Scientific African 28 (2025) e02662

Contents lists available at ScienceDirect

Scientific African

journal homepage: www.elsevier.com/locate/sciaf

Clustering and spatial dynamics of informal trading in Zimbabwe during the COVID-19 pandemic

Meshach Ndlovu

Gwanda State University, Faculty of Computational Sciences, Department of Mathematics and Statistics

A R T I C L E I N F O

Editor: DR B Gyampoh

Keywords:
Spatial dynamics
Informal financial market
Geostatistics
Persistence diagram
Topological data analysis

A B S T R A C T

The COVID-19 pandemic led to significant disruptions across various sectors, with Zimbabwe's informal economy being acutely impacted. In response, the Zimbabwean government implemented measures to mitigate the economic challenges faced by informal traders, including the establishment of new trading spaces such as Nkulumane Sekusile in Bulawayo. This study investigates the spatial dynamics, understood as the clustering and grouping of different types of traders within Sekusile market during the pandemic. The research examines how the influx of traders from various regions altered the market's structure and competition. By employing advanced statistical clustering methods—such as hotspot analysis, the K-means algorithm, ordinary kriging, and persistence diagrams—this study provides critical insights into these spatial variations, informing policy recommendations aimed at enhancing the socio-economic resilience of informal traders.

Introduction

Informal trading plays a critical role in Zimbabwe's economy, providing livelihoods for a significant portion of the population. With limited formal employment opportunities, many individuals engage in informal trade, occupying public spaces and markets to sell various goods and services. According to the 2023 ZimStat Second Quarter report, informal sector employment (non-agriculture) stood at 1298,828 in the first quarter and increased to 1375,493 in the second quarter, while formal sector employment (non-agriculture) rose from 912,036 to 931,507 in the same period. The labor force participation rates by sex for the two main provinces in Zimbabwe indicate that Harare recorded 73.3 % male participation and 50.0 % female participation, with an overall rate of 61.0 %. In Bulawayo, the labor force participation rates were 72.8 % for males and 58.8 % for females, with a total of 64.9 %. The employment distribution in the informal sector (non-agriculture) by industry shows that wholesale and retail trade, sale, and repair of motor vehicles and motorcycles contributed approximately 49.8 %, making it the dominant industry in informal employment [1].

Sekusile Market in Bulawayo serves as a crucial focal point for studying these spatial dynamics. As one of the largest and most active informal trading hubs in the region, it accommodates a diverse range of traders engaged in the sale of clothing, food items, fresh farm produce, and financial services such as money exchange. Its significance extends beyond Bulawayo, as it plays a pivotal role in the broader informal economy of Zimbabwe, facilitating regional trade and contributing to employment generation. For the Zimbabwean government, markets like Sekusile represent both a challenge and an opportunity: while they provide vital economic sustenance for thousands, they also highlight the need for effective urban planning and regulatory measures to support sustainable market operations. During the peak of the COVID-19 pandemic [2], stringent lockdown measures disrupted the conventional functioning of the market,

E-mail address: meshach.ndlovu@gsu.ac.zw.

https://doi.org/10.1016/j.sciaf.2025.e02662
Received 27 October 2024; Received in revised form 18 February 2025; Accepted 21 March 2025
Available online 22 March 2025

leading traders to adapt by forming clusters that operated without formal oversight. This unregulated environment often resulted in conflicts over limited space, intensified competition, and evolving trade dynamics. By analyzing Sekusile Market, this study offers insights into the resilience of informal traders, the economic impact of spatial clustering, and potential policy interventions to enhance the sustainability of informal trading spaces.

The study seeks to understand the spatial distribution and organization of informal traders in Zimbabwe, particularly in response to economic disruptions such as the COVID-19 pandemic, and to propose data-driven policy recommendations for enhancing market sustainability and urban planning strategies. To achieve this, the research objective is to analyze and model the spatial patterns of informal trading using a combination of geostatistical techniques and topological data analysis (TDA), comparing their effectiveness in capturing market dynamics and informing policy decisions. A key research question guiding this study is: How can geostatistical and topological data analysis methods be utilized to reveal meaningful patterns in the spatial distribution of informal traders and inform strategies for market resilience and equitable urban planning? To justify the methodological approach, a comparative analysis of alternative spatial analysis techniques is conducted. Traditional geostatistical methods such as hotspot analysis and K-means clustering effectively identify density variations and segment data into clusters but often rely on predefined parameters and fail to capture higher-order spatial relationships. In contrast, TDA, particularly through persistence diagrams, provides a more flexible and robust framework for identifying persistent topological features across multiple spatial scales, offering deeper insights into market organization and resilience. By integrating TDA with geostatistical methods, this study presents a more comprehensive approach to understanding informal market dynamics and informing sustainable urban planning initiatives.

Understanding informal trading patterns from existing studies

This section reviews literature to enhance our understanding of how informal trading patterns form and evolve. It examines key factors such as market dynamics, economic conditions, and external disruptions like the COVID-19 pandemic, providing a foundation for the analytical approach in this study.

The influx of new traders from various suburbs and cities disrupted the established trading patterns of local vendors, some of whom had operated at Sekusile for over two decades. The heightened competition led to price volatility, with smaller traders struggling to remain viable. The spatial clustering of traders became a key feature of the market, as groups of traders adapted to the changing dynamics by forming clusters based on product types, pricing strategies, and consumer targets. The study by Kabonga, Zvokuomba and Nyagadza highlighted the unique struggles faced by young informal traders, a group often overlooked by governing authorities at both the local and national levels. Many young traders contend with limited access to capital, harassment from municipal authorities, and intense competition from established vendors. Despite these barriers, they employ various survival strategies, including securing financial support from friends and family. By examining these multifaceted challenges and adaptive strategies, this study extends the understanding of how informal traders navigate complex economic and regulatory landscapes, ultimately contributing to broader discussions on market sustainability and urban planning [3]

To analyze these spatial dynamics, the study employs a suite of statistical clustering methodologies, including hotspot analysis, the K-means algorithm, the ordinary kriging method, and persistence diagrams. Persistence diagrams, in particular, provide a robust visual tool for tracking the birth and death of topological features within spatial data, enabling a detailed analysis of the connectivity and persistence of trader groups [4–6]. The insights generated from this analysis aim to inform the development of policies that enhance the socio-economic interactions and resilience of informal traders. Using a dynamic stochastic general equilibrium model, Ngalawa and Viegi investigated the relationship between informal and formal markets and uncovered some interesting results, such as the existence of interest rates that differ between informal and formal markets and an increasing reliance on formal financial loans [7]. An informal trader is a person who participates in the financial market legally but in an ungoverned approach. This includes people who sell vegetables, groceries, clothes, and change money. Johnson's study highlights the importance of informal groups in Kenya's economic dynamics, particularly the reasons why different traders choose to be part of informal trading groups. The study also examined the organizational features of these informal groups, including record-keeping practices, certificates of registration (vendor licenses), bank accounts or money boxes, and audit strategies. Moreover, experience in informal market trading appears to be a critical skill that can determine the success or failure of a trader in such a market [8].

Spatial dynamics, particularly in the context of financial markets is the distribution and clustering of economic activities across different geographic locations. These dynamics are influenced by various factors, including market forces, regulatory environments, and socio-economic conditions. In informal markets, spatial dynamics often emerge as a result of traders clustering in specific areas to maximize access to resources, customers, and opportunities. This clustering can lead to the formation of distinct economic zones within a market, each characterized by specific types of trading activities. Understanding these spatial dynamics is crucial for developing policies that support the sustainable development of informal markets [9–11]. The spatial competition among trader groups is a critical aspect of informal markets. As traders cluster in specific areas, competition for limited space and consumer attention intensifies. This competition often leads to the formation of strategic alliances among traders, where groups with similar products or services collaborate to enhance their market position. However, it can also result in conflicts, particularly when new entrants disrupt established trading patterns. The spatial arrangement of traders within a market is thus a dynamic process, influenced by both internal and external factors. This study aims to capture these dynamics by analyzing how traders' groups form, compete, and evolve over time [12].

Several studies have explored the spatial characteristics of informal markets, focusing on how traders organize themselves within specific geographic locations [9,11,13,14]. For instance, research on informal markets in sub-Saharan Africa has highlighted the role of spatial clustering in shaping market dynamics. These studies have shown that traders often cluster in areas that offer strategic advantages, such as high foot traffic, proximity to suppliers, or access to specific consumer demographics. Spatial analysis techniques, including clustering algorithms and geospatial mapping, have been used to identify and analyze these patterns, providing insights into the underlying forces that drive market organization [12,15–17]. Karanagoda emphasizes that the unequal distribution of resources drives communities to innovate, leading to the development of informal markets, policies, and infrastructure. Furthermore, the issues associated with informal markets and clustering tend to be confined to specific geographical areas and groups of people over time [18]. The paper referenced in [19] explores the connection between qualitative characteristics of useful financial information using a content analysis method. These fundamental characteristics include trust, responsibility, accountability, public trust, humility, equity, diligence, and discipline. These traits are crucial for the success and dynamics of any trader operating within an informal market. Derrida's deconstruction theory can be employed to qualitatively analyze these characteristics. Additionally, the TDA approach can be used to explore the qualitative features of traders and informal market datasets.

Spatial analysis of informal markets extends beyond mapping vendor locations; it also examines how traders respond to external shocks such as economic downturns, policy interventions, and regulatory changes. The COVID-19 pandemic, for instance, significantly disrupted informal markets, leading to shifts in trader distribution as vendors adapted to restrictions and changing consumer behavior. Some markets fragmented due to social distancing measures, while others saw increased clustering in areas with relaxed enforcement. These spatial transformations influenced pricing strategies, product diversification, and customer engagement, highlighting the resilience of informal economies. Analyzing these shifts provides insights into the role of location in sustaining businesses, the impact of government policies on trader mobility, and the effectiveness of informal networks in responding to crises [20–22]. Beyond spatial patterns, research has also explored the socio-economic characteristics of market participants, revealing that traders often cluster based on shared factors such as ethnicity, product type, or business strategy. While these clusters strengthen social networks and facilitate resource sharing, they can also create entry barriers for new vendors, particularly those lacking capital or connections [23–25]. By integrating spatial analysis with socio-economic research, scholars can develop a comprehensive understanding of informal market structures, informing policies that balance market regulation with livelihood preservation while addressing economic inequality and social exclusion.

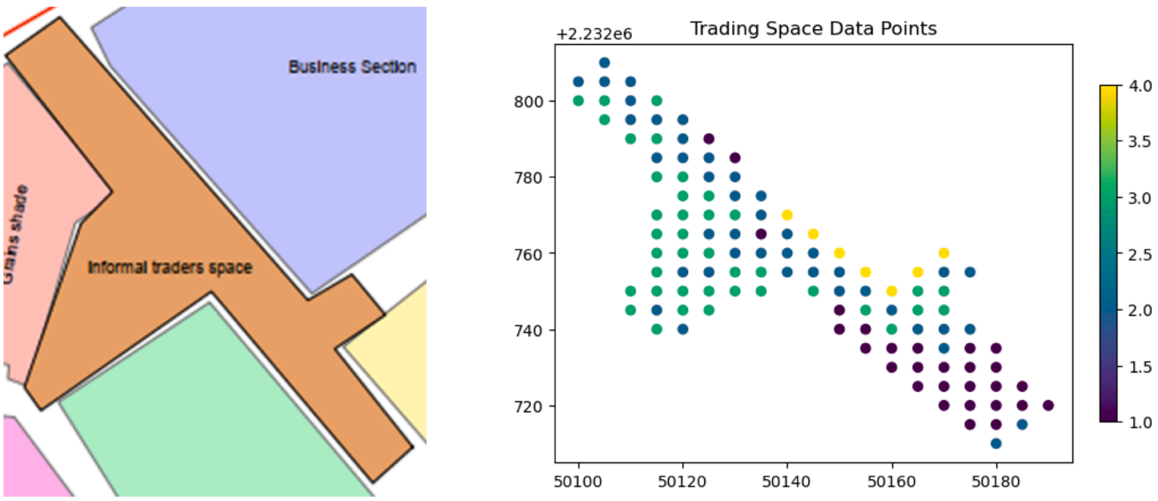
Methodology

In this section, we describe different methods of clustering data and tracking data structures and connectivity.

Research design and data collection

The informal trade area is encircled by a business section, grains and vegetables trading shades, and clothes trading shades in Fig. 1 (a), which describes the layout of the study area. Moreover, the study focused only on four types of traders who flooded this arena namely, those who traded clothing items coded as 1, those who traded food items like groceries coded as 2, those who traded fresh farm produce coded as 3, and money changers coded as 4.

A spatial regular tessellation is a process of dividing a geographical space into regular polygons (e.g., squares). In this case study, the informal trading space is divided into a regular grid of squares of area one square meter. Each square is observed to determine which trader or traders dominate this area. Hence, the spatial data points of informal traders are given in Fig. 1(b). Spatial tessellation is an important tool that can be used for zoning and partitioning of urban areas for land use planning. This tool provides capabilities for



((a)) Sokusile Informal Trading Space ((b)) Data Points

Fig. 1. Sokusile Informal Trading Space Data Points.

network analysis, visualization, and geographical data organization.

The informal trading space of Sekusile Market Place was represented by geographical data points embedded in a map shown in Fig. 1, which reflected the market dynamics. The distribution and dominance of traders in this trading arena are crucially shown in Fig. 1(b). By transforming this dataset into a matrix representation, important details on the trading landscape, market dynamics, trading hotspots, and geographical trader groups inside the Sekusile marketplace may be obtained through further data analysis techniques like hotspots analysis, K-means algorithm, ordinary kriging, and persistence homology.

Analytical technique

We begin this section by stating some of the advantages and disadvantages of different clustering methods.

Table 1 highlights the advantages of methods like K-means, Variogram/Kriging and Topological Data Analysis (TDA). K-means is known for its simplicity, speed, and effectiveness with large datasets, while Variogram/Kriging works well in handling spatial data and providing uncertainty estimates. Finally, TDA stands is critical when capturing topological features and dealing with complex data structures.

Table 2 presents some of the disadvantages of different clustering methods. K-means assumes spherical clusters and performs poorly with non-linear data, while Variogram/Kriging is computationally intensive and less effective with high-dimensional datasets. On the other hand, TDA is robust, computationally demanding, and its interpretation can be challenging for non-experts.

Table 1 and Table 2, provide an overview of different clustering methods strengths and weaknesses.

Hotspots analysis

Our references in the section are [26–28]. The classical hotspot analysis is a method used to identify areas of statistically significant clustering in spatial data. In this study, we make use of the Getis-Ord *Gi** statistic to perform hotspot analysis. The *Gi** statistic formula for a location i is

$$Gi^{*}(i) = \frac{\sum_{j=1}^{n} W_{ij}x_j - \overline{x}\sum_{j=1}^{n} W_{ij}}{s\sqrt{\frac{\sum_{j=1}^{n} W_{ij}^2 - \left(\sum_{j=1}^{n} W_{ij}\right)^2}{n}}} \tag{1}$$

where x_j represents the attribute value at location j, $\overline{x}$ is the mean, s is the standard deviation, W_{ij} is the spatial weight between locations i and j, and n is the total number of locations. From Eq. (1) a positive value highlights clusters of high values (hotspots), whereas a negative value highlights clusters of low values (cold spots). Hypothesis testing can be used to determine the significance of the *Gi** statistic.

K-means clustering

A spatial dataset may be divided into K different groups using the K-means clustering algorithm, an unsupervised machine learning approach. Also, the implementation is by reducing the within-cluster sum of squares as the primary goal of the K-means algorithm. Let $X = \{x_{11}, x_{12}, \ldots, x_{ij}\}$ be the dataset, where each x_{ij} is a data point in D-dimensional space. In the K-means algorithm, the following steps are involved:

1. Using the heuristic method, select K initial centroids $C = \{c_1, c_2, \ldots, c_k\}$.
2. For each point x_{ij} in the dataset, compute the centroid distance given by

$$d(x_{ij},\ c_k) = \sqrt{\sum_k (x_{ij} - c_k)^2} \tag{2}$$

The aim is to assign each data point x_{ij} to the nearest centroid $\min_k(d(x_{ij}, c_k))$.

Table 1
Advantages of different clustering methods.

Clustering Method	Advantages
K-means	Simple and fast to compute, Works well with large datasets, Easily interpretable clusters
Variogram/Kriging	Accounts for spatial correlation, Provides predictions with uncertainty estimates, Effective for geostatistical data
TDA (Persistence Homology)	Captures topological features, Handles complex data structures, Robust to noise and outliers

Table 2
Disadvantages of different clustering methods.

Clustering Method	Disadvantages
K-means	Assumes spherical clusters, Sensitive to initial centroids, Not suitable for non-linear data structures
Variogram/Kriging	Computationally intensive, Requires assumptions about variogram model, Less effective for non-spatial or high dimensional data
TDA (Persistence Homology)	Computationally demanding, Interpretation can be challenging, Less intuitive for non-topological data analysts

1. If S_j is the set of data points assigned to cluster k then recalculate the centroids based on the mean of the data points assigned to each cluster
2. $c_k = \frac{1}{|S_k|}\sum_{x_{ij}\in S_k} x_{ij}$.The last step is to repeat steps 2 and 3 until the convergence conditions are fulfilled (i.e., no change in cluster assignments or centroids).

The goal of K-means algorithms is to minimize the within-cluster sum of squares (WCSS)

$$WCSS = \sum_{k}\sum_{x_{ij}\in S_k} \|x_{ij} - c_k\|^2 \tag{3}$$

Overall, the K-means algorithm iteratively updates cluster assignments and centroids and partitions the dataset into K clusters. For more details on the K clustering method, the reader can refer to [29,30].

Geostatistics

For the concepts in this section, we refer to [31,32]. In geostatistics, the ordinary kriging method is used for spatial interpolation. specifically, the method generates estimated values at unobserved locations based on the observed data points. The process of performing ordinary kriging involves several key steps, which include computing semi-variogram, semi-variogram model fitting and computing kriging predicted values.

The semi-variogram $\gamma(h)$ measures the spatial variability between data points as a function of distance or vector h. Then, the semi-variogram is computed by:

$$\gamma(h) = \frac{1}{2N(h)}\sum_{i=1}^{N(h)}[z(x_i) - z(x_i + h)]^2 \tag{4}$$

where $N(h)$ is the number of pairs of sample points separated by distance h, $z(x_i)$ and $z(x_i + h)$ are the values of the variable of interest at locations x_i and $x_i + h$, respectively. Therefore, after calculating the semi-variogram, we fit a model to the experimental semi-variogram to represent the spatial correlation structure. Let a be the range defined as the distance that separates the correlated and uncorrelated variables. If C is the sill defined as the constant value of the variogram for any distance greater than the range a. Then we state three commonly fitted variogram models other than the linear model:

(a) Spherical variogram model

$$\gamma(h) = \begin{cases} C\left[\frac{3h}{2a} - \frac{h^3}{2a^3}\right] & for\ h \le 0 \\ C & for\ h > 0 \end{cases} \tag{5}$$

(b) Exponential variogram model

$$\gamma(h) = C\left[1 - e^{\left(-\frac{h}{a}\right)}\right] \tag{6}$$

(c) Gaussian variogram model

$$\gamma(h) = C\left[1 - e^{\left(-\frac{h}{a}\right)^2}\right] \tag{7}$$

Next, we give steps for performing the ordinary kriging method. To estimate unobserved locations, weights for observed data points are used based on their geographic proximity and spatial correlation with the target location. Let λ_i denote the kriging weights for each observed point. For ordinary kriging, the kriging weights λ_i should satisfy the following conditions:

1. Unbiasedness condition $\sum_{i=1}^{n} \lambda_i = 1$.
2. Let μ be a Lagrange multiplier $\sum_{i=1}^{n} \lambda_i \gamma(x_i - x_j) - \mu = \gamma(x_i - x)$ for $i = 1, \ldots, n$ where x is the target location.

These conditions, when combined, yield a system of equations called the Kriging system. Furthermore, after calculating the kriging weights λ_i, the projected values at the unobserved points may be determined by taking the linear combination of the observed values and the kriging weights given by:

$$\hat{z}(x) = \sum_{i=1}^{n} \lambda_i z(x_i).$$

This equation forms the basis of ordinary kriging by producing the predicted values at unobserved locations while considering spatial correlation and variability in the data.

Topological data analysis (TDA) persistence diagrams

In topological data analysis (TDA), there is a method for tracking persistent features of the data called persistence homology by utilizing persistence diagrams. The main objects used in constructing persistence diagrams are simplicial complexes. A simplicial complex $C(X)$ is an algebraic structure representing topological features in data X. It is a collection of simplices, which are basic geometric objects such as vertices, edges, and triangles, as well as their higher-dimensional equivalents. The standard $(n-1)$-dimensional simplex whose vertices are $[v_0, v_1, \ldots, v_{n-1}]$ are the n standard unit vectors in R^n, such that

$$\left\{ \sum x_i v_i : (x_0, x_1, \cdots, x_{n-1}) \in \mathbb{R}^n, \sum_{i=1}^{n-1} x_i = 1 \ \forall \ x_i \geq 0 \right\}.$$

So, $[v_0, v_1, \cdots, v_{n-1}]$ consists of all linear combinations of the form $x_0 v_0 + x_1 v_1 + \cdots + x_{n-1} v_{n-1}$. Again, a filtration of simplicial complexes is defined by the following sequence $C_0(X) \subseteq C_1(X) \subseteq \ldots \subseteq C_1(X)$ [33]. The primary advantage of the filtration condition for simplicial complexes is that it captures the development of topological features in data when a parameter, such as distance or size, changes.

Let R be a commutative ring and X be a topological space, then there are R-morphisms

$$\partial_n : C_n(X,R) \rightarrow C_{n-1}(X,R),$$

such that $\partial_n \circ \partial_{n+1} = 0$, the sequence

$$\cdots \overset{\partial_{n+2}}{\rightarrow} C_{n+1}(X,R) \overset{\partial_{n+1}}{\rightarrow} C_n(X,R) \overset{\partial_n}{\rightarrow} C_{n-1}(X,R) \overset{\partial_{n-1}}{\rightarrow} \cdots$$

is called a chain complex [33]. The n^{th} homology group $H_n(X,R)$ of R with coefficients in R is defined by

$$H_n(X,R) = ker\partial_n / Im\partial_{n+1}.$$

The R-modules ker ∂_n and $\mathrm{Im}\partial_{n+1}$ are called cycles and boundaries, respectively. Note that the homology group $H_n(X, R)$ is an algebraic invariant that describes the n-dimensional holes or cycles in the simplicial complex $C(X, R)$ [33].

In data science, high-dimensional point cloud data can be analyzed to uncover topological features using a technique called *persistent homology*. While traditional homology measures "holes" or voids in a dataset at fixed dimensions, persistent homology captures how these topological features evolve across scales or resolutions. Specifically, *Betti numbers* are used to count these holes. The n^{th} *Betti numbers* β_n is given by

$$\beta_n = dim_R H_n(X; R),$$

which quantifies the n-dimensional holes in a simplicial complex. Moreover, the *Betti numbers* are useful tools in computing the Euler characteristic

$$\chi(X) = \sum_{n=0}^{\infty} (-1)^n \beta_n$$

which is the alternating sum of *Betti numbers* [5].

In topological data analysis (TDA), persistent homology reveals the topological signature of a dataset. A loop attached to a fixed point with a varying radius r connects nearby data points. As the radius increases, changes in the topology are observed; some loops emerge (birth), while others disappear (death). This process is visually represented by a *barcode diagram*, where each bar tracks the lifespan of a topological feature, typically in the form of 1-dimensional loops. The *Rips complex* is another key concept in TDA, consisting of simplices formed by n-nodes, with edges determined by pairwise distances between points, often measured using Euclidean distance. By adjusting the radius, the Rips complex grows, providing insights into the dataset's topology across different dimensions. These mathematical tools form the foundation of TDA, offering a robust framework for extracting and analyzing topological information from complex datasets [34,35].

Results

Here we present our main findings together with comparative analysis of the following clustering methods: hotspots analysis, K clustering, Kriging method and Persistence homology.

Hotspots and clusters of trader groups

This section examines the spatial concentration of traders by identifying hotspots and clustering patterns within the market. By applying spatial analysis techniques, we uncover regions of high trader density and explore the underlying factors driving these formations.

Fig. 2 presents the results of 3-means clustering applied to the traders' dataset, which consists of four distinct types of traders: clothing sellers, grocery vendors, vegetable traders, and money changers. In this clustering, the traders are divided into three distinct clusters. This clustering likely represents broader trade categories, possibly grouping grocery and vegetable vendors together due to their similar product types, while money changers might form a separate cluster. The spatial arrangement of these clusters indicates the density of trader activity in particular sections of the market, reflecting how traders coalesce based on both product type and customer traffic. The clustering also reveals that some types of traders tend to dominate certain areas, influencing spatial competition in the marketplace.

In Fig. 3, the 4-means clustering approach is applied to the same traders' dataset, effectively partitioning the data into four clusters. Each cluster corresponds more closely to the specific types of traders observed in the Sekusile market: clothing sellers, grocery vendors, vegetable traders, and money changers. The four clusters show a more detailed and precise breakdown of trader distribution, with each group occupying distinct areas of the market. This clustering reflects not only the natural segmentation of traders based on the goods they offer but also the competitive dynamics at play. Traders of similar products are likely grouped together, indicating that each cluster corresponds to a localized economic zone where traders vie for space and customers.

Spatial variability in trading behavior

The spatial distribution variability in trading behavior was assessed using geostatistical techniques, specifically through the fitting of a variogram and the application of ordinary kriging. These methods provided a quantitative measure of spatial dependence, allowing for the identification of clustering patterns and market dynamics.

The variogram in Fig. 4 shows the spatial analysis of informal traders' correlation across different points within the trading space. In particular, the variogram shows the variance, and the sill value is estimated to be around 0.9, which is the spatial correlation corresponding to insignificant data points. Since this variogram reaches a sill in the range of about 42 m This implies that the data points beyond these distances are uncorrelated with a variance of about 0.9. This smaller scale for the range 42 implies a less broad spatial correlation i.e., localized spatial correlation. The fitted model is the Gaussian model. Therefore, the spatial correlation decreases rapidly in the neighborhood of the origin, indicating strong spatial dependence at short distances. Note that about three points on the variogarm lie far away from the fitted model, which may indicate small-scale variability, which might not also be captured by the spatial model. These findings have important policy and planning implications, as the significant range suggests the need for zoning strategies. So as to optimize resource allocation and crowd management. The small nugget effect implies predictable trading behaviors,

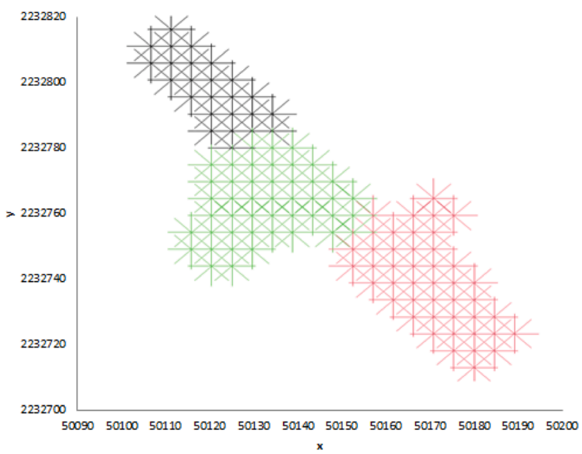


Fig. 2. 3-means clustering.

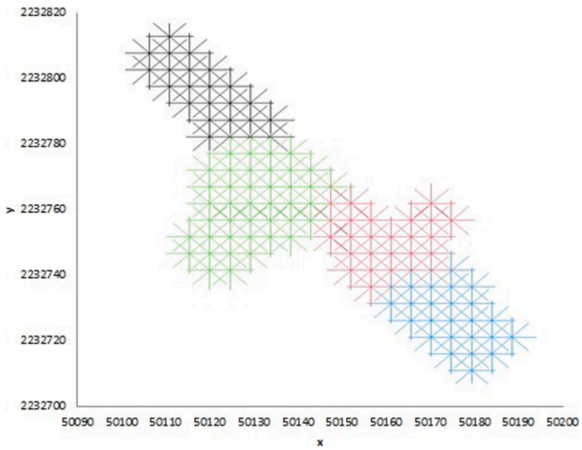


Fig. 3. 4-means clustering.

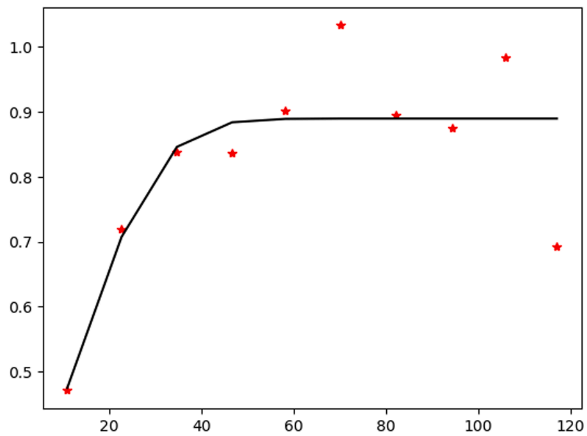


Fig. 4. The variogram.

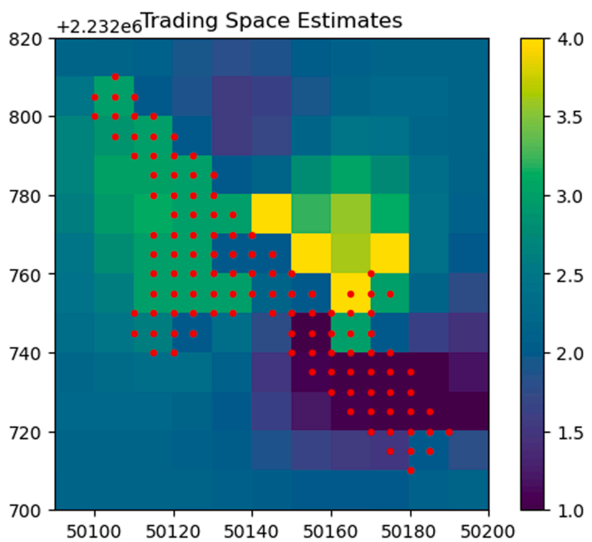


Fig. 5. Ordinary kriging.

allowing for improved market infrastructure planning. Additionally, the moderate sill value supports targeted intervention strategies to prevent excessive clustering and maintain balanced trader distribution. Overall, these insights are valuable for market regulation, congestion control, and resilience planning, particularly during external disruptions like pandemics or economic shocks.

Fig. 5 presents the results of spatial interpolation using the ordinary Kriging method, offering key insights into the organization of traders within the market space. The interpolated grid highlights distinct clustering patterns that have significant practical implications for market regulation and resource allocation. Traders labeled 1 form a prominent cluster in the lower-right corner, suggesting a well-defined trading zone, potentially benefiting from strategic positioning and high customer footfall. Traders labeled 2 are concentrated in the central area, indicating a competitive core market region where accessibility and customer engagement may be maximized. Traders labeled 3 exhibit two separate clusters, occupying a large portion of the space, making them the most dominant group in terms of spatial influence. Their widespread presence suggests potential market congestion issues or the need for zoning interventions to prevent overcrowding. In contrast, traders labeled 4 do not form a complete grid, suggesting minimal influence on the broader market dynamics. Their scattered presence may indicate lower demand, weaker customer engagement, or regulatory constraints affecting their placement. The dominance of traders labeled 3 suggests that strategic interventions, such as redistributing trading areas, optimizing stall allocations, or enhancing accessibility, could improve market efficiency and fairness. Additionally, policymakers and urban planners can use this spatial data to design infrastructure improvements, such as walkways, designated trading areas, and congestion mitigation strategies, ensuring a more balanced and sustainable trading environment.

Topological features of the market structure

In this section, we present the topological features results for understanding traders' spatial dynamics in the informal market. In addition, we show the spatial distributions of each trader type by presenting the persistence diagram alongside the barcode diagram.

According to Fig. 6, the persistence diagram for Trader 1 illustrates the birth and death of topological features. Points closer to the diagonal line (Birth = Death) correspond to features that are short-lived, while points further away from the diagonal represent more persistent features. The proximity of points to the diagonal indicates a more complex structure in the data distribution. Specifically, the H_0 connected components are divided into two groups, reflecting a balance between high and low levels of complexity in the data. The H_1 connected components are situated closer to the diagonal, indicating lower complexity in these data points. In the barcode diagram, the longer H_0 bars represent persistent features, suggesting that Trader 1′s data points have more significant topological features. Conversely, the shorter H_1 bars indicate less significant one-dimensional components. Overall, the clustering of Trader 1 (second-hand clothing vendors) within the informal trading space suggests that these vendors are connected both locally and globally, but the connectedness and clustering at the H_1 level are short-lived.

Fig. 7 shows increasing persistence in the connected components, while the H_1 points

(loops) are closer to the diagonal, indicating short-lived features. The barcode for Trader 2 features a broken line representing H_1 (one-dimensional homology), indicating the presence of loops that appear and then disappear as the filtration progresses. The short bars in H_1 imply that the connected components are not very persistent, meaning they quickly merge with other components as the filtration parameter increases. The H_0 homology, rep- resenting connected components in the data, is depicted with a longer bar, indicating the significance of this feature. This longer bar indicates the persistence of these components over a range of filtration parameters. This analysis helps in understanding the underlying structure of the data, particularly how data points cluster together and how these clusters evolve as the scale of observation changes. This suggests that two significant clusters can be detected by loops in H_1. Overall, the clustering of Trader 2 (grocery vendors) within the informal trading space indicates that these vendors are connected both locally, and the H_1 loops measure two significant clusters at this level.

The diagrams in Fig. 8 provide insights into the topological structure of Trader 3′s dataset. Most of the connected components are persistent, as they lie far from the diagonal line. Additionally, the H_1 component shows some significant persistence, as many of them lie far above the diagonal line. The barcode diagram complements the persistence diagram, showing the lifespan of features, with connected components H_0 having longer lifespans and loops or one-dimensional components having shorter lifespans. Overall, the clustering of Trader 3 (vegetable vendors) within the informal trading space indicates that these vendors are connected both locally and globally.

The persistence diagram for Trader 4, shown in Fig. 9, is represented by a few data points in the top right, indicating persistent H_0 connected components, with no H_1 points representing the topological features of the data. The barcode diagram corresponds to the results presented in the persistence diagram, with one longer H_0 bar indicating more significant connected component features. This suggests that the spatial distribution of Trader 4 (money changers) in this network is compact and localized to one cluster.

Discussion

In this study, we analyzed the spatial dynamics of traders in Sekusile market. Our focus was on informal traders during the COVID-19 pandemic, we provided significant insights into how clustering patterns influence market organization and competition. The use of clustering techniques like persistence diagrams, K-means, and ordinary kriging has revealed the intricate ways in which traders come together to create unique spatial clusters within the market. The clustering patterns observed in the results section figures correspond with previous studies on informal markets in sub-Saharan Africa, which indicated that spatial clustering is a frequent characteristic of these markets. Research such as that by Potts [16] and Kesteltoot & Meert [9] have illustrated how traders in informal markets tend to cluster in particular manners that provide strategic benefits, like significant consumer foot traffic and closeness to suppliers. The competitive dynamics emphasized in this research, where vendors cluster according to product type and pricing strategies, further

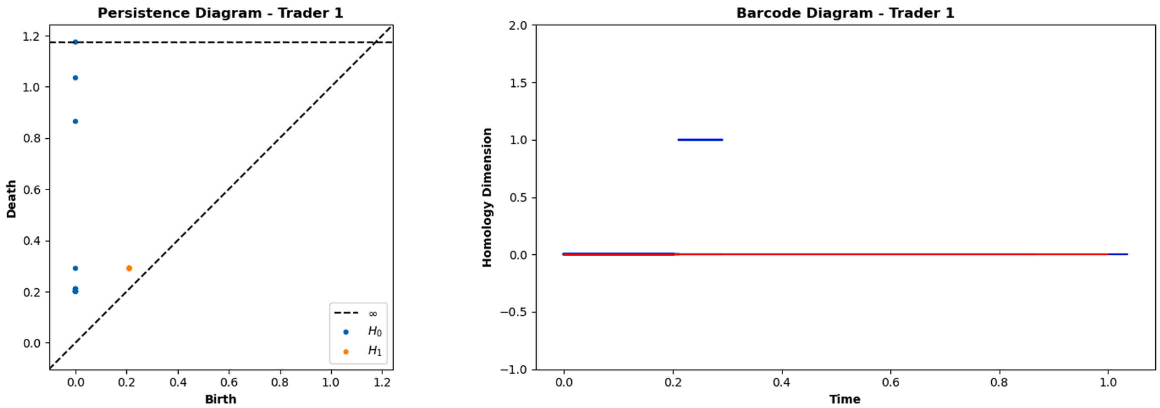


Fig. 6. Trader 1 Persistence and barcode diagrams.

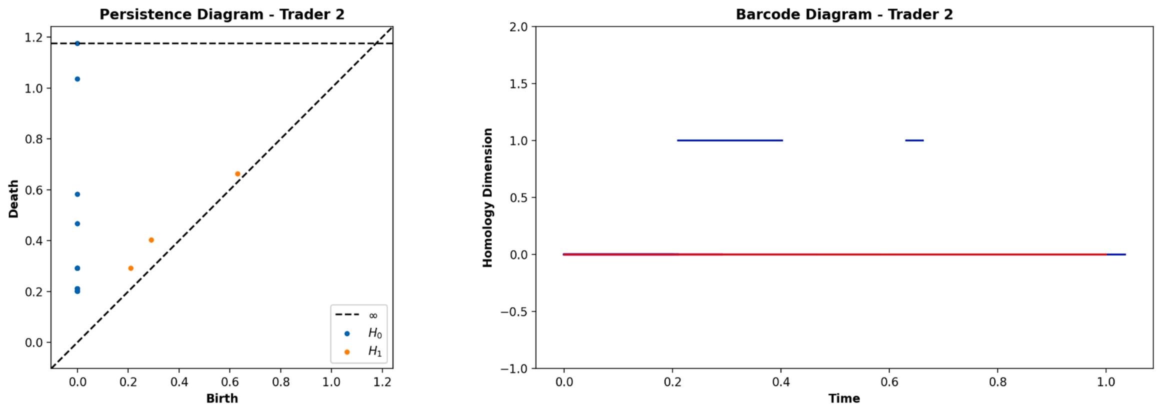


Fig. 7. Trader 2 Persistence and barcode diagrams.

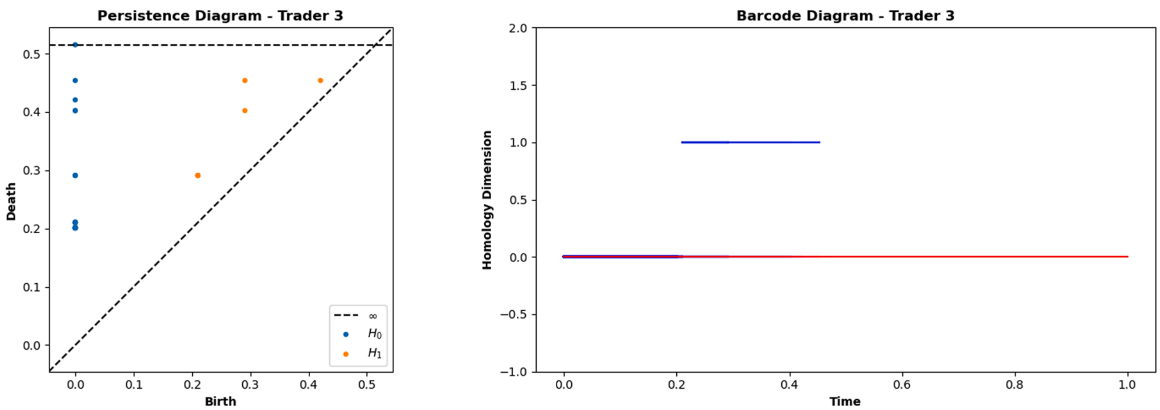


Fig. 8. Trader 3 Persistence and barcode diagrams.

reinforce the idea that spatial clustering serves as a fundamental organizing principle in informal markets. Previous work by Mukhtarova in [20] and Narula in [21] emphasized the resilience of informal traders in the face of external shocks, such as economic crises and regulatory changes. As seen during the COVID-19 pandemic, the spatial shifts experienced in Sekusile market reflect similar patterns, where traders adapted to changes in consumer behavior and market regulations by forming new clusters, creating a more fluid and dynamic market structure. Furthermore, studies by Johnson [8] and Kima & Sugiri [24] highlighted the role of social networks and shared characteristics in shaping the organization of traders, further explaining why certain types of traders, such as vegetable vendors, tend to cluster together in specific zones. Topological data analysis (TDA), in particular the use of persistence

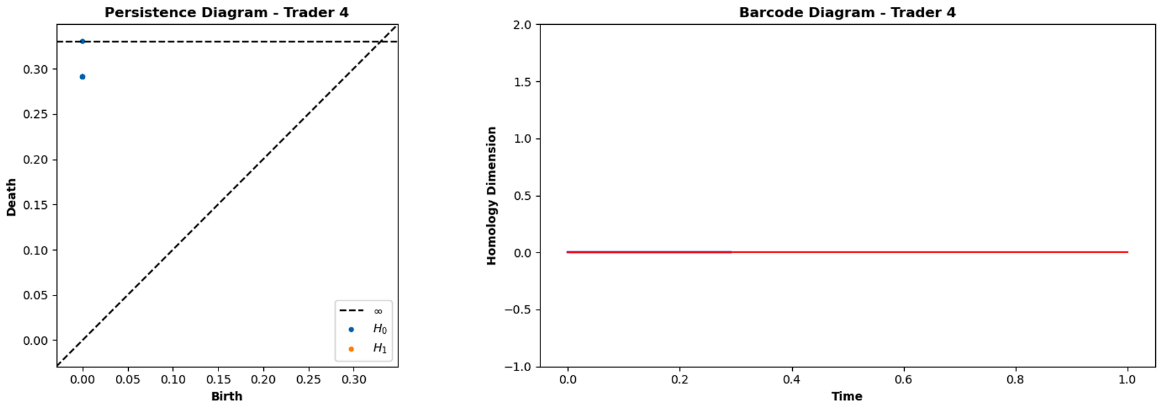


Fig. 9. Trader 4 persistence and barcode diagrams.

diagrams in this study, adds another layer of depth to our understanding of market dynamics. Persistence homology, which tracks the evolution of topological features, reveals the connectivity and endurance of trader clusters. This approach has been shown to be effective in capturing spatial patterns in other complex datasets, including those in geostatistics. By applying TDA, this study provided a novel method for visualizing how trader clusters form and persist over time, offering a unique perspective on informal market dynamics that has not been widely explored in the existing literature.

Conclusion and implications

This study has provided a comprehensive analysis of the spatial dynamics of informal traders in Sekusile Market during the COVID-19 pandemic, highlighting the intricate patterns of clustering and competition that shape informal economic activities. By employing advanced clustering methods such as K-means, ordinary kriging, and persistence diagrams, we identified distinct trader clusters based on product type and spatial organization. These findings underscore the role of spatial competition in informal markets and contribute to the growing literature on market resilience in the face of economic shocks. Beyond academic insights, this study offers practical and policy-relevant implications for urban planning, economic regulation, and market sustainability. The clustering patterns observed suggest that informal traders naturally form localized economic zones, which can be leveraged to design better market layouts that minimize conflicts over trading space while optimizing customer access. Policymakers can use spatial analysis to inform zoning strategies, designating specific areas for different types of traders to reduce congestion and enhance market efficiency.

Furthermore, the resilience of informal markets, as observed during the pandemic, highlights the need for adaptive policy frameworks that support rather than suppress informal trade. Authorities should consider formalizing aspects of informal markets by providing legal recognition, infrastructure improvements, and access to financial services. This can help integrate informal traders into the broader economy while maintaining the flexibility that enables them to respond to external shocks. Additionally, regulatory measures should be data-driven, using real-time market analytics to monitor fluctuations in trader distribution and customer flow. Future research could expand on these findings by examining the long-term effects of economic disruptions on trader mobility and exploring how digital platforms can enhance informal market resilience. By integrating real-time data collection and geospatial analysis, policymakers can develop dynamic regulatory frameworks that balance economic inclusivity with urban development.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work, the author(s) used ChatGPT-4o mini as language editing tool. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

CRediT authorship contribution statement

Meshach Ndlovu: Data curation, Formal analysis, Methodology, Writing – original draft, Writing – review & editing.

Declaration of competing interest

I, Meshach Ndlovu, the author of the manuscript titled "**Clustering and Spatial Dynamics of Informal Trading: A Case Study of Sekusile Market, Bulawayo, During the COVID-19 Pandemic**", declare that I have no financial, personal, or professional conflicts of interest that could have influenced the research presented in this study. Additionally, this research did not receive external funding and was solely developed by me.

Acknowledgments

I would like to express my sincere gratitude to the authors and researchers whose insightful works provided valuable references to this study. I am especially grateful to Mr. E. Zulu for suggesting the research problem, which served as the foundation for this project.

References

[1] 2023 S Quarter, Quarterly Labour Force Survey Report. Tech. Rep. (2023).
[2] M. Ndlovu, R. Moyo, M. Mpofu, Modelling covid-19 infection with season- ality in Zimbabwe, Phys. Chem.Earth, Parts A/B/C 127 (2022).
[3] I. Kabonga, K. Zvokuomba, B. Nyagadza, The challenges faced by young entrepreneurs in informal trading in Bindura, Zimbabwe, J. Asian. Afr. Stud. 56 (8) (2021) 1780–1794.
[4] I. Clark, et al., Practical Geostatistics, 3, Applied Science Publishers London, 1979.
[5] H. Edelsbrunner, D. Letscher, A. Zomorodian, Topological persistence and simplification, Discrete. Comput. Geom. 28 (2002) 511–533.
[6] J.A. Hartigan, M.A. Wong, Algorithm as 136: a k-means clustering algorithm, J. royal stat. society. series c 28 (1979) 100–108.
[7] H. Ngalawa, N. Viegi, Interaction of formal and informal financial markets in quasi-emerging market economies, Econ. Model 31 (2013) 614–624.
[8] S. Johnson, M. Malkamaki, M. Niño-Zarazua, The role of informal groups in financial markets: evidence from Kenya, University of Bath (2010).
[9] C. Kesteltoot, H. Meert, Informal spaces: the geography of informal eco- nomic activities in brussels, Int. J. Urban Reg. Res. 23 (1999) 232–251.
[10] S. Patel, R. Furlan, M. Grosvald, A framework for enhancing the spatial urban form of informal economies in India: the case of Krishna Rajendra Market, Bangalore, Sage Open. 11 (2021) 21582440211023184.
[11] E. Sik, The spatial distribution of informal marketplaces and informal foreign traders in contemporary Hungary. Underground Economies in Transition, Routledge, 2019, pp. 275–306.
[12] M.E. Porter, et al., Clusters and the New Economics of Competition, 76, Harvard Business Review Boston, 1998.
[13] J.E. Fantcho, P.K. N'Gouan, Spatial characteristics and dynamic modeling of informal economies, Int. Econ. Econ. Policy 21 (2024) 483–514.
[14] D. Hantono, W. Nuryanti, D. Pramitasari, Spatial character of informal markets in Jakarta, in: AIP Conference Proceedings 2689, AIP Publishing, 2023.
[15] E.K. Avenyo, M. Konte, P. Mohnen, Product innovation and informal market competition in sub-Saharan Africa, J. Evol. Econ. 31 (2021) 605–637.
[16] D. Potts, The urban informal sector in sub-Saharan Africa: from bad to good (and back again?), Dev. South Afr. 25 (2008) 151–167.
[17] J.L. Roitman, The politics of informal markets in sub-Saharan Africa, J. Mod. Afr. Stud. 28 (1990) 671–696.
[18] K. Karunagoda, Informal markets for financial services in Sri Lanka, Deputy Director of Central Bank of Sri Lanka, 19th Anniversary Convention (2007).
[19] E.O. Frank, An assessment of the conceptual linkages between the qualitative characteristics of useful financial information and ethical behavior within informal institutions, Ekonomski horizonti 22 (2020) 137–148.
[20] T. Mukhtarova, Covid-19 and the Informal Sector, Georgetown Institute for Women, Peace and Security, 2020.
[21] R. Narula, Policy opportunities and challenges from the covid-19 pandemic for economies with large informal sectors, J. Int. Business policy 3 (2020) 302–310.
[22] C. Skinner, V. Watson, Planning and informal food traders under covid-19: the South African case, Town Planning Rev. 92 (2021) 301–307.
[23] G.A. Garcça, A.I. Moreno-Monroy, Intra-metropolitan clustering of formal and informal manufacturing activity: evidence from Cali, Colombia, University of Groningen, Groningen, The Netherlands (2012).
[24] U. Kima, A. Sugiri, A study on informal market cluster activities concentrating around mendi market area: is relocation the best option? Tataloka 17 (2015) 12–20.
[25] Mukim, M. Industry and the urge to cluster: a study of the informal sector in India, 2011.
[26] N.L. Harris, E. Goldman, C. Gabris, J. Nordling, S. Minnemeyer, S. Ansari, M. Lippmann, L. Bennett, M. Raad, M. Hansen, et al., Using spatial statistics to identify emerging hot spots of forest loss, Environ. Res. Lett. 12 (2017) 024012.
[27] U. Manepalli, G.H. Bham, S. Kandada, Evaluation of hotspots identification using kernel density estimation (k) and getis-ord (gi*) on i-630, in: 3rd International Conference on Road Safety and Simulation 21, National Academy of Sciences Indianapolis Indiana, United States, 2011, pp. 14–16.
[28] B.W. Silverman, Density Estimation For Statistics and Data Analysis, Routledge, 2018.
[29] S. Na, L. Xumin, G. Yong, Research on k-means clustering algorithm: an improved k-means clustering algorithm, in: 2010 Third International Symposium on intelligent information technology and security informatics, 2010, pp. 63–67.
[30] J. Wu, Advances in K-means clustering: a Data Mining Thinking, Springer Theses. Springer Science & Business Media, 2012.
[31] N.A.C. Cressie, Statistics For Spatial Data, John Wiley & Sons, Hoboken, NJ, 2015 revised paperback edition ed.
[32] D.D. Sarma, Geostatistics With Applications in Earth Sciences, Springer Science & Business Media, 2010.
[33] Y. Felix, S. Halperin, J.C. Thomas, Rational homotopy theory, of, in: Graduate Texts in Mathematics, 205, Springer, New York, 2001.
[34] E. Munch, A user's guide to topological data analysis, J. Learn. Analy. 4 (2017) 47–61.
[35] A. Zomorodian, G. Carlsson, Computing persistent homology, in: Proceedings of the twentieth annual symposium on Computational geometry, 2004, pp. 347–356.