International Journal of Applied Earth Observation and Geoinformation 128 (2024) 103764

Contents lists available at ScienceDirect

International Journal of Applied Earth Observation and Geoinformation

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

Spatiotemporal mapping of urban trade and shopping patterns: A geospatial big data approach

Bakhtiar Feizizadeh [a], Davoud Omarzadeh [b], Thomas Blaschke [c,*]

[a] Department of Remote Sensing and GIS, University of Tabriz, Iran
[b] IN3, Universitat Oberta de Catalunya, Barcelona, Spain
[c] Department of Geoinformatics, University of Salzburg, Austria

ARTICLE INFO

Keywords:
Shopping pattern mapping
GIS
Geospatial big data
Data-driven approaches
Spatially explicit
GIScience
Novel methodology

ABSTRACT

The economic viability of an urban area in terms of trade and shopping significantly impacts its residents' quality of life and is crucial for any sustainable development initiative. Geographic information systems (GIS) are well established, but the use of GIS technology within finance and trade analysis is still in its infancy. In this article, we highlighted the potential of GIS technology and big data analytics and demonstrated the importance of thinking in spatial terms for analysing patterns within the trade and finance industries. We studied spatiotemporal trade and shopping patterns in the city of Tabriz using data generated by customer purchase transactions obtained from 5200 stores, shopping, business and service centres. We employed time series transaction data collected from the points of sale in stores, shopping, service and business centres located in different areas of the city. We applied four well known geospatial big data driven approaches including machine learning nearest neighbour, kernel density estimation, space–time pattern mining and spatiotemporal coupling tele-coupling for detecting and mapping of spatial trade hotspot patterns. The results of this study indicated the potential of GIScience methods for the explicit spatial mapping of trade and shopping patterns. The results revealed that the city centre, particularly the Bazaar of Tabriz, acts as the city's heart of trade, and we identify additional major business hotspots. Furthermore, the results allow for studying the impacts of unbalanced urban development in Tabriz, where the wealthy suburbs with high quality of life, such as Valiasr and Elguli, host the major shopping hotspots. The spatial patterns obtained enable local stakeholders, decision makers and authorities to develop strategic plans for urban sustainable development in Tabriz. The geospatial big data approach used can stimulate novel and progressive research. Results of this study demonstrate methodological advancements in GIScience by 'spatializing' individual purchase data and therefor proposing an explicit geospatial big data analysis approach.

1. Introduction

The socioeconomic status of an urban area which is known as status of a population group (e.g. education, household income, employment status and etc) significantly impacts its residents' quality of life and is crucial for any sustainable development initiative. The term of 'socio-economic status addresses an individual's position in a society which is determined based on the wealth, occupation, and social class in order to measure of an individual's or group's standing in the community (Sarsani, 2011). Since the early 1980s, improvements in global communication have significantly influenced globalisation, labour market upheaval, economic liberalisation and socioeconomic segregation (van Ham et al., 2021).

Nowadays, socioeconomic characteristics can be observed and mapped as big data through advanced geographic information science (GIScience) methods, such as data acquisition, mining, pattern and trend detection. In this context, the recent advance in geocommunication technology allows geotagged information to be generated using tools such as mobile devices, sensors, e-banking devices etc. More recently, spatial datasets have been generated from location-based social networks and analyses of transactions with location characteristics from credit and debit cards. Obviously, this has opened new opportunities for socioeconomic analysis which leads to the potential of spatiotemporal trade and shopping patterns to be evaluated and analysed (Martí et al., 2019; Hladík et al., 2021). Such data have been utilised in many studies to better understand human and environmental dynamics (Qiang and

* Corresponding author.
E-mail address: Thomas.Blaschke@sbg.ac.at (T. Blaschke).

https://doi.org/10.1016/j.jag.2024.103764
Received 6 June 2023; Received in revised form 11 February 2024; Accepted 8 March 2024
Available online 12 March 2024

Van de Weghe, 2019).

In this context big data are distinguished by pace and volume, and the instant accessibility through sensors and dispersion via the internet such as credit and debit cards. Technically speaking, the terms of big data and its utmost elementary meaning, references to a significant expanse of information that are collected from different resource. Basically, big data can be found in variety of data types and comprise a wide range of data across a variety of resource (Deng et al., 2023). Early studies indicated several characteristics for big such as: volume, variety, variability, velocity, veracity, value, volatility and validity (Deng et al., 2023). It is well understood that intensive progress and availability of social media platforms, personal computers, smartphones, tablets, intelligent autonomous sensors and pervasive network interactions with individuals implies the most of human activities from offline to online mood (Miller and Goodchild, 2015; Liu et al., 2022). Nowadays based on the capability of Internet of Things (IoT) support that the massive extraction of various information and data via the advancement and proliferation of linked mobile devices, detectors and similar platforms (Deng et al., 2023). The ever-increasing flow of big data, from various types of sensors, messaging and social media systems and platforms, as well as more traditional measurement and observation systems, has invaded many aspects of daily life to be mapped and analysed.

Big data has great potential to benefit applications such as the analysis and mitigation of climate change and its impacts, disease monitoring, disaster response, geomarketing, life quality, crime analysis or the monitoring of critical infrastructures and transportation (Li et al., 2016). Socioeconomical perspective is on ethe major area for big data anysis and application. Intensive progress in big data and availability of wide range of economic big data (e.g. credit and debit cards, online shopping and etc) created opportunities for economists and policy-makers to learn about economic systems and spatial patterns more efficiently (Harding and Hersh, 2018). One the major advantage of big data is that this dataset is that they provide very critical detailed information about economic behaviours and decisions which has spurred research aiming at answering long-standing economic questions. In addition, they might be available in real time, a property that can be employed to construct economic indicators at high frequencies. In the highly aggregate context of macroeconomic analysis, big data provide the opportunity to integrate to light the heterogeneity in consumers and firms that might be neglected in official and traditional statistics. The high granularity of big data can be subjugated to construct indicators that are better designed to express certain economical phenomena such as geographical trading and geomakrting (Manzan, 2023).Big data analytics is a trend in GIScience that provides methodological approaches for GIS-based spatial analysis of structured and unstructured big data and can enable new approaches to scientific inquiry and thought that depart from conventional research. Big data is based on a data-driven research pattern of thought that allows the data to speak for itself and it is considered different from just higher volumes of data (Mayer Schonberger and Cukier, 2013). Big data analysis allows for exploring statistical correlations between different data rather than just causality (Miller and Goodchild, 2015; Xu and Zhang, 2021). Based on the availability of big data implied through of the advanced location based social media, the cyberspace has evolved as a relatively new approach in GIScience which leads to analyse the dimensions beyond geographic space (Liu et al., 2022). Accordingly, the 'cyber thinking' received a significant attention in the past years to emerge the human activities (e.g., shopping) and the geographic space (e.g., location of the markets and their characteristics) (Lu et al., 2019; Goodchild, 2022).Based on the ongoing developments in digital transformation, Geomarkting, decentralized, sensor-based technologies and related technologies, geospatial big data (GBD) has been identified as paradigm shifting while throwing a new light on GIScience (Dangermond and Goodchild, 2019; Liu et al., 2022). In this context, GBD with its defining characteristics of being large (voluminous), heterogeneous (variety), real-time processed (velocity), inconsistent (variability), and thus also of variable quality (veracity), must suffer even more from uncertainty, asynchronicity, and incompleteness (Liu et al., 2022). Technically speaking, the GDB, being abstractions and observations of a continuous reality is by nature uncertain, ideally time-stamped and often incomplete (Frank, 2001; Liu et al., 2022).

Despite the related privacy concerns, technical and methodological developments that analyse individuals' digital footprints have revolutionised the potential amount and granularity of the data potentially available for various socioeconomic analyses (Blazquez and Domenechl, 2018). Such data can reveal the socioeconomic characteristics of a society in ways that were not previously possible. As almost all stores and markets make use of credit and debit cards for payment and transactions, a large number of daily socioeconomic datasets have become available, and credit and debit card transaction data are clearly 'Big data'. Our society produces 'Big data' every day. This fact has triggered a paradigm shift in GIScience that provides a setting for socioeconomic policy, strategy development, business management and decision-making analysis (Blazquez and Domenechl, 2018). Using spatial patterns of shopping and quantifying 'customer behaviour characteristics', which is often regarded as a qualitative variable, can be mapped in the context of big data science through geospatial technologies such as GIScience and geomarketing. From a geomarketing perspective, such information allows for better decisions about where to choose a store location with potentially better returns and market management, as well as future design layout. In this context, location-based analysis, such as query, ideal product routeing, spatial distribution and geomarketing, has mostly dominated geographic information system (GIS) applications in the commercial sector. Research and application of geospatial technology on customer behaviour, such as trade patterns, trend mapping and shopping hotspots, are some examples indicating the significance of geomarketing and how critical setting up a business can be (Habibpour et al., 2021; Feizizadeh et al., 2023).

Due to the high number of financial transactions in the form of e-banking, e-commerce and urban sensors (credit card readers, public transport card readers, tolls, retail scanners, etc.), these data can be employed to analyse the major socioeconomic aspects through big data technology (e.g., Van Vlasselaer et al., 2015; Blazquez and Domenechl, 2018). Credit/debit card readers, such as point of sale (POS), are among the most widely used devices in almost all markets around the world. The amount and characteristics of this 'new data' obtained from credit/debit cards and their respective transactions can support a variety of applications, such as fraudulent purchases (Van Vlasselaer et al., 2015), personal bankruptcy (Xiong et al., 2013) and default repayment – reaching a new level of automated checking and control analytics. Retail scanners record shoppers' daily purchases. The resulting data can be employed to analyse and actively influence shopping/trade patterns and customer behaviour based on initial insights into typical and unusual temporal and spatial concentrations. While the analysis of individual behaviour is very critical in terms of privacy and is restricted in most parts of the world, applications such as identifying the hotspots of sales in an urban area or predicting consumer behaviour are increasingly used in operational business (Blazquez and Domenechl, 2018). Big data architectures have been adapted to specific domains and purposes and have been utilised to generate high-quality data with unprecedented temporal and spatial resolutions. Such architectures should be able to manage the entire data lifecycle in an organisation, including ingestion, analysis and storage (Blazquez and Domenechl, 2018).

The number of studies that use spatial data generated from business and economics has been significantly increasing in recent years. In addition, researchers with a limited background in spatial sciences have become interested in using such new data-generation methods that can reveal business information and revenues, as well as spatial coordinates to identify supply and demand patterns and economic and commercial areas in the context of geomarketing (Lim, 2023; Lim et al., 2022; Gómez et al., 2020; Chacón-García, 2017; Omidipoor et al., 2019). Big data analysis of credit card transactions and their spatial patterns can provide previously unattainable insights into the spending patterns and mobility patterns of a sizeable population (Alefo et al., 2019). Furthermore, exploring such spatiotemporal patterns at the community level necessitates complex system modelling and parametrisation, as well as an accurate depiction of the temporal dynamics. Particularly in developing societies with traditional marketing systems with high numbers of small stores and businesses (e.g., Iran), the processing and computing of data collocated through customers' transactions allows for recognising hotspots in a shopping area and business spatial pattern. A hotspot can, in fact, be intermittent (present only occasionally due to local events) or permanent (present throughout the observation period) (Alefo et al., 2019). Therefore, temporal and spatial analyses are typically combined and sometimes referred to as spatiotemporal analyses.

GIScience, including GIS technology and other geoinformatics methods, has already been considered an efficient approach to big data processing and identifying spatial patterns involving a variety of variables and trend assessment. As data quality is critical, GIScience offers the means to analyse data quality and accuracy. Business and economic utilities depend on reliable data, cutting-edge technology and science-based decision support systems for successful business goals (Kamil et al., 2021). It has been claimed that the ability of GIScience to incorporate data from disparate sources, interpret it and present map-based knowledge to help marketing-specific decisions sets it apart from other technologies (Omidipoor et al., 2019). In addition, the capability of GIScience to link attribute data, such as demographic and socioeconomic data, with spatial entities and context (e.g., topological, geometric or geographical properties) allows for minimising the uncertainty in making decisions about developing and extending markets, such as real estate and retail (Suárez-Vega et al., 2012; Roig-Tierno et al., 2013; Cliquet 2013; Chacón-García, 2017). Considering the significance of 'Big data' and the need for more efficient methodologies for this new type of data generation, developing an approach that can be effectively applied to spatially analyse shopping patterns is critical. Based on the significance of GBD for the field of GIScience, it is expected that new data acquisition, mining and spatiotemporal pattern analysis approaches that are recently developed – also in this study – can respond to this raising demand. Thus, the present study aims to develop and apply GIScience methods utilising a –driven approach for spatial sale and shopping patterns for the city of Tabriz, Iran, using data generated by customers' purchase transactions obtained from 5200 stores, shopping, business and service centres. More specifically, we aim to identify the citizen's shopping pattern behaviour in the city and analyse the hotspots of trade applying different geospatial analyses.

2. Materials and methods

2.1. Study area

The study area was Tabriz, which is the capital of the East Azarabijan Province and the fourth-largest city in Iran. The city has about 2 million inhabitants and is known as a major trade centre in the northwest of the country (Iranian Census Centre, 2016). Fig. 1 shows the geographic location of Tabriz, which is also known as a major economic centre due to its intensive industrial, cultural, political, commercial, educational and tourism activities (Pourmoradian, 2018). The Bazaar of Tabriz is inscribed as a United Nations Educational, Scientific and Cultural Organization (UNESCO) World Heritage site and is known as the largest covered traditional market in the world. It serves as the basis for major commercial activities in northwest Iran (Pourmoradian, 2018). Due to the central role of commercial and industrial activities and the availability of job opportunities, large numbers of immigrants from other cities, particularly from villages, have moved to Tabriz over the past three decades (Feizizadeh et al., 2022a). Climate change in the form of intensive droughts has especially impacted agricultural activities over the past decade and has forced many people to leave the villages and small towns in the Lake Urmia region and to immigrate to Tabriz in search of jobs and a better quality of life. These socioeconomic issues have led to the development of enormous informal settlements around Tabriz, which have accordingly triggered some unsustainable urban development effects in the greater Tabriz area. Based on this situation, it can be concluded that the development of this city is unbalanced, resulting in further enhancement of differences in lifestyle and quality of life in Tabriz (Feizizadeh et al., 2021a; Feizizadeh et al., 2021b).

3. Dataset and methodology

3.1. Data acquisition and preparation

In the present study, we employed several categories of datasets to carry out the analysis of the trade patterns. The POS transaction dataset of the local banks was obtained from the central bank in Tabriz. It includes the transactions of 5200 stores in the city for the year 2021. These POS locations and their transaction values, together with information on the category of the related business, were employed to analyze and carry out the spatial trade patterns. We employed the postal code of the stores and POS locations to create spatial GIS point data. Accordingly, sums of monthly and annual transactions for each POS were obtained for stores, shopping and services centres, as recorded by the local bank. We analysed the annual numbers and amounts of POS transactions. In addition, to analyse the spatial correlation of the shopping patterns with demographic characteristics, such as population density and age groups, urban land use and traffic data were obtained in the form of GIS datasets from the Municipality of Tabriz. Demographic characteristics (e.g., age, gender, population density) are considered the most important factors when analysing the consumer behaviour of a city (Hassan and Salim, 2015; Wei et al., 2018). Thus, combining demographic data with land use is crucial to understanding purchasing patterns. The demography data were initially published by the Iran Statistical Centre as a result of the national census[1] which are added to the attribute table of urban parcels on the scale of 1/2000 by GIS and the information centre of the Municipality of Tabriz. We also employed the residential population data and the land use map of the city to analyze the spatial correlation of shopping patterns with land use classes. The latest version of the urban land use map in GIS format was also obtained from the Municipality of Tabriz. The land value data for each area of the city were also published y tax department of the city which are also added to the attribute table of the urban parcels in the scale of 1/2000 and we received this data from the Municipality of Tabriz. We employed traffic data to analyse the interaction relation between shopping patterns and traffic volume. Table 1 lists the initial sources of the dataset employed in the current study.

4. Methods

4.1. Big data–driven approaches

Shopping and sales patterns can now be investigated to reveal hidden information and structures of large socioeconomical datasets by overlaying several types of geographic data (Xu and Zhang, 2021). Thus, in following the research objective for applying and integrating the GBD-driven approaches, the methodology was developed based on the well-known GBD approaches including machine learning nearest neighbour, kernel density estimation, space–time pattern mining, and spatiotemporal coupling tele-coupling methods. The selection of these methods was based on the review of research literature (e.g. Zhou et al., 2016; 2020; Yuxuan et al., 2017; Li et al., 2018; Yang et al., 2019; Sharma et al., 2019; Goodchild, 2019; Xu and Zhang, 2021; Xu and Zhang, 2021; Ariza-López et al. 2021. Lu et al. 2021; Mete 2023) as well

[1] please remove all uncited refrences, including this link as well. https://www.amar.org.ir/

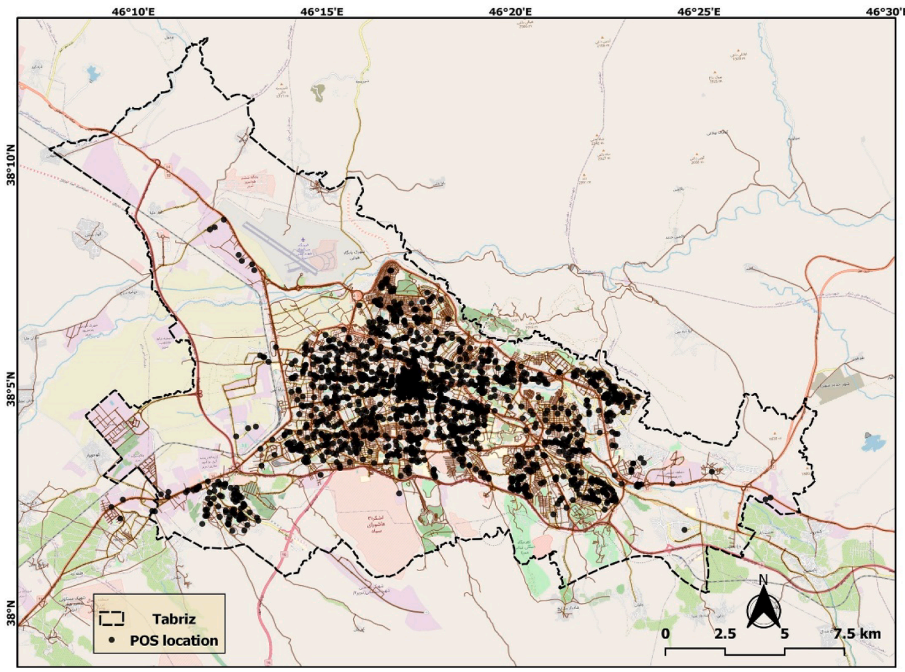

Fig. 1. Study area location and distribution of POS locations in Tabriz.

as the functionality of methods for integrating and developing a comprehensive GBD approach. We also applied the Moran's Index for trend detection of the shopping hotspots. Finally, in order to analyze the impact of socioeconomic factors on shopping patterns, we employed spatial correlation and geomarketing analysis to compute and measure the spatial relationship between the shopping hotspots/cold spots. Fig. 2 shows the major methodological schemes that are applied and integrated as effective GBD-driven approaches for shopping and trade analysis in using the POSs transaction and related socioeconomic dataset.

4.2. Machine learning nearest neighbour

In the first step, we applied the nearest neighbour method to determine the distribution of stores and business/services centres (POS locations). First, the spatial distribution and distance between all sales points (POS locations) were computed. To achieve comparable values, we statistically analysed the geographic distribution of retail centres using the standard deviation ellipse and the machine learning closest neighbour method while sorting the point data into two categories based on the type of business and the time of the transactions. The standard deviation ellipse allowed for delineating the scattering area and identifying the scattering direction for shopping centres and sale points (Taunk et al., 2019; Feizizadeh et al., 2022). We employed the machine learning-based average nearest neighbour (ML-ANN) method to compute the spatial distance between the POS locations.

The efficiency of this approach has been addressed in earlier studies (Taunk et al., 2019; Nazmfar et al., 2020a; Elmezughi et al., 2022; Ghouchan Nezhad Noor Nia et al., 2022). The ML-ANN, p-value and z-score were the results of the average nearest neighbour analysis. The ML-ANN was derived by dividing the observed mean distance by the anticipated mean distance. The expected mean distance, which is the average distance between neighbours in a geographical distribution that is random, reveals the degree of scattering or clustering (Mohammadian et al., 2021). The results of such an analysis indicate average nearest neighbours, along with the ML-ANN, p-values and z-scores. A clustered pattern can be observed in the spatial distribution if the ML-ANN is less than one, while an evenly distributed pattern will be obtained if the ML-ANN is equal to one, and a uniform pattern will be achieved in the dataset if the ML-ANN is computed to be more than one (Yang et al., 2019; Nazmfar et al., 2020a; Mohammadian et al., 2021). Fig. 3 shows the functionality of the ML-ANN analysis in the data used.

As previously discussed, the ML-ANN integrated approach can be efficiently applied for spatial distribution and density assessment. Computationally, ML-ANN operates on the following equations:

$$\mathrm{ANN} = \frac{\overline{\mathrm{D}}_{\mathrm{O}}}{\overline{\mathrm{D}}_{\mathrm{E}}} \tag{1}$$

$$\overline{\mathrm{D}}_{\mathrm{O}} = \frac{\sum_{i=1}^{n}\mathrm{D}_i}{\mathrm{n}} \tag{2}$$

$$\overline{\mathrm{D}}_{\mathrm{E}} = \frac{0.5}{\sqrt{\frac{\mathrm{n}}{\mathrm{A}}}} \tag{3}$$

$$\mathrm{Z} = \frac{\overline{\mathrm{D}}_{\mathrm{O}} - \overline{\mathrm{D}}_{\mathrm{E}}}{\mathrm{SE}} \tag{4}$$

$$\mathrm{SE} = \frac{0.26136}{\sqrt{\frac{\mathrm{n}^2}{\mathrm{A}}}} \tag{5}$$

In our study, $\overline{D}_O$ represents the computed mean distance between POS locations (e.g., shopping centres, malls, stores, etc.) and its nearest neighbour. $\overline{D}_E$ stands for the anticipated mean distance for the POS locations, which results in a random pattern. D_i indicates the distance between feature i and its nearest neighbouring feature, such as shopping centres, malls and stores where the POSs are located. n represents the total number of POSs located in different business classes and categories (Table 1). A also shows the area of a minimum enclosing rectangle around all features such proximity and service area of markets in neighbourhood (e.g. urban parcels) were recorded as POS locations. Finally, Z is the average nearest neighbour z-score

4.3. Kernel density estimation

Interpolation techniques, spatial analysis techniques and mapping cluster techniques are the three primary classes of hotspot mapping approaches (Amiri et al., 2021). In the present study, we aimed to map

Table 1
Dataset employed in this study.

Dataset	Categories	Number of POS data	Year	Scale	Sources
Monthly and annual number and amount of POS transactions which are obtained for each market based on the postal address.		5200	2021	–	The initial data were obtained from the central bank and developed as GIS data based on the address of stores.
Stores and business based on the POSs location	official services	154	2021	–	Municipality of Tabriz and POS attribute
	Grocery store	1474			
	Cosmetics	248			
	Nuts	152			
	Educational Service	180			
	Garment	868			
	Jewelry	378			
	Home appliances	565			
	Car service	231			
	Construction services	390			
	Property service	257			
	Medical Services	293			
Demographic information	Population density (N/km^2) Age groups (25–35, >35–45 and >45) and Gender (m/f/d?)		2016	1/2000	Based on the national census published by statistical center of Iran. This data developed in GIS format by Municipality of Tabriz
Land properties	Land use (5 types…) Land value residential property value, and Commercial property value, based on the information announced by tax department of Tabriz		2021	1/2000	The land value data is computed by tax department for all urban neighborhoods-based m the potential trade value and this data also developed in GIS format for all urban parcel by Municipality of Tabriz.
Traffic volume	Road network and traffic volume (Monthly average)		2022	1/2000	Traffic Police of Tabriz

trade hotspots in the city and examined attributes, locations and linkages in spatial data to extract the required information. Our hotspot analysis utilised kernel density estimation (KDE), line density estimation and point density estimation. KDE has already been applied and found useful for depicting the most intensive regions of sales (Mohammadian et al., 2021). Generally, KDE is a well-known nonparametric method for estimating spatial density (Dudzińska et al., 2020; Cellmer and Trojanek, 2020; Nazmfar et al., 2020b). KDE produces a smooth surface by calculating the severity of sales amounts within an accurate research bandwidth in the study regions. In our case, a kernel function was used to assign a weight based on the distance from the point event to the area surrounding the event. The value then gently falls to zero within the radius of the research circle after reaching its maximum value at the place of the event. In the end, individual kernels in the research region are added to provide a smoothly continuous intensity surface (Anderson, 2009). The density at a definite position is computed using the following equation:

$$f(x,y) = \frac{1}{\mathrm{nh}^2}\sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{k}\left(\frac{\mathrm{d_i}}{\mathrm{h}}\right) \tag{6}$$

where $f(x, y)$ denotes the estimated density at the location (x, y), n denotes the number of observations, h denotes the bandwidth or kernel size, K denotes the kernel function, and di is the separation between the location of the ith observation and the location of the location (x, y). In essence, the KDE technique generates rasters. According to earlier studies, the selection of the bandwidth r is more critical than the selection of the kernel function k (O'Sullivan and Unwin, 2010). Three kernel functions can be used to conduct KDE: the Gaussian function, the quartic function and the minimum variance function (Schabenberger and Gotway, 2017), and we chose the Gaussian function for our study (Equation (7).

$$\mathrm{K}\left(\frac{\mathrm{d_i}}{\mathrm{h}}\right) = \frac{1}{\sqrt{2\mathrm{n}}}\exp\left(-\frac{\mathrm{d_i^2}}{2\mathrm{h}^2}\right), \ \mathrm{When} 0 < \mathrm{d_i} \leq \mathrm{h}$$

$$\mathrm{K}\left(\frac{\mathrm{d_i}}{\mathrm{h}}\right) = 0, \ \mathrm{When}\mathrm{d_i} > \mathrm{h} \tag{7}$$

We employed the KDE method to analyse the spatial variety of shopping and trade centres based on the POS data. To achieve this goal, all business categories were analysed and accordingly classified into 12 subcategories to develop a spatially explicit sales map for each category. These subclasses are shown in Table 1. These business categories were used for evaluating and analysing their spatial patterns throughout the city, and we identified outliers based on the descriptive statistical analysis.

4.4. Space–time pattern mining (STPM)

STPM is a fundamental spatiotemporal analysis technique in GIScience and has been employed in a variety of studies in geography, socioeconomic analysis and urban planning applications (Hamdi et al., 2021; Blazy and Łabuz, 2022). In this regard, STPM is an efficient machine learning method for spatial analysis and pattern mapping for Big data analysis with GIS. STPM clusters the spatiotemporal data and illustrates the data as three-dimensional cuboids, whereby the x and y axes indicate the spatial dimension, and the z-axis represents the time dimension. STPM reveals the spatial pattern in the form of hot and cold spots. It allows for examining increasing or decreasing values in spatial datasets and the severity of contrasts (Sharma et al., 2019). Based on the research aims and objective for applying data driven approaches for GBD, we applied STPM to analyse the shopping and trade patterns in Tabriz as well as. For this goal, the POSs locations and their characteristics were used to develop of trade hotspots clusters in a Python environment.

4.5. Spatiotemporal coupling telecoupling (SPTC)

SPTC is known as one of the most efficient spatial GBD mining methods. It addresses instances that occur in close spatial and temporal proximity. The coupling patterns can be employed for unordered (co-occurrences as cascading patterns) or ordered (sequential patterns) spatial datasets. SPTC also aims to reveal positive or negative correla-

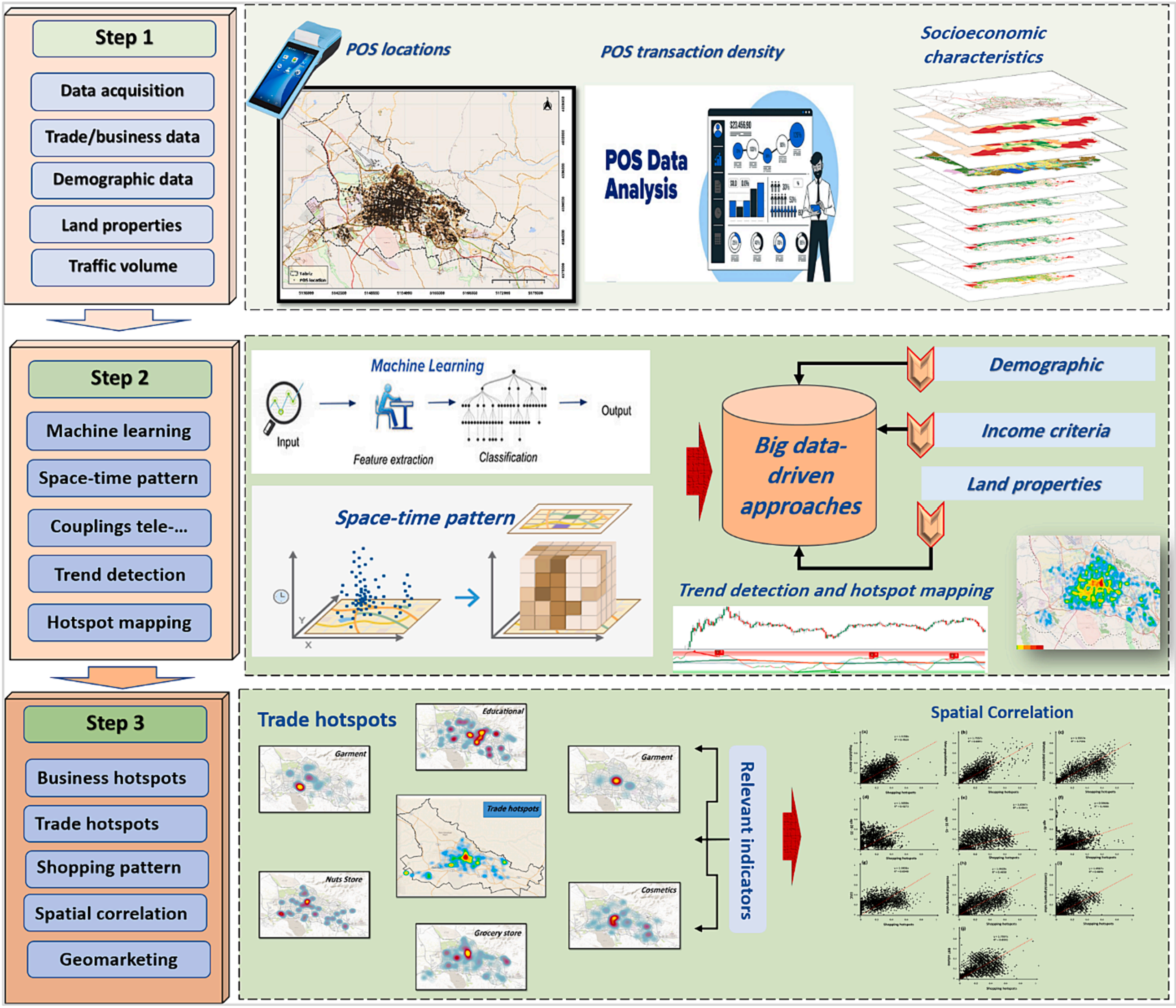

Fig. 2. Main scheme of the research methodology.

tions within a spatial time series (Sharma et al., 2019). The SPTC approach efficiently analyses the spatial context of a point dataset, as demonstrated by previous studies (Chen et al., 2017; Li et al., 2018; Yuxuan et al., 2017; Sharma et al., 2019). Based on the concept of SPTC for addressing different phenomena that are located at separated places but are linked through process (e.g. stores in shopping centres), we aimed to apply the SPTC for trade pattern mapping in shopping centres and particularly Bazaar of Tabriz. As indicated in the case study area statement, the Bazaar of Tabriz acts as a major trade centre for north-west Iran, including intensive trade relations between the bazaar and other cities. The number of POS transactions from all stores and markets was indicated by the variable *T(x)*. In addition, the shaping pattern is represented by *S(y)* as follows:

$$\mathrm{T}(x) = \sum\nolimits_{i=1}^{n} w_i f_i \tag{8}$$

$$\mathrm{S}(x) = \sum\nolimits_{i=1}^{m} w_i f_i \tag{9}$$

In Equations (8) and (9), the weight of the *j*th index in the transaction is represented by the w_i, f_i is the normalised value of the *i*th index in the transaction, and *j* stands for the weight of the *j*th index in the shopping, such as the amount of purchase. *Sj* is also the shopping subsystem's *j*th index's normalised value. The weight in these equations was also computed using the entropy model. In our study, we employed the results of SPTC together with those of STPM to obtain a variety of shopping and trade patterns in the bazzar of city and major shopping centres which high density of stores. The process was competed for integration of SPTC and STPM in the Python environment.

4.6. Trend detection and hotspot mapping

A spatial trend is a consistent change in one or more non-spatial characteristics as an item moves away from it in space. A method for identifying patterns of attribute changes in relation to the vicinity of a geographic object is called spatial trend detection. The spatial auto-correlation approach examines the positions and values of features simultaneously. Whether the data are dispersed, clustered or randomly distributed, this approach returns the pattern that is reflected by the data. As an inferential statistical technique, Moran's I interprets the findings of the analysis according to the null hypothesis (Soltani and Askari, 2014; Afolayan et al., 2022). The null hypothesis takes full spatial randomisation as a given. In other words, values are distributed erratically among features, suggesting an erratically generated spatial process. The goal of this strategy was to determine the statistical

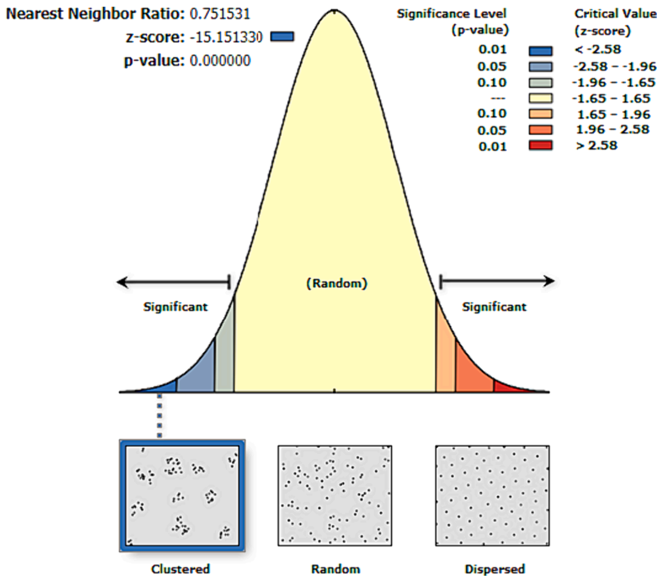

Fig. 3. Average nearest neighbour distribution diagram applied for POS location density assessment (ESRI (https://desktop.arcgis.com/)).

significance of clusters with hot or cold shopping spots because the values linked with the shopping spots were not fairly uniformly distributed across the research area. Moran's I is known as the high or low cluster (Getis Ord. General G). Basically, the general 'G' statistics are more appropriate for creating such a distribution, where the local value spikes are chosen as clusters of high values. However, Moran's I reports the features as clustered, dispersed or randomly distributed and correlates the feature values globally with a fixed distance band or the average nearest neighbour, making it appropriate for evaluating clusters in the dataset (Perumal et al., 2015; Muthu et al., 2022; Afolayan et al., 2022). Moran's I can range from 1 to –1. A value closer to −1 shows that areas have dispersed patterns, while a value closer to 1 shows that areas have similar values (high or low). A number close to zero indicates random patterns. In principle, spatial correlation measures the spatial dependency between the values of random variables in different places (Nazmfar et al., 2020b; Feizizadeh et al., 2022b). The Moran's I spatial autocorrelation index can be calculated as follows:

$$I = \frac{n\sum\sum w_{ij}(x_i - \overline{x})(x_j - \overline{x})}{w\sum(x_i - \overline{x})^2} \tag{10}$$

Where *Xi* is the relative or distance correlation in area units of *I*, N is the number of area units, and *Wij* stands for the weight (Balyani et al., 2014; Nazmfar et al., 2020a). As one of the main objectives of this study, to identify the shopping hotspots, trend detection and hotspot analysis were conducted to determine the main trade hotpots. The results of this step indicate that spatial autocorrelation finds the most neighbouring locations for determining the shopping hotspots in the city.

4.7. Spatial correlation analysis

We performed a spatial correlation analysis to measure correlations between the shopping hotspots/cold spots and the selected impacting criteria in a spatial context (Fig. 4). The selection of impacting criteria was performed based on the expert knowledge of local stakeholders, experts and authorities in Municipality of Tabriz, departments of economic sciences and urban planning in the university of Tabriz as well as a review of the early studies (e.g. Fotheringham et al., 2013; Gray et al., 2020; Berumen et al., 2020; Comber et al., 2002; Feizizadeh et al., 2023; Mete et al. 2023). Due to the variety of scales for the impacting criteria, we standardised the criteria using a fuzzy approach and rescaled all variables to 0–1. We then used GIS to analyse the spatial correlation of shopping patterns and their relationships with various factors (). As we intended to explore the spatial relationship of shopping hotspots against the impacting criteria (Fig. 4), thus we applied the spatial correlation assessment technique. Based on this objective, the geographically weighted regression (GWR) model was applied, to address the relationships between certain sets of variables related to hotspot shopping patterns. GWR is a spatially variable coefficient method that uses a kernel-based strategy to produce a number of local regression models from which the local coefficient estimates of the predictor variable can be retrieved and mapped (Comber et al., 2002). For this goal, we employed population density, male and female population, different age categories (25–35, 35–45 and > 45) and land characteristics, including a land use map, land value for residential and commercial properties, and road network and traffic volume, as relevant indicators for shopping/trade spatial patterns (Fig. 4).

5. Results

One of the main objectives of this research was to address the spatial patterns of sales and shopping and their spatial correlation with the relevant indicators. Based on the POS information, we determined which businesses had the highest sales in different parts of the city. We also identified hot and cold spots in Tabriz. Fig. 5 shows the results when applying Kernel density estimation for the POS density estimation and nearest neighbour analysis. Fig. 6 represents the spatial patterns for the selected business categories, according to Table 1. The business categories of official services, jewellery, grocery stores and home appliances and agencies yielded high levels of transactions, basically in the central part of the city. The major jewellery shops are located in the Bazaar of Tabriz (Bazaar Amir), and high levels of POS transactions in this area could be justified. In terms of home appliances and agencies, several hotspots were recognised in different parts of the city. However, the major hotspot that obtained based on the number and amount of POSs transaction was still recognised in the centre area of the city, which might also mean a high correlation with the land value. For nuts stores and educational activities, the obtained shopping patterns were relatively widely distributed throughout the city. However, the number of transactions was higher in the Valiasr and Elguli subtowns, which are two wealthy towns in Tabriz with high-quality life conditions (Khedmatzadeh and Feizizadeh, 2021; Habibpour et al., 2021).

For garment stores, the results revealed a high number of transactions for Nesfeh-Rah and its vicinity. It has to be indicated that the area of Nesfeh-Rah is famous for its garment trade. These novel markets (e.g., Setarh Baran) were developed over the past few years. However, due to the variety of garment brands and types, this area turned out to be the most favourable area for garment shopping. The central area, particularly the Bazaar of Tabriz, was the second hotspot identified for garments, as it also hosts a large number of garment stores. For cosmetics, the results showed that this business is more equally distributed throughout the city, but three hotspots could be observed. These were again the bazaar and the Valiasr and Elguli subtowns. Analysing the hotspots for medical services yielded the highest POS transaction concentrations in the 17 Shahriver area, which is well known for a variety of health services. As expected, the major hotspots for care services were observed in the suburbs, particularly reflecting care service centres located between Osku and Tabriz along a major transport road. The Valiasr and Elguli subtowns were also found to be hotspots for educational activities. However, there were also many other small hotspots in other parts of the city. In recent years, private schools have received significant attention due to their high reputation compared to public schools. It has recently been observed that an increasing number of families prefer private schools for their children (Nikdel, 2022). This category also includes related activities, such as schools for foreign languages (e.g., English, German), music, arts, information technology, and so on.

Fig. 7 shows the overall and annual trade and shopping hotspots in

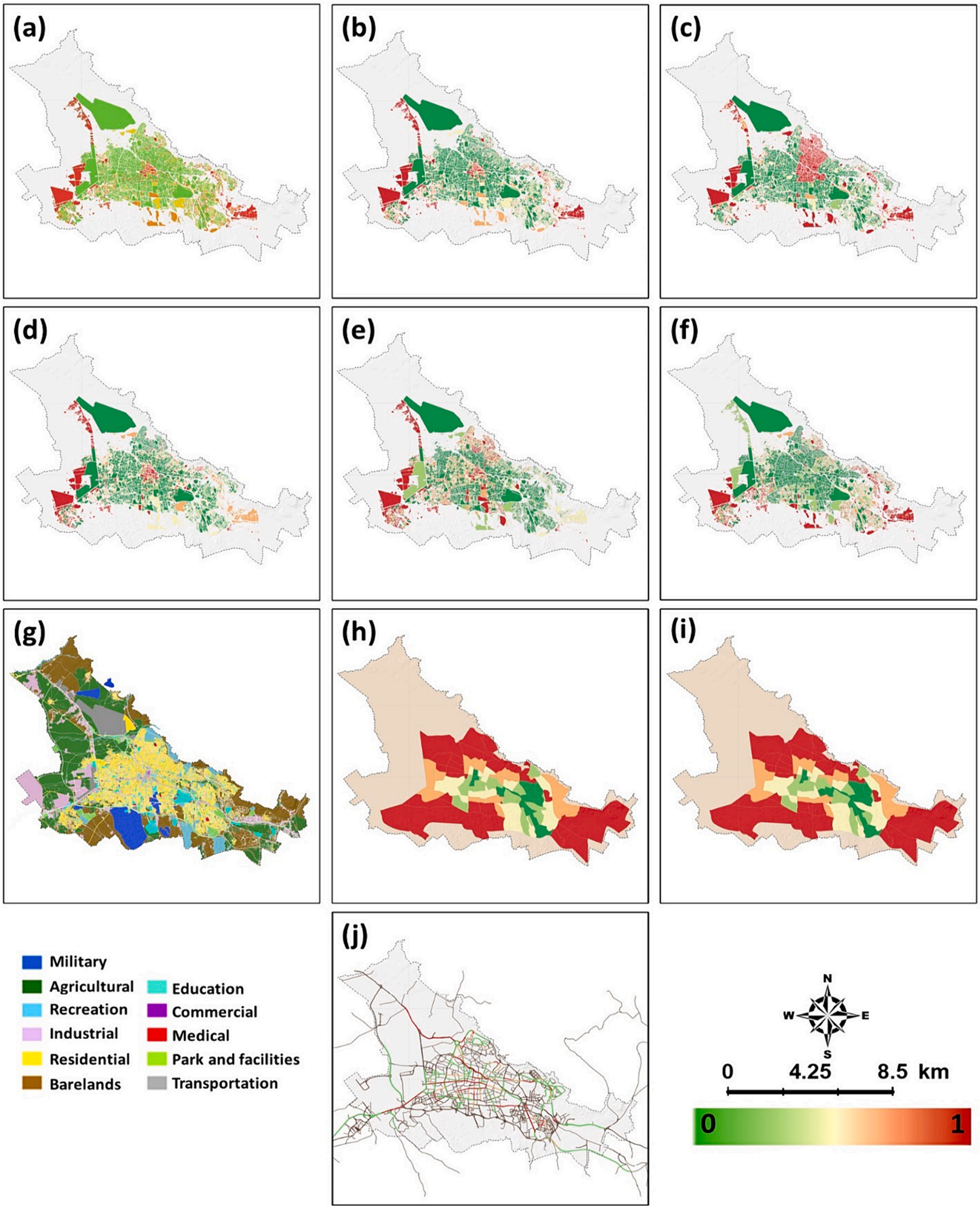

Fig. 4. Standardised spatial distribution of the relevant criteria, including a) overall population density, b) density of male population, c) density of female population, d) density of population aged 25–35, e) density of population aged 35–45, f) density of population aged >45, g) land use map, h) residential property land value, i) commercial property land value and j) road network and traffic volume.

Tabriz for the year 2021, and Fig. 8 enlarges several identified trade hotspots. It becomes visually clear that the highest numbers and values of POS transactions were found in the Bazaar of Tabriz as the major trade centre of the city. It is worth mentioning that this bazaar, which has an area of 500 ha, is believed to be the largest bazaar in the world and is a UNESCO World Heritage site (Pourmoradian, 2018). Fig. 8 also highlights the wealthy subtowns Valiasr and Elguli as major shopping hotspots. As this figure clearly indicates, the central area and several sub city centres are mostly allocated with trade and administrative activities. In addition, several major official and administrative organisations, such as the governor of the East Azerbaijan province, tax offices, central branches of national banks and currency exchange companies,

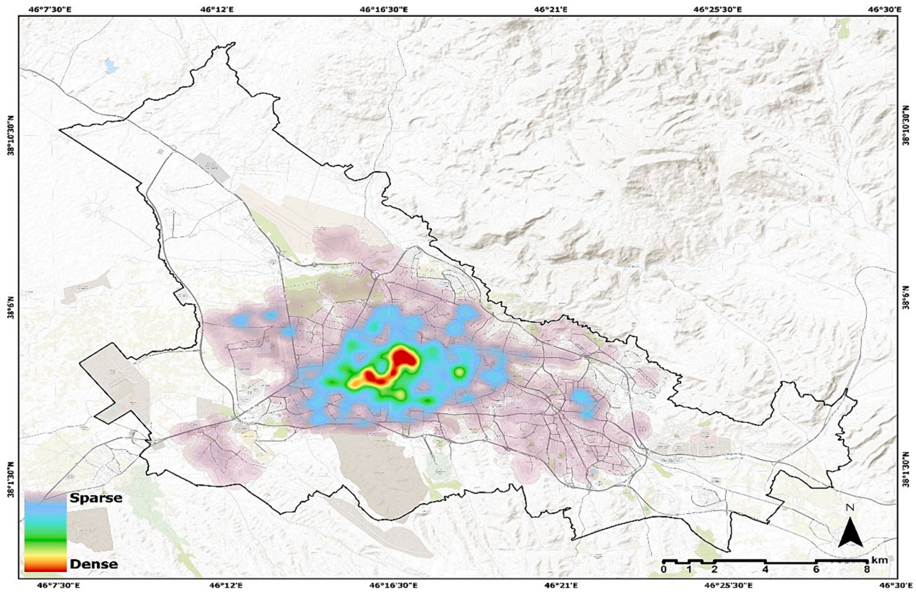

Fig. 5. Results of applying ML-ANN for the POS spatial distribution and density estimation of POS locations.

are located in the central area of Tabriz next to the bazaar, which also causes intensive daily traffic (Feizizadeh et al., 2022b). In addition, due to the trade context, this area also represents very high land values. From an urban planning perspective, most of the major administrative organisations (the city municipality, the governor of the state, etc.) are located in the central part of the city. Fig. 9 shows the spatial correlation of the annual hotspot of shopping and trade patterns and related factors, including demographic characteristics, land properties and traffic volume. As this figure clearly shows, the annual hotspots yield significant correlations with impacting factors, such as population density, land use map and traffic volumes. As Fig. 9h and 9 show, there were very significant correlations between trade and shopping hotspots with commercial and residential land values, which indicates the impacts of property and land value on trade patterns. These trade hotspots have developed over time. Consequently, these well-known areas for trade and shopping have also strongly influenced their land value.

We investigated the spatial correlations of the trade patterns with the relevant indicators (Figs. 4 and 9). The annual trade patterns and population density exhibited significant spatial correlations between the type of business and hotspot areas (Fig. 9a). The Bazaar of Tabriz still acts as a major trade and economic heart of the city. As a UNESCO World Heritage site, it is also a major tourism destination, which also increases the POS transaction figures. The Bazaar of Tabriz is known as the largest international trade centre in northwest Iran, and many products, such as carpets, jewellery, nuts and garments, are being traded here mostly as wholesale. The centralisation of administrative and trade centres in this area contributes to intensive traffic and air pollution (Feizizadeh and Blaschke, 2013). In addition, the bazaar's economy also impacts the land value and increases urban development imbalances. Parallel to this traditional and well-known socioeconomic situation of the central part of the city, this study could demonstrate that modern markets in wealthy sub towns (e.g.Valisar, Roshdieyh and Elguli) have a strong economic effect. These areas have been developing into novel trade hotspots in the city, and our results show that several businesses, such as garment, construction and property services, are not centralised, as we could identify hotspots in different parts of the city. Interestingly, the major trade spots for garments were observed in Nesfeh-Rah Square and its surroundings. Even though the bazaar is still the number one place for many businesses, it is safe to say that the modern and famous markets (e.g., Laleh Park, Roshdeyeh shopping centre, Palladium, Atlas, Tabriz international centre and Milad) in wealthy sub towns have been developing into trade hotspots for several business categories.

6. Discussion

The results of this study confirmed the appropriateness of the GIScience methods used for the spatially explicit analysis of the trade and shopping patterns in Tabriz. The results indicated that the city centre, particularly the Bazaar of Tabriz, acts as the heart of trade for the city, and various hotspots of major business were identified in this area. While this could have been expected, the two other hotspots, the Elguli and Valisar sub towns, were not expected – at least not that they would be so significant. These suburbs are famous for their high property values and high quality of life, as their residents are wealthier. In contrast, the western, northern and southern areas are residential areas interwoven with intensive industrial activities and are generally home to low-income families. From a social and cultural perspective, there is generally a significant difference in lifestyle and quality of life for residents throughout Tabriz (Azami-Aghdash et al., 2019). It should be considered that Tabriz has been facing a critical socioeconomic situation in recent decades as a major immigrant destination due to its industrial and economic attractiveness (Feizizadeh et al. 2023). As a result, about 502 ha of slums and informal settlements and 2500 ha of worn-out urban texture, with a population between 400,000 and 700,000 persons, has developed over the past four decades (Municipality of Tabriz, 2021). These slums and informal settlements face a dualism in lifestyle, urban fabric and morphology (Feizizadeh et al., 2021a). According to our results, the urban stores located in these areas fail to achieve high customer numbers and accordingly have relatively low numbers and amounts of POS transactions. As a result of unbalanced development, the low purchasing power and low quality of life in these low-income areas have prevented the development of modern markets. Nowadays, big data is a buzzword everywhere on the internet, social media as well as communication platforms. Big data has been suggested as a predominant source of innovation, competition and productivity (Manyika et al., 2011), and has caused a paradigm shift to data-driven research (Goodchild, 2013; Kitchin, 2014). In the context of GBD it is widely acknowledged to be a hot topic in GIScience. However, based on the variety of data, challenges of data acquisition, accuracy of dataset and data mining methods, new data driven approaches demanded to be developed and proposed as sate of art. For this goal, early studies addressed a novel methodological framework such as geospatial big data

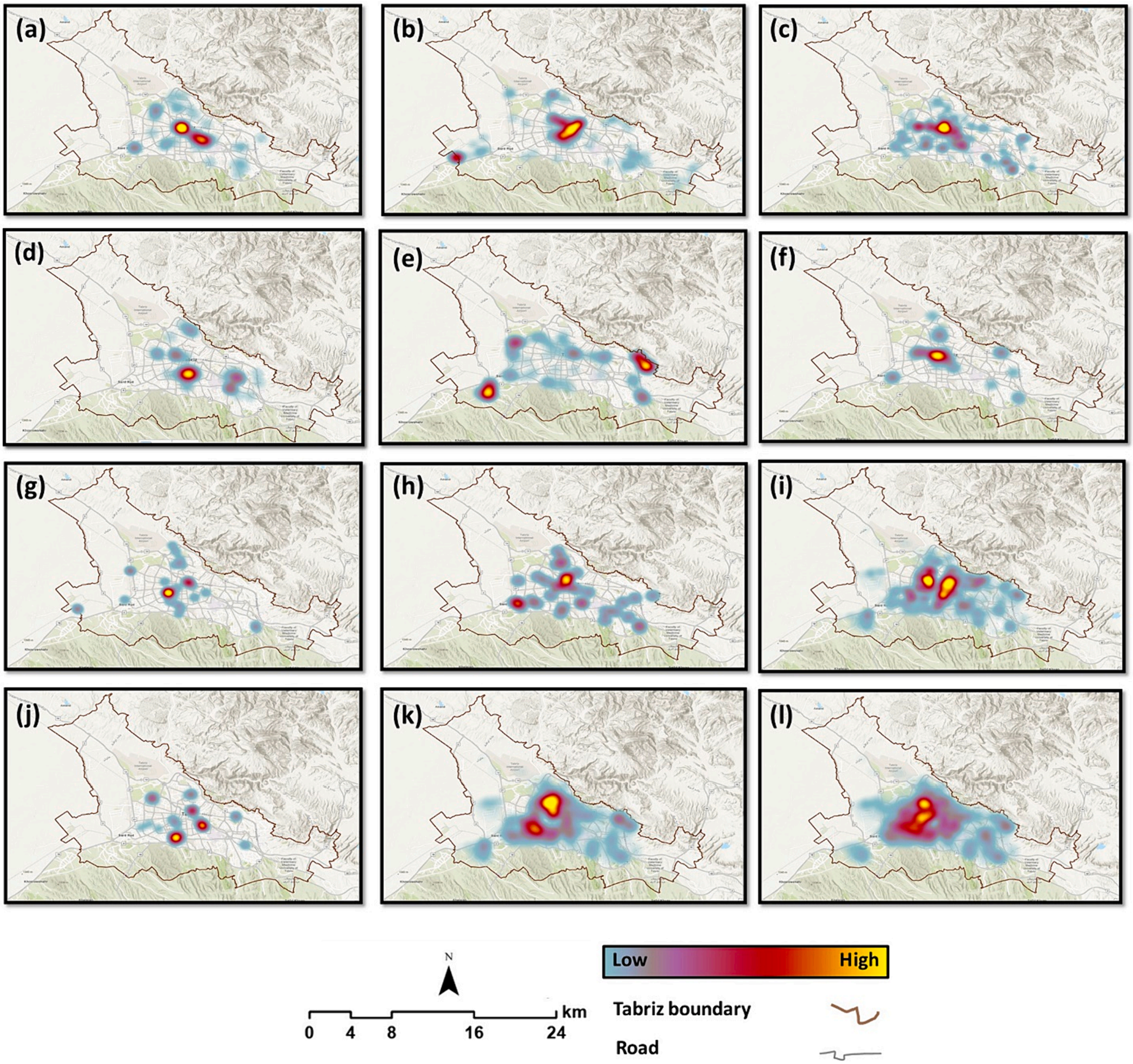

Fig. 6. Hotspot maps of business types in Tabriz: a) official services, b) car services, c) home appliances, d) medical services, e) property services, f) jewellery, g) construction services, h) nuts stores, i) educational services, j) garments, k) grocery stores and l) cosmetics.

management and computer vision algorithms (Nica and Vahancik, 2023), platform for geospatial services and big data (Wang and Jiang, 2021), data quality management strategy (Payani et al., 2022), positional accuracy assessment of geospatial data (Ariza-López et al., 2021), integration of volunteered geographic Information and data from geosensor networks (Li et al., 2021), automatic evaluation of geospatial data quality using web services (Xavier et al., 2019). Within this study we also applied and proposed several GBD data driven GIScience approaches. From a methodological perspective, we can conclude that big data has brought about a paradigm shift in GIScience, bringing both opportunities and challenges (Xu and Zhang, 2021).

From the methodological context, the results of this study indicated that the integration of the proposed four methods can be applied as an effective GBD approach. Results of our study indicated that the integrated approach of these GBD leads to improving the functionality of the approach to overcoming the inherent issue and uncertainty associated with the individual method. For example, the machine learning-based average nearest neighbour (ML-ANN) is a popular data mining approach employed for big data analysis by early studies (Anchalia and Roy, 2014; Neeb and Kurrus, 2016; Deng et al. 2016; Chatzigeorgakidis et al., 2018). Based on our experience in this study, the ML-ANN method might face computational challenges and uncertainty on big data processing while it does not learn from the training set immediately instead it stores the dataset at the time of classification. Basically, it performs an action on the nearest neighbor process as the non-parametric algorithm, which means it does not make any assumption on underlying data. The issue of ML-ANN for big data analysis has also early studies as well as (e.g. Chatzigeorgakidis et al., 2018; Saadatfar et al., 2020). Based on our experience for implementation of, kernel density estimation (KDE), this technique has also turned out to be an effective technique for visualizing data distributions. Technically, the KDE is a robust alternative to the commonly used 'Probability Density Plot' in order to visualize the frequency and spatial pattern of data. The KDE computes data frequency by summing a set of Gaussian distributions, but in contrast to the

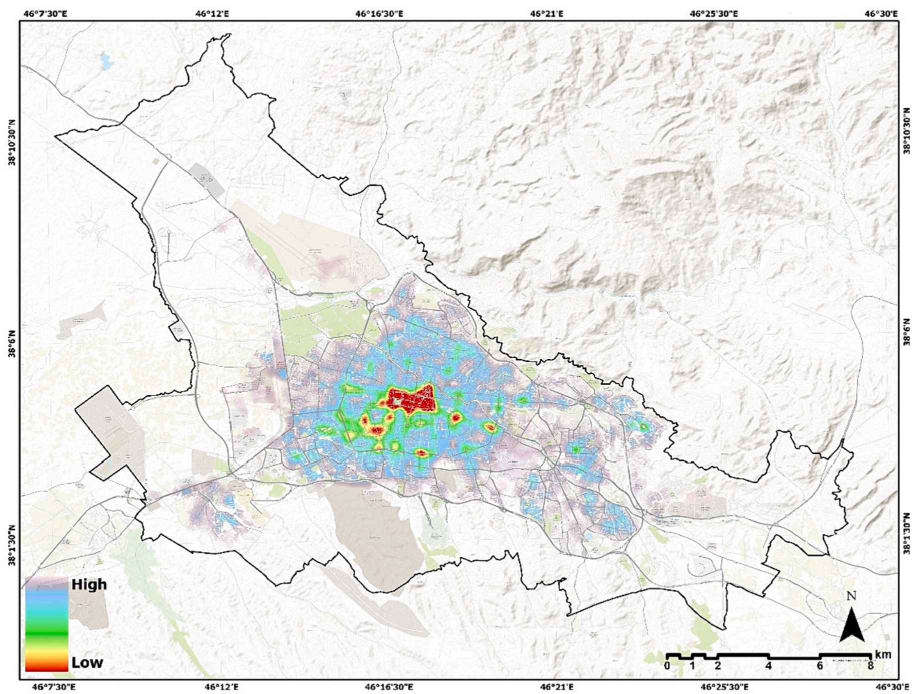

Fig. 7. Results of integrated Geospatial big data driven approaches as overall map of trade hotspots and cold spots in Tabriz.

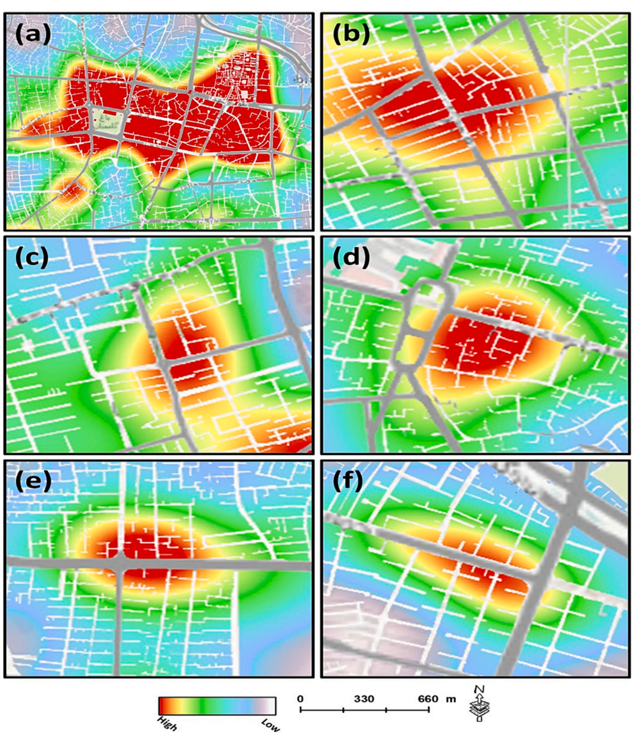

Fig. 8. Enlarged trade hotspots in Tabriz in Fig. 7 to shows more detail of shopping patterns, including a) Bazaar of Tabriz, b) Nesfeh-Rah Square, c) Garagaj Street, d) Valisar sub town, d) Elguli sub town, e) Roshdeyeh town and f) Mansur Square.

'Probability Density Plot', does not take into account the analytical uncertainty (Spencer et al., 2017). This is particularly useful in looking for a cluster of analyses in spectra of data and can be effectively applied to the distribution assessment of the GBD. The space time pattern mining is also proposed as effective GBD data driven approach, which is similar to the time series clustering, under unsupervised machine learning (Sharma et al., 2018; Bennett, 2018; Stular et al., 2022). Spatiotemporal coupling telecoupling technique is also effective data mining approach that can be described as a process of identifying non-trivial, interesting and meaningful patterns from the massive spatiotemporal databases (Shekhar et al., 2015; Das and Ghosh, 2020).

Accordingly, big data developed as a major trend, a foundation to examine, and one that throws new light on GIScience based on the ongoing developments in digital transformation, decentralized, and sensor-based technologies (Dangermond and Goodchild, 2019; Liu et al., 2022). Based on the significance of big data and cyber thinking in the domain of novel GIScience, it is widely expected that new data accusation, mining, and spatiotemporal pattern analysis approaches will be developed to respond to this ongoing and intensive demand. The results of this study verify the efficiency of GIS-based big data approaches for mapping and analysing the spatial context of POS transaction datasets. Such a study would not have been possible a few years ago. Generally, time- and geotagged data are produced at a previously unheard-of rate from many platforms, providing a wealth of options for studying human and environmental dynamics. Spatial data mining can empower a variety of data analysis approaches, including clustering and classification, as well as trend detection in various domains (Perumal et al., 2016). This study also confirms that GIS-based big geo data mining can aid in knowledge discovery in spatial databases and may allow for obtaining implicit knowledge or spatiotemporal patterns from big data in a way that has not been previously possible.

7. Conclusion and future work

Big data encompasses unstructured, semi-structured and structured data. The rapidly growing amount of big data, originating from the many different types of sensors, messaging systems and social networks in addition to more traditional measurement and observation systems, have already invaded many aspects of our everyday existence. The significance of big data and the increase of this novel data in recent years demand the development of new and efficient methods. In this context, the current research aimed to investigate GIS-based GBD analysis of the spatial pattern of trade and shopping. The spatial analysis approach in this article is one way to utilise such data, based on location as the key

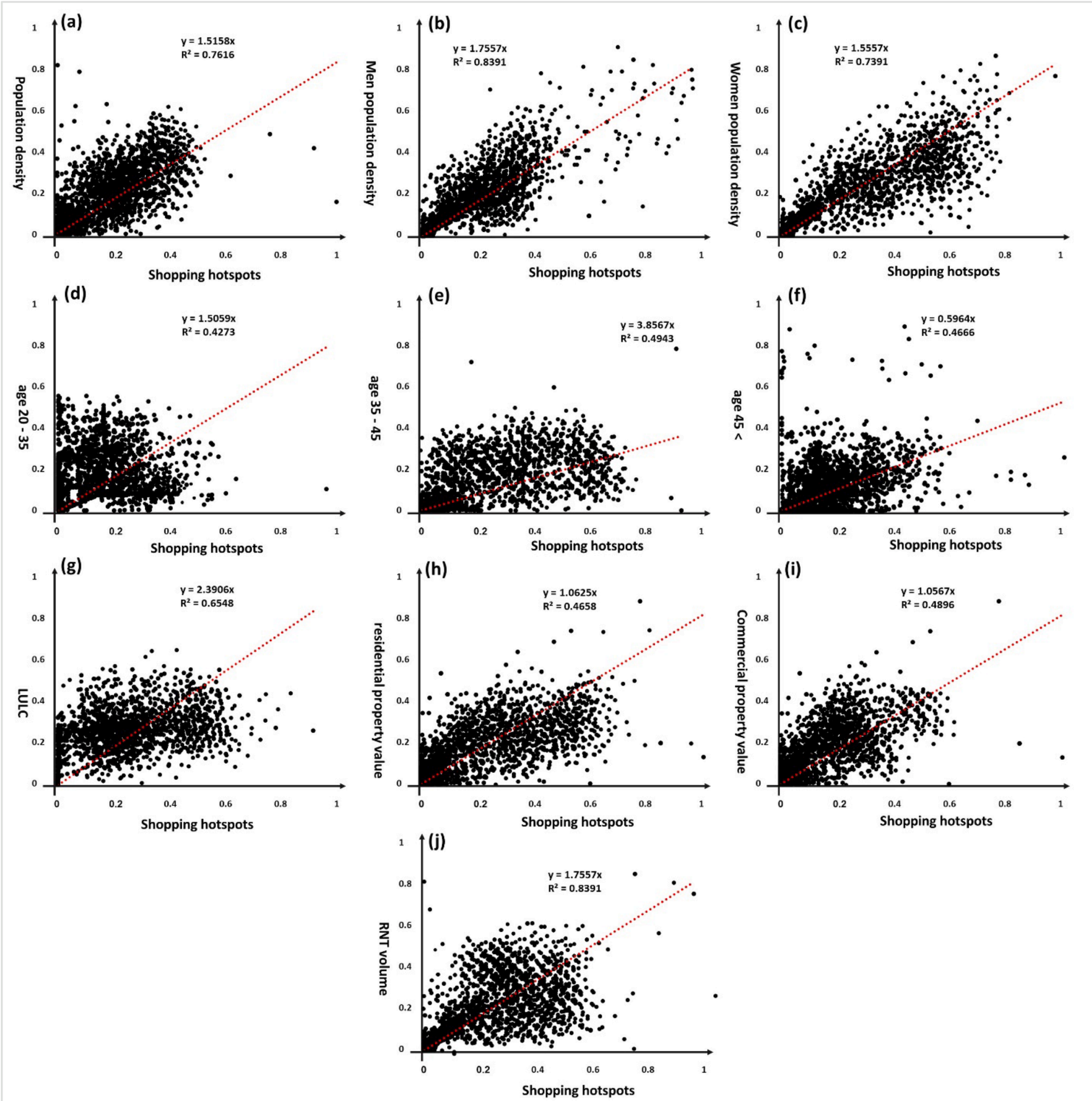

Fig. 9. Spatial correlation of the annual shopping pattern and the relevant impacting criteria computed by the geographically weighted regression method, including a) overall population density, b) density of male population, c) density of female population, d) density of population aged 25–35, e) density of population aged 35–45, f) density of population aged >45, g) land use map, h) residential property land value, i) commercial property land value and j) road network and traffic volume.

variable. We proposed and applied four novel big data-driven methodological frameworks including ML-ANN, KDE, space–time pattern mining and spatiotemporal coupling telecoupling technique for mapping the trade spatial pattern based on the transactions performed by POSs. The results of the study indicated that the big data-driven approach can be customized and applied to similar topics. However, since the GBD approach is still a novel topic in the domain of GIScience, we concluded that further methods and techniques need to be developed and their efficiency to be discussed based on the validation and accuracy assessment methods in future studies.

From an urban planning perspective, the results indicate the impacts of unbalanced developments in Tabriz that significantly affect residents' quality of life. Our results showed that the city centre, particularly the Bazaar of Tabriz, still acts as the city's heart of trade and identified hotspots of major business sectors. Based on the results of this study our future research will focus to apply GBD data driven approaches to time series shopping data in order to analyse and map the spatiotemporal patterns of trade in Tabriz. We also aim to apply different GBD data driven approaches and to compare the efficiency of the methods as part of our future research. The information obtained from this study will benefit local stakeholders, decision makers and authorities in Tabriz in developing strategic plans and programmes for sustainable urban development. The use of GIScience methods allows for investigating trade and shopping patterns in such a way that implicit or hidden information can be revealed. This information has the potential to be used for planning business development and urban development. This research underpinned the importance of inherent spatial patterns in urban environments for location-based services and for the benefit of citizens' quality of life.

CRediT authorship contribution statement

Bakhtiar Feizizadeh: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Validation, Visualization, Writing – original draft, Writing – review & editing. **Davoud Omarzadeh:** Data curation, Resources, Software, Validation. **Thomas Blaschke:** Funding acquisition, Supervision, Visualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Manuscript entitled: A GIS-based big data driven approaches applied for the spatiotemporal mapping of trade and shopping patterns in urban environments, written by Bakhtiar Feizizadeh, Davoud Omarzadeh, Tobia Lakes and Thomas Blaschke. This research is funded by Alexander Von Humboldt Foundation as experienced fellowship which is carried out in Humboldt University of Berlin.

Data availability

Data will be made available on request.

References

Afolayan, A., Easa, S.M., Abiola, O.S., Alayaki, F.M., Folorunso, O., 2022. GIS-Based Spatial Analysis of Accident Hotspots: A Nigerian Case Study. Infrastructures 7, 103. https://doi.org/10.3390/infrastructures7080103.

Alefo, A.L., Cimino, M.G.C.A., Lepri, A., Pentland, A.S., Vaglini, G., 2019. Detecting Permanent and Intermittent Purchase Hotspots via Computational Stigmergy, ICPRAM The 8th International Conference on Pattern Recognition Applications and Methods, doi.org/10.48550/arXiv.1902.01601.

Amiri, A.M., Nadimi, N., Khalifeh, V., Shams, M., 2021. GIS-based crash hotspot identification: a comparison among mapping clusters and spatial analysis techniques. International Journal of Injury Control and Safety Promotion 1–14.

Anchalia, P.P., Roy, K., 2014. The k-Nearest Neighbor Algorithm Using MapReduce Paradigm. In Proceedings of the 2014 5th International Conference on Intelligent Systems, Modelling and Simulation, Langkawi, Malaysia, 27–29 January 2014,513–518.

Anderson, T.K., 2009. Kernel density estimation and K-means clustering to profile road accident hotspots. Accident Analysis & Prevention 41 (3), 359–364.

Ariza-López, F.J., García-Balboa, J.L., Rodríguez-Avi, J., Robledo J., 2021. Guide for the Positional Accuracy Assessment of Geospatial Data. Pan American Institute of Geography andHistory, Occasional Publication # 563.

Azami-Aghdash, S., Gharaee, H., Aghaei, M.H., Derakhshani, N., 2019. Cardiovascular Disease Patient's Quality of Life in Tabriz City in Iran in 2018. Journal of Community Health Research. https://doi.org/10.18502/jchr.v8i4.2080.

Balyani, Y., Hakim Dust, Y., Alijani, B., 2014. Principles of Spatial Data Processing Using Spatial Analysis Methods, publication Azadpeyma.

Bennett, L.,2018. Machine Learning in ArcGIS. ArcUser, the Magazine for Esri Software Users. 2018; 21(2): 8–9. Available from: https://www.esri.com/about/newsroom/arcuser/machine-learning-in-arcgis/.

Berumen, J., Gray, H.L.H., Lovett, S.M., Himmelgreen, D., Biswas, D., Bohn, J., Buro, A. W., 2020. A Mixed-methods Study to Understand Food Environments and Grocery Shopping Patterns of Community Residents in Underserved Neighborhoods in Tampa, Florida. Ecology of Food and Nutrition 1–19.

Blazquez, D., Domenechl, J., 2018. Big Data sources and methods for social and economic analyses. Technological Forecasting and Social Change 130, 99–113.

Blazy, R., Łabuz, R., 2022. Spatial Distribution and Land Development Parameters of Shopping Centers Based on GIS Analysis: A Case Study on Kraków, Poland, Sustainability 14, 7539.

Cellmer, R., Trojanek, R., 2020. Towards Increasing Residential Market Transparency: Mapping Local Housing Prices and Dynamics. ISPRS International Journal of Geo-Information 9 (1), 2.

Chacón-García, J., 2017. Geomarketing Techniques to Locate Retail Companies in Regulated Markets. Australasian Marketing Journal 25 (3), 185–193. https://doi.org/10.1016/j.ausmj.2017.06.001.

Chatzigeorgakidis, G., Karagiorgou, S., Athanasiou, S., Skiadopoulos, S., 2018. FML-kNN: Scalable machine learning on Big Data using k-nearest neighbor joins. Journal of Big Data 5, 4.

Chen, T., Peng, L.,Wang, Q., Liu, S., 2017. Measuring the Coordinated Development of Ecological and Economic Systems in Hengduan Mountain Area. Sustainability, 9, 1270. (CrossRef).

Cliquet, G. (Ed.), 2013. Geomarketing: Methods and Strategies in Spatial Marketing. John Wiley & Sons.

Comber,A., Harris,P., Murakami,D., Tsutsumida, N., Brunsdon; C., 2002. Geographically Varying Coefficient Regression: GWR-Exit and GAM-On?, 15th International Conference on Spatial Information Theory. Editors: Toru Ishikawa, Sara Irina Fabrikant, and Stephan Winter; Article No. 13; pp. 13:1–13:10.

Dangermond, J., Goodchild, M.F., 2019. Building geospatial infrastructure. Geo-Spatial Information Science. 23, 1–9. https://doi.org/10.1080/10095020.2019.1698274.

Das, M., Ghosh, S.K., 2020. Data-driven approaches for spatio-temporal analysis: A survey of the state-of-the-arts. Journal of Computer Science and Technology 35 (3), 665–696.

Deng, Y., Zheng, H., Yan, J., 2023. Applications of Big Data in Economic Information Analysis and Decision-Making under the Background of Wireless Communication Networks. Wireless Communications and Mobile Computing. https://doi.org/10.1155/2022/7084969.

Deng, Z., Zhu, X., Cheng, D., Zong, M., Zhang, S., 2016. Efficient kNN classification algorithm for big data. Neurocomputing 195, 143–148.

Dudzińska, M., Prus, B., Cellmer, R., Bacior, S., Kocur-Bera, K., Klimach, A., Trystuła, A., 2020. The Impact of Flood Risk on the Activity of the Residential Land Market in a Polish Cultural Heritage Town. Sustainability 12 (23), 10098.

Elmezughi, M.K., Salih, O., Afullo, T.J., Duffy, K.J., 2022. Comparative Analysis of Major Machine-LearningBased Path Loss Models for Enclosed Indoor Channels. Sensors, 22, 4967. https://doi.org/10.3390/s22134967.

Feizizadeh, B., Blaschke, T., 2013. Examining Urban Heat Island Relations to Land Use and Air Pollution: Multiple Endmember Spectral Mixture Analysis for Thermal Remote Sensing. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 6 (3), 1–8.

Feizizadeh, B., Ronagh, Z., Pourmoradian, S., Gheshlaghi, H.A., Lakes, T., Blaschke, T., 2021a. An efficient GIS-based approach for sustainability assessment of urban drinking water consumption patterns: A study in Tabriz city. Iran. Sustainable Cities and Society 64, 102584.

Feizizadeh, B., Omarzadeh, D., Ronag, Z., Sharifi, A., Blaschke, T., Lakes, T., 2021b. A scenario-based approach for urban water management in the context of the COVID-19 pandemic and a case study for the Tabriz metropolitan area. Iran, Sciences of Total Environment. https://doi.org/10.1016/j.scitotenv.2021.148272.

Feizizadeh, B., Lakes, T., Omarzadeh, D., Sharifi, A., Blaschke, T., Karmizadeh, S.M., 2022a. Scenario-based analysis of the impacts of lake drying on sustainable food prodection. Nature of Scientific Report. https://doi.org/10.1038/s41598-022-10159-2.

Feizizadeh, B., Omarzadeh, D., Sharifi, A., Rahmani, A., Lakes, T., Blaschke, T., 2022b. A GIS-based spatiotemporal analysis of COVID-19 impacts on the urban traffic accident hotspots and transport network sustainability assessment in Tabriz. Sustanbility. https://doi.org/10.3390/su14127468.

Feizizadeh, B., Omrazadeh, D., Ghasemi, M., Bageri, S., Lakes, T., Kitzmann, R., Ghanbari, A., Blaschke, T., 2023. Urban restaurants and online food delivery during the COVID-19 pandemic: a spatial and socio-demographic analysis. International Journal of Digital Earth 16 (1), 1725–1751. https://doi.org/10.1080/17538947.2023.2210313.

Fotheringham, A.S., Kelly, M.H., Charlton, M., 2013. The demographic impacts of the Irish famine: towards a greater geographical understanding. Transactions of the Institute of British Geographers 38 (2), 221–237.

Ghouchan Nezhad Noor Nia, R., Jalali, M., Houshmand, M. A., 2022. Graph-Based k-Nearest Neighbor Approach for Predicting Phases in High-Entropy Alloys. Appl. Sci. 2, 12, 8021. https://doi.org/10.3390/app12168021.

Gómez, Y.E., Sánchez, J.F.M., Molano, J.I.R., 2020. Geveca Servidor Web Espacial para el GeoMarketing. Revista Ibérica De Sistemas e Tecnologias De Informação E35, 225–233.

Goodchild, M.F., 2013. The quality of big (geo)data. Dialogues in Human Geography 3 (3), 280–284.

Goodchild, M.F., 2022. Commentary: general principles and analytical frameworks in geography and GIScience. Annual of GIS 28, 85–87.

Gray, H., Berumen, J., Lovett, S., Himmelgreen, d., Biswas, D., Bohn, J., Peacock, C., Buro, A., Van Arsdale, W., 2020. Understanding Access to Healthy Foods and Grocery Shopping Patterns Among Community Residents in Underserved Neighborhoods in Tampa, Florida.". Current Developments in Nutrition 4 (2), 194.

Habibpour, F., Feizizadeh, B., Jabarzadeh, Y., 2021. GIS Spatial and Network Analysis Applied for Bookstores Geomarketing. JGST. 10 (3), 99–109.

Hamdi, H., Khaled, S., Abdelkarim, E., Amr, M., Shakila, K.R., Salimm, F., 2021. Spatiotemporal Data Mining: A Survey on Challenges and Open Problems. Artifcial Intelligence Review 55, 1441–1488.

Harding, M., Hersh, J., 2018. Big Data in economics. IZA World of Labor 2018, 451. https://doi.org/10.15185/izawol.451.

Hassan, K., Salim, R., 2015. Population ageing, income growth and CO2 emission: Empirical evidence from high income OECD countries. Journal of Economic Studies 42 (1), 54–67.

Hladík, H., Snopková, D., Lichter, M., Herman, L., Konečn, M., 2021. Spatial-temporal analysis of retail and services using Facebook Places data: a case study in Brno. Czech Republic, Annals of GIS. https://doi.org/10.1080/19475683.2021.1921846.

Kamil, N. M. F. N. M., Razak, K. A., Yatim, F. M., Hidzir, A. H., 2021. GIS-Based Crisis Response System for Electric Distribution Utility: A Case Study of Covid-19

Pandemic In Malaysia. In IOP Conference Series: Materials Science and Engineering, 1051, No. 1, p. 012046, IOP Publishing.
Khedmatzadeh, A., Feizizadeh, B., 2021. Assessing urban quality of life using remote sensing and GIS (Case study: Urmia Urban Region). Journal of Applied Researches in Geographical Sciences 22 (67), 259–279.
Li, S., et al., 2016. Geospatial big data handling theory and methods: A review and research challenges. ISPRS Journal of Photogrammetry and Remote Sensing 115, 119–133.
Li, L., Lei, Y., Wu, S., He, C., Yan, D., 2018. Study on the coordinated development of economy, environment and resource in coal-based areas in Shanxi Province in China: Based on the multi-objective optimization model. Resour. Policy 2018 (55), 80–86.
Lim, W.M., Rasul, T., Kumar, S., Ala, M., 2022. Past, present, and future of customer engagement. Journal of Business Research 140, 439–458.
Lim, W.M.,2023. Transformative marketing in the new normal: A novel practice-scholarly integrative review of business-to-business marketing mix challenges, opportunities, and solutions", Journal of Business Research, 160,113638.
Liu, X., Chen, M., Claramunt, C., et al., 2022. Geographic information science in the era of geospatial big data: A cyberspace perspective. The Innovation 3 (5), 100279.
Lu, G., Batty, M., Strobl, J., Lin, H., Zhu, A.X., Chen, M., 2019. Reflections and speculations on the progress in Geographic Information Systems (GIS): a geographic perspective. Int. J. Geogr. Inf. Sci. 33, 346–367.
Manzan, S., 2023. Big Data and Computational Social Science for Economic Analysis and Policy, Chapter in Handbook edited by: Bertoni, E., Fontana, M., Gabrielli, L., Signorelli, S., Vespe, M., 2023. Handbook of Computational Social Science for Policy, ISBN 978-3-031-16623-5 ISBN 978-3-031-16624-2 (eBook) https://doi.org/10.1007/978-3-031-16624-2.
Martí, P., Serrano-Estrada, L., Social, N.-C.-A., 2019. Media Data: Challenges, Opportunities and Limitations in Urban Studies. Computers, Environment and Urban Systems 74, 161–174.
Mete, A., 2023. Geosptial big data analytics for sustainable smart cities, he International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLVIII-4/W7-2023FOSS4G (Free and Open Source Software for Geospatial) 2023 – Academic Track, 26 June–2 July 2023, Prizren, Kosovo.
Mohammadian, A., Nadimi, A.N., Khalifeh, V., Shams, M., 2021. GIS-based crash hotspot identification: a comparison among mapping clusters and spatial analysis techniques. International Journal of Injury Control and Safety Promotion 27 (26). https://doi.org/10.1080/17457300.2021.1925924.
Muthu, S., Saravanakumar, T.P.A., Ganie, P., Yadav, V., Baghel, P.K., Jeyaraman, M., 2022. Thematic trend mapping and hotspot analysis in bone marrow aspirate concentrate therapy: A scientometric literature analysis and advances in osteoarthritis. Cytotherapy 24 (5), 445–455.
Nazmfar, H., Alavi, A., Feizizadeh, B., Eshgi, A., 2020a. An Integrated Approach of Geographical Information System and Barometer Model for the Spatial Analysis of Drug Crimes in Tehran Parks. The Professional Geographer 72 (1), 1–15.
Nazmfar, H., Alavi, S., Feizizadeh, B., Masodifar, R., Eshghei, A., 2020b. Spatial Analysis of Security and Insecurity in Urban Parks: A Case Study of Tehran. Iran, Professional Geographer 33–0124, 16–21.
Neeb, H.; Kurrus, C., 2016.Distributed K-Nearest Neighbors; Stanford University Publishing: Stanford, CA, USA, 2016.
Nica, E., Vahancik, J., 2023. Geospatial Big Data Management and Computer Vision Algorithms, Remote Sensing and Image Recognition Technologies, and Event Modeling and Forecasting Tools in the Virtual Economy of the Metaverse". Linguistic and Philosophical Investigations 22, 9–25. https://doi.org/10.22381/lpi2220231.
Nikdel, A., 2022. GIS applied for spatiotemporal mapping of the privet school's distribution in Tabriz, master thesis. Department of Remote Sensing and GIS. University of Tabriz.
O'Sullivan, D., Unwin, D.J., 2010. Geographic information analysis. John Wiley and Sons, New York, NY.
Omidipoor, M., Jelokhani-Niaraki, M., Samany, N.N., 2019. A Web-based geo-marketing decision support system for land selection: a case study of Tehran. Iran. Annals of GIS 25 (2), 179–193.
Payani, A.S., Mayanda, A.M., Adiresta, A., Ruldeviyani, Y., 2022. Data quality management strategy to improve remote sensing data quality: A case study on national remote sensing data bank. PTEK the Journal of Technology and Science 33 (3), 2022 (e/pISSN:2088–2033, 0853–4098).
Perumal, M., Velumani, B., Sadhasivam, A., Ramaswamy, K., 2015. Spatial Data Mining Approaches for GIS – A Brief Review. Advances in Intelligent Systems and Computing 338. https://doi.org/10.1007/978-3-319-13731-5_63.
Perumal, M., Velumani, B., Sadhasivam, A., Ramaswamy, K., 2016. Spatial Data Mining Approaches for GIS – A Brief Review. Advances in Intelligent Systems and Computing 338. https://doi.org/10.1007/978-3-319-13731-5_63.
Pourmoradian, S., 2018. Tourism Policymaking Assessment of UNESCO world Heritage Sites in Tabriz, Iran. University of Salzburg, Austria. PhD thesis.
Qiang, Y., Van de Weghe, N., 2019. Re-Arranging Space, Time and Scales in GIS: Alternative Models for Multi-Scale Spatio-Temporal Modeling and Analyses, ISPRS Int. J. Geo-Inf. 2019, 8, 72; doi:10.3390/ijgi8020072.
Roig-Tierno, N., Baviera-Puig, A., Buitrago-Vera, J., Mas-Verdu, F., 2013. The retail site location decision process using GIS and the analytical hierarchy process. Applied Geography 40, 191–198.
Saadatfar, H., Khosravi, S., Joloudari, J.H., Mosavi, A., Shamshirband, S., 2020. A New K-Nearest Neighbors Classifier for Big Data Based on Efficient Data Pruning. Mathematics. 8 (2), 286. https://doi.org/10.3390/math8020286.
Sarsani, M.R., 2011. Socio-Economic Status and Performance on Creativity Tests, Editor (s): Mark A. Runco, Steven R. Pritzker, Encyclopedia of Creativity (Second Edition), Academic Press, 2011, Pages 360-363, ISBN 9780123750389, https://doi.org/10.1016/B978-0-12-375038-9.00179-5.
Schabenberger, O., Gotway, C.A., 2017. Statistical methods for spatial data analysis. CRC Press.
Sharma, A., Jiang,Z., Shekhar.S., 2019. Spatiotemporal Data Mining: A Survey. J. ACM 37, 4, Article 1 (August 2018), 14 pages. https://doi.org/10.1145/1122445.1122456.
Shekhar, S., Jiang, Z., Ali, R.Y., Eftelioglu, E., Tang, X., Gunturi, V., Zhou, X., 2015. Spatiotemporal data mining: A computational perspective. ISPRS International Journal of GeoInformation 4 (4), 2306–2338.
Soltani, A., Askari, S., 2014. Analysis of Intra-Urban Traffic Accidents Using Spatiotemporal Visualization Techniques. Transport and Telecommunication Journal. https://doi.org/10.2478/ttj-2014-0020.
Spencer, C.J., Yakymchuk, C., Ghaznavi, M., 2017. Visualising data distributions with kernel density estimation and reduced chi-squared statistic, Geoscience Frontiers, 8 (6), 1247–1252. Sharma, A., Jiang, Z., Shekhar, S., 2018. Spatiotemporal Data Mining: A Survey., y. J. ACM 37, 4. Article 1, 1–14. https://doi.org/10.1145/1122445.1122456.
Suárez-Vega, R., Santos-Peñate, D.R., Dorta-González, P., 2012. Location models and GIS tools for retail site location. Applied Geography 35 (1–2), 12–22.
Taunk, K., De, S., Verma, S., 2019. Machine learning classification with K-nearest neighbors, 2019 International Conference on Intelligent Computing and Control Systems (ICCS),DOI: 0.1109/ICCS45141.2019.9065747.
Van Ham, M., Tammaru, T., Ubareviciene, R., Janssen, H., 2021. Urban Socio-Economic Segregation and Income Inequality. The Urban Book Series. https://doi.org/10.1007/978-3-030-64569-4_1.
Van Vlasselaer, V., Bravo, C., Caelen, O., Eliassi-Rad, T., Akoglu, L., Snoeck, M., Baesens, B., 2015. APATE: a novel approach for automated credit card transaction fraud detection using network-based extensions. Decis. Support. Syst. 75, 38–48. https://doi.org/10.1016/j.dss.2015.04.013.
Wang, Q., Jiang, Y., 2021. GIS cloud computing based government Big Dataanalysis platform, 021 IEEE 2nd International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE 2021).
Wei, T., Zhu, Q., Glomsrød, S., 2018. How Will Demographic Characteristics of the Labor Force Matter for the Global Economy and Carbon Dioxide Emissions? Ecological Economics 147, 197–207.
Xavier, E.M.A., Ariza-López, F.J., Ureña-Cámara, M.A., 2019. Automatic evaluation of geospatial data quality using web services, Revista Cartográfica 98, 59-73, ISSN (impresa) 0080-2085 ISSN (en línea) 2663-3981, DOI: https://doi.org/10.35424/rcarto.i98.141.
Xiong, T., Wang, S., Mayers, A., Monga, E., 2013. Personal bankruptcy prediction by mining credit card data. Expert Syst. Appl. 40, 665–676. https://doi.org/10.1016/j.eswa.2012.07.072.
Xu, H., Zhang, C., 2021. Development and applications of GIS-based spatial analysis in environmental geochemistry in the Big data era. Environ Geochem Health. https://doi.org/10.1007/s10653-021-01183-8.
Yang, J., Zhang, Z., Wang, B., Li, Q., Yu, Q., Ou, X., Ali, K., 2019. Spatial distribution patterns and intra-specific competition of pine (Pinus yunnanensis) in abandoned farmland under the Sloping Land Conservation Program. Ecological Engineering 135, 17–27.
Yuxuan, L., Zhongyuan, J., Zheng, Y., 2017. Inferring traffic cascading patterns. In Proceedings of the 25th acm sigspatial international conference on advances in geographic information systems. 1-10.
Zhou, C., Su, F., Harvey, F., Xu, J., 2016. Spatial Data Handling in Big Data Era. Springer, Beijing.
Zhou, C., Su, F., Pei, T., et al., 2020. COVID-19: Challenges to GIS with Big Data. Geography and Sustainability. https://doi.org/10.1016/j.geosus.2020.03.005.