILR balances offer more information about mixing in station areas, thereby complementing the SEI. Table 2 compares the cluster interpretation based on SEI and ILR balances.

Table 2: Comparison of the interpretation of ILR balances and SEI

Cluster	ILR interpretation	SEI interpretation	Combined interpretation of ILR and SEI
C0	Commercially dominated with high presence of offices and bussines uses	Low diversity	Commercially specialized hub with high proportion of offices and bussines uses
C1	Residential dominated with less food and hospitality activities	High diversity	Residential dominated hub with similar proportions of sub residential categories
C2	Residential commercial balanced cluster with high high-density housing	High diversity	Mixed use buffers with broader functional range.

5. Conclusions

This study introduced isometric log-ratio (ILR) balances as a novel, methodologically sound framework for analyzing land-use composition around transit stations, applying compositional data analysis methods. Applied to station buffers in the Osaka DID, ILR-based clustering identified three distinct clusters with varying population density and land prices, confirming that ILR balances are not merely methodological artifacts but rather reflect real structural differences in urban economic geography. The ILR and SEI comparison demonstrates that the entropy values often overlapped across clusters and that ILR is beneficial for distinguishing the dominant functional use category and the sub-compositional balances. Therefore, the sign-magnitude interpretation of each ILR balance has direct implications for TOD as a planning decision-support tool. The category sensitivity encoded in ILR1 enables planners to classify station areas both broadly and closely monitor the sub-compositional patterns. Further, the full set of ILR balances enables sub-category-specific policy design that focuses on micro-functional uses. Most importantly, ILR balances are recommended over raw proportions and entropy for use in statistical models because they reduce spurious correlations and provide a more detailed analysis than a single scalar index. Kahatagahawatte et al. (2025) analyzed relationships between centrality and land-use composition using GAM models, with ILR balances as response variables.

ILR-based analysis is better regarded as a complement to, rather than a replacement for, the entropy index. Entropy provides a useful first-order summary of overall functional evenness and remains computationally efficient for large-scale studies (Zhuo et al., 2022). In contrast, ILR decomposes compositional structure into interpretable orthogonal contrasts, thereby identifying several functional trade-offs and providing more information about the functional mixing.

Several limitations are acknowledged. First, the analysis was restricted to residential and commercial uses, restricting a full-scale analysis. Future research should investigate how including a broader range of land-use types affects the ILR framework and interpretability of ILR-derived typologies. Second, the functional balances constructed using the Sequential Binary Partition (SBP) matrix represented only one possible compositional decomposition scheme. An alternative SBP matrix can yield different ILR balances. Third, the analysis was limited to the Osaka DID, which constrains generalizability and transferability to other cities with different urban morphologies or transit systems.

Despite these limitations, this study demonstrates the complementarity of ILR balances for the Shannon entropy index and provides a methodologically rigorous, planning-interpretable, and statistically deployable framework for mixed-use analysis in station areas.

Acknowledgements

We acknowledge the Urban Planning Department at Osaka City Hall, Japan, for providing land use data.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Conflicts of Interest

The authors report no conflicts of interest.

Data Availability Statement

The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

Institutional Review Board Statement

Not applicable.

CRediT Author Statement

Bhagya Kahatagahawatte: Conceptualization, Methodology, Writing – original draft, Writing – review & editing, Visualization, Data curation. Ryoji Matsunaka: Methodology, Writing – review & editing, Supervision. Nobuhiro Uno: Methodology, Writing – review & editing, Supervision. Tomoki Nishigaki: Writing – review & editing, supervision, Data curation. All authors have read and approved the final version of the manuscript.

References

Aitchison, J. (1982). The statistical analysis of compositional data. *Journal of the Royal Statistical Society: Series B (Methodological), 44*(2), 139–160. https://doi.org/10.1111/j.2517-6161.1982.tb01195.x

Akyol, M. A., Duzgun, S., & Baykal, N. (2024). landusemix: A Python package for calculating land use mix. *SoftwareX, 27*, Article 101861. https://doi.org/10.1016/j.softx.2024.101861

Alonso, W. (1964). *Location and land use: Toward a general theory of land rent*. Harvard University Press. https://doi.org/10.4159/harvard.9780674730854

Anderson, M. J. (2001). A new method for non-parametric multivariate analysis of variance. *Austral Ecology, 26*(1), 32–46. https://doi.org/10.1111/j.1442-9993.2001.01070.pp.x

Anderson, M. J., Ellingsen, K. E., & McArdle, B. H. (2006). Multivariate dispersion as a measure of beta diversity. *Ecology Letters, 9*(6), 683–693. https://doi.org/10.1111/j.1461-0248.2006.00926.x

Bardaka, E. (2024). Transit-induced gentrification and displacement: Future directions in research and practice. *Transport Reviews, 44*(3), 567–571. https://doi.org/10.1080/01441647.2023.2282285

Bertolini, L. (1999). Spatial development patterns and public transport: The application of an analytical model in the Netherlands. *Planning Practice & Research, 14*(2), 199–210. https://doi.org/10.1080/02697459915724

Cervero, R., & Kockelman, K. (1997). Travel demand and the 3Ds: Density, diversity, and design. *Transportation Research Part D: Transport and Environment, 2*(3), 199–219. https://doi.org/10.1016/S1361-9209(97)00009-6

Chatman, D. G., & Noland, R. B. (2011). Do public transport improvements increase agglomeration economies? A review of literature and an agenda for research. *Transport Reviews, 31*(6), 725–742. https://doi.org/10.1080/01441647.2011.587908

Chen, Y., & Song, Y. (2021). From measuring land use mix to measuring land use pattern: New methods for assessing land use. *arXiv*. https://doi.org/10.48550/arXiv.2105.10383

Dovey, K., & Pafka, E. (2017). What is functional mix? An assemblage approach. *Planning Theory & Practice, 18*(2), 249–267. https://doi.org/10.1080/14649357.2017.1281996

Egozcue, J. J., & Pawlowsky-Glahn, V. (2005). Groups of parts and their balances in compositional data analysis. *Mathematical Geology, 37*(7), 795–828. https://doi.org/10.1007/s11004-005-7381-9

Egozcue, J. J., Pawlowsky-Glahn, V., Mateu-Figueras, G., & Barceló-Vidal, C. (2003). Isometric logratio transformations for compositional data analysis. *Mathematical Geology, 35*(3), 279–300. https://doi.org/10.1023/A:1023818214614

Filzmoser, P., Hron, K., & Templ, M. (2018). *Applied compositional data analysis: With worked examples in R*. Springer International Publishing. https://doi.org/10.1007/978-3-319-96422-5

Gao, Y., Cui, X., & Sun, X. (2024). Land use characteristics of commuter rail station areas and their impact on station ridership: A case study of Japan Railways in the Tokyo Metropolitan Area. *Land, 13*(12), Article 2045. https://doi.org/10.3390/land13122045

Grunsky, E., Greenacre, M., & Kjarsgaard, B. (2024). GeoCoDA: Recognizing and validating structural processes in geochemical data: A workflow on compositional data analysis in lithogeochemistry. *Applied Computing and Geosciences, 22*, Article 100149. https://doi.org/10.1016/j.acags.2023.100149

Gu, X., Lin, S., & Wang, C. (2024). Integrated impact of urban mixed land use on TOD ridership: A multi-radius comparative analysis. *Journal of Transport and Land Use, 17*(1), 457–481. https://doi.org/10.5198/jtlu.2024.2462

Ibraeva, A., Correia, G. H. de A., Silva, C., & Antunes, A. P. (2020). Transit-oriented development: A review of research achievements and challenges. *Transportation Research Part A: Policy and Practice, 132*, 110–130. https://doi.org/10.1016/j.tra.2019.10.018

Im, H. N., & Choi, C. G. (2019). The hidden side of the entropy-based land-use mix index: Clarifying the relationship between pedestrian volume and land-use mix. *Urban Studies, 56*(9), 1865–1881. https://doi.org/10.1177/0042098018763319

Jiao, J., Rollo, J., & Fu, B. (2021). The hidden characteristics of land-use mix indices: An overview and validity analysis based on the land use in Melbourne, Australia. *Sustainability, 13*(4), Article 1898. https://doi.org/10.3390/su13041898

Kahatagahawatte, B., Matsunaka, R., Uno, N., & Nishigaki, T. (2025). A CoDA-based analysis of the influence of train network centrality on urban functional use composition in station buffers. In

Proceedings of the 29th International Conference of the Hong Kong Society for Transportation Studies (HKSTS), Paper 219.

Kashef, M. (2023). Mixed-use and street network attributes of vibrant urban settings. *Architecture and Urban Planning, 19*(1), 188–199. https://doi.org/10.2478/aup-2023-0017

Ketchen, D. J., & Shook, C. L. (1996). The application of cluster analysis in strategic management research: An analysis and critique. *Strategic Management Journal, 17*(6), 441–458. https://doi.org/10.1002/(SICI)1097-0266(199606)17:6%3C441::AID-SMJ819%3E3.0.CO;2-G

Khan, F. M., Pafka, E., & Dovey, K. (2022). Extremes of mixed-use architecture: A spatial analysis of vertical functional mix in Dhaka. *City, Territory and Architecture, 9*(1), Article 31. https://doi.org/10.1186/s40410-022-00177-y

Kida, M., Merder, J., Dittmar, T., Pawlowsky-Glahn, V., & Egozcue, J. J. (2026). Reframing natural organic matter research through compositional data analysis. *Frontiers in Environmental Science, 14*, Article 1724142. https://doi.org/10.3389/fenvs.2026.1724142

Kim, J. H., & Li, X. (2021). Building more housing near transit: A spatial analysis of residential densification dynamics. *Transport Policy, 114*, 15–24. https://doi.org/10.1016/j.tranpol.2021.08.018

Kretzer, G., Kanashiro, M., & Tibiriçá, R. (2023). Complementarity between urban land uses: A temporal analysis. *Journal of Urban Design, 29*(4), 468–485. https://doi.org/10.1080/13574809.2023.2280272

Kumakoshi, Y., Koizumi, H., & Yoshimura, Y. (2021). Diversity and density of urban functions in station areas. *Computers, Environment and Urban Systems, 89*, Article 101679. https://doi.org/10.1016/j.compenvurbsys.2021.101679

Legendre, P. (2018). Numerical ecology. In *Reference Module in Earth Systems and Environmental Sciences*. Elsevier. https://doi.org/10.1016/B978-0-12-409548-9.10595-0

Martín-Fernández, J. A., Vidal, C. B., & Pawlowsky-Glahn, V. (2003). Dealing with zeros and missing values in compositional data sets using nonparametric imputation. *Mathematical Geology, 35*(3), 253–278. https://doi.org/10.1023/A:1023866030544

Mohabey, D. P., Nongkynrih, J. M., & Kumar, U. (2023). Urban growth trend analysis of proposed Greater Silchar City, India, using landscape metrics and Shannon entropy model. *Environment, Development and Sustainability, 26*(10), 25335–25366. https://doi.org/10.1007/s10668-023-03681-4

Motieyan, H., & Azmoodeh, M. (2021). Mixed-use distribution index: A novel bilevel measure to address urban land-use mix pattern: A case study in Tehran, Iran. *Land Use Policy, 109*, Article 105724. https://doi.org/10.1016/j.landusepol.2021.105724

Pawlowsky-Glahn, V., & Buccianti, A. (Eds.). (2011). *Compositional data analysis: Theory and applications*. Wiley. https://doi.org/10.1002/9781119976462

Pearson, K. (1897). Mathematical contributions to the theory of evolution: On a form of spurious correlation which may arise when indices are used in the measurement of organs. *Proceedings of the Royal Society of London, 60*(359–367), 489–498. https://doi.org/10.1098/rspl.1896.0076

Pedreira, J. U., Louro, T. V., Assis, L. B. M. D., Brito, P. L., & Bomfim, F. G. (2026). BGBI: A citywide-referenced and bidirectional land use mix index for planning and policy evaluation. *Land Use Policy, 169*, Article 108135. https://doi.org/10.1016/j.landusepol.2026.108135

Raj, S. I., Mohino, I., & Khatun, F. (2024). Socio-economic impact of the Mass Rapid Transit (MRT) system in Dhaka: A case study of Mirpur stations. *Journal of Contemporary Urban Affairs, 8*(2), 436–459. https://doi.org/10.25034/ijcua.2024.v8n2-9

Thomas-Agnan, C., Laurent, T., Ruiz-Gazen, A., Nguyen, T. H. A., Chakir, R., & Lungarska, A. (2021). Spatial simultaneous autoregressive models for compositional data: Application to land use. In P. Filzmoser, K. Hron, J. A. Martín-Fernández, & J. Palarea-Albaladejo (Eds.), *Advances in compositional data analysis* (pp. 225–249). Springer. https://doi.org/10.1007/978-3-030-71175-7_12

Wu, J., Jia, P., Feng, T., Li, H., & Kuang, H. (2023). Spatiotemporal analysis of built environment restrained traffic carbon emissions and policy implications. *Transportation Research Part D: Transport and Environment, 121*, Article 103839. https://doi.org/10.1016/j.trd.2023.103839

Yang, H., Li, X., Li, C., Huo, J., & Liu, Y. (2021). How do different treatments of catchment area affect the station level demand modeling of urban rail transit? *Journal of Advanced Transportation, 2021*, Article 2763304. https://doi.org/10.1155/2021/2763304

Yang, Y., Zhong, C., & Gao, Q.-L. (2022). An extended node-place model for comparative studies of transit-oriented development. *Transportation Research Part D: Transport and Environment, 113*, Article 103514. https://doi.org/10.1016/j.trd.2022.103514

Yoshida, T., & Tsutsumi, M. (2018). On the effects of spatial relationships in spatial compositional multivariate models. *Letters in Spatial and Resource Sciences, 11*(1), 57–70. https://doi.org/10.1007/s12076-017-0199-5

Zhao, X., Xia, N., & Li, M. (2023). 3-D multi-aspect mix degree index: A method for measuring land use mix at street block level. *Computers, Environment and Urban Systems, 104*, Article 102005. https://doi.org/10.1016/j.compenvurbsys.2023.102005

Zhuo, Y., Jing, X., Wang, X., Li, G., Xu, Z., Chen, Y., & Wang, X. (2022). The rise and fall of land use mix: Review and prospects. *Land, 11*(12), Article 2198. https://doi.org/10.3390/land11122198

Check for Updates

How to cite this article? (APA Style)

Kahatagahawatte, B., Matsunaka, R., Uno, N., & Nishigaki, T. (2026). Introducing a novel compositional data analysis framework for measuring land use mix in station areas. *Journal of Contemporary Urban Affairs*, 10(1), 153–170. https://doi.org/10.25034/ijcua.2026.v10n1-b7