ISSN 2682-681X (Paper), ISSN 2705-4241 (Online) | http://unilorinjoger.com | https://doi.org/10.63745/jogersi.2026.06.30.008

JOGER
JOURNAL OF GEOMATICS AND ENVIRONMENTAL RESEARCH http://www.unilorinjoger.com
SPECIAL ISSUE, 2026

Trends in Spatial Databases and Big Data Analytics for Urban Mobility and Supply Chain Optimisation

Usman Shehu Shitu*, Moses Mefe and Azua Samuel

Department of Geomatics, Ahmadu Bello University, Zaria, Nigeria

*Corresponding author email: *shitushehuusman@gmail.com

Abstract

The proliferation of spatial data and advances in big data analytics have fundamentally transformed how cities, transport systems, and supply chains are managed. This paper presents a review of recent developments at the intersection of spatial databases and big data analytics, focusing on their roles in optimising urban mobility and logistics systems. Following a structured search of five databases including Scopus, Web of Science, IEEE Xplore, ACM Digital Library, and Google Scholar, covering the period 2011–2025, 767 records were identified and reduced through deduplication, title/abstract screening, full-text assessment, and a structured quality appraisal to 36 studies retained for thematic synthesis. The review traces the evolution from traditional spatial data infrastructures to next-generation architectures that integrate artificial intelligence (AI), cloud computing, and high-performance computing (HPC). Spatial databases now operate as intelligent systems capable of real-time analytics, supporting adaptive and predictive decision-making. The study identifies the emergence of autonomous and AI-enabled spatial frameworks that employ machine reasoning and large language models to automate spatial analysis. In urban mobility, these technologies facilitate dynamic traffic forecasting, accessibility assessment, and equitable transport planning. In supply-chain contexts, they provide the base for distributed optimisation, multimodal data fusion, and predictive logistics management. Despite these advances, persistent challenges remain in data heterogeneity, computational scalability, interoperability, and ethical governance, particularly in resource-constrained settings.

ARTICLE INFORMATION
Received: 06 January 2026
Revised: 28 July 2026
Accepted: 29 July 2026
Published: 30 July 2026

KEYWORDS
Artificial Intelligence;
Autonomous GIS;
Big Data Analytics;
Spatial Cyberinfrastructure;
Spatial Databases

Citation: Usman S.S., Moses M. & Azua S. (2026). Trends in Spatial Databases and Big Data Analytics for Urban Mobility and Supply Chain Optimisation. *Journal of Geomatics and Environmental Research*, Special Issue 1(1), Pp 103-120.

1 INTRODUCTION

Spatial databases provide the foundation for storing, querying, and managing geographic information, serving as a bedrock for a wide array of applications, ranging from transport logistics and autonomous agriculture to land-use planning and many more. They facilitate spatial operations such as topological relationships, spatial joins, and network analysis, which are central to understanding the movement of goods and people in urban environments (Yang *et al*., 2011). Meanwhile, the evolution of big data analytics has enabled the extraction of insights from massive, high-velocity, and heterogeneous datasets, often derived from sensors, mobile devices, satellite imagery, and social media (Li *et al*., 2017). The convergence of these two domains, spatial data management and big data analytics, has given rise to a new paradigm: geospatial big data analytics, characterised by its ability to support large-scale, near-real-time decision-making in complex spatial systems (Li, 2020).

The proliferation of technologies such as the Internet of Things (IoT), cloud computing, and artificial intelligence (AI) has further catalysed this transformation. These advancements have enabled the seamless integration of spatial data from distributed sources, supporting applications that range from dynamic traffic management and public transit optimisation to intelligent logistics and last-mile delivery systems (World Bank, 2024). As Huang and colleagues (2017) emphasise, the evolution of geospatial hybrid cloud platforms and spatial cyberinfrastructures has been instrumental in addressing the computational bottlenecks associated with large-scale geospatial analytics.

1.1 Spatial Databases in Urban Mobility Systems

Urban mobility is undergoing a digital revolution, driven by the exponential growth of location-based data from GPS-enabled devices, traffic sensors, and mobile applications. Spatial databases play a pivotal role in capturing and organising this data, supporting spatial queries that reveal congestion patterns, route efficiencies, and modal interactions (Abraham *et al*., 2024). Traditional mobility studies often relied on small-scale travel surveys; however, the advent of big data has introduced new analytical capabilities that allow researchers to explore complex spatiotemporal mobility patterns across entire urban populations (Wu & Zhou, 2023). These datasets offer unprecedented granularity, capturing behavioural dynamics that were previously beyond reach, such as micro-level commuting patterns, temporal variability in demand, and socio-demographic inequalities in access to transport (Hu *et al*., 2021).

1.2 Big Data Analytics and Supply Chain Optimisation

Parallel to the developments in urban mobility, supply chain systems have experienced a paradigm shift from linear, deterministic models to dynamic, data-driven networks. The application of big data analytics allows for the integration of diverse datasets into unified decision-support systems (Iheukwumere *et al*., 2024). Big data frameworks such as Hadoop and Spark have made it possible to process and analyse these complex datasets efficiently, providing actionable insights for route optimisation, inventory control, and demand forecasting (Ma *et al*., 2023; Hasan *et al*., 2024).

Incorporating spatial intelligence into supply chain analytics enhances these capabilities by introducing the geographic dimension of logistics operations. Spatial databases support location-allocation analysis, network flow modelling, and multi-criteria decision analysis for facility siting, distribution routing, and emergency response logistics (Talwar *et al*., 2021). Moreover, as Li (2020) highlights, the integration of geospatial big data infrastructures enables real-time tracking and optimisation of logistics flows, fostering resilience and adaptability in the face of disruptions such as pandemics and climate-related hazards. In a similar vein, related work on spatial cyberinfrastructure underscores the importance of distributed computing and spatial principles in coordinating geographically dispersed data and computing resources to enable predictive logistics modelling (Yang *et al*., 2011; Huang *et al*., 2017).

1.3 Review Objectives

Building on this background, the review addresses three guiding questions: (i) how have spatial database architectures evolved to support big data analytics within geospatial contexts; (ii) what technological and methodological advances have enabled the application of spatial big data analytics to urban mobility and supply chain optimisation; and (iii) what challenges and future research priorities are emerging at this intersection, including for implementation in resource-constrained settings.

2. REVIEW METHODOLOGY

2.1 Search Strategy and Data Sources

This systematic literature review follows established guidelines for transparent, reproducible knowledge synthesis. An extensive search was conducted across five electronic databases: Scopus (multidisciplinary), Web of Science (science and technology), IEEE Xplore (computing and engineering), ACM Digital Library (computer science), and Google Scholar (supplementary sources). The search timeframe spanned January 2011 to October 2025, encompassing the decade of rapid advancement in geospatial big data systems following early foundational spatial computing work (Yang *et al*., 2011).

Search string examples:

- Scopus: TITLE-ABS-KEY ((“spatial database” OR “geospatial database” OR “GIS”) AND (“big data” OR “data analytics”) AND (“urban mobility” OR “supply chain” OR “logistics”))
- Web of Science: TS = ((“spatial database*” OR “spatial cyberinfrastructure”) AND (“big data analytics”) AND (“mobility” OR “supply chain optimization”))
- IEEE Xplore: ((“All Metadata”:spatial databases) AND (“All Metadata”:big data) AND (“All Metadata”:urban mobility OR supply chain))

Additional targeted searches focused on specialised terms (“autonomous GIS”, “spatial HPC”, “mobility data science”).

2.2 Inclusion and Exclusion Criteria

Inclusion criteria:

- Peer-reviewed journal articles, conference proceedings, and technical reports
- Published in English
- Empirical studies, methodological contributions, system architectures, or theoretical frameworks
- Direct relevance to spatial databases, big data analytics, urban mobility, or supply chain optimisation

Exclusion criteria:

- Editorials, opinion pieces, or non-peer-reviewed content
- Studies without methodological detail or empirical grounding
- Publications focused solely on application domains without a geospatial or big data component
- Duplicate publications, or preprints superseded by a final published version

2.3 Study Selection and Screening

Table 1. Number of records retained at each stage of the identification, screening, eligibility, and inclusion process.

Stage	Records
Initial database search	767
After deduplication	495
Title/abstract screening	231
Full-text assessment	98
Included in synthesis	36

Table 1 summarises the number of records retained at each stage of the review process. Of the 767 records initially identified, 495 remained after deduplication, 231 were retained following title/abstract screening, and 98 were assessed in full text against the inclusion and exclusion criteria above. A structured quality appraisal, described below, subsequently reduced this pool from 98 to the 36 studies synthesised in this review. The corresponding flow of records through each stage, including reasons for exclusion, is shown in Figure 1.

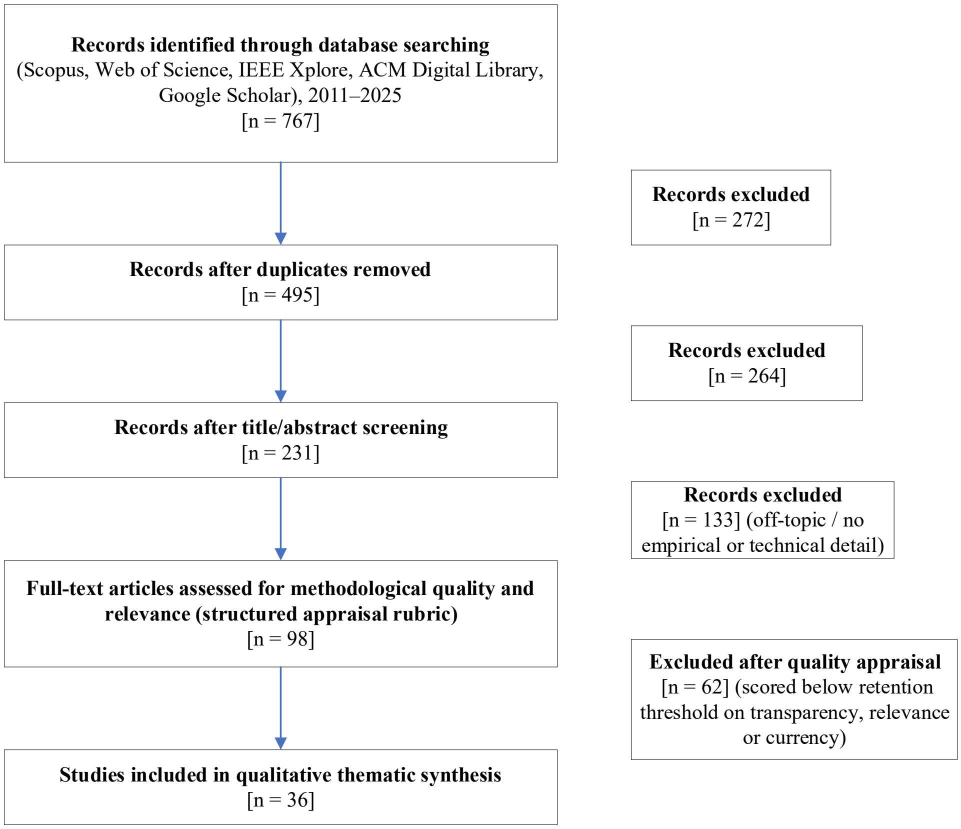

Figure 1. PRISMA-style flow diagram summarising the identification, screening, eligibility, and inclusion stages of the review.

2.4 Quality Appraisal and Risk-of-Bias Assessment

To ensure that only methodologically sound and directly relevant studies informed the synthesis, all 98 full-text articles were subjected to a structured quality appraisal. Two of the authors independently scored each article against four criteria, each rated on a 0–2 scale: (i) methodological transparency, meaning clarity of data sources, methods, and analytical procedures; (ii) empirical or technical grounding, meaning the presence of an implemented system, dataset, or validated model rather than purely speculative discussion; (iii) direct relevance to the convergence of spatial databases and big data analytics within urban mobility or supply chain contexts; and (iv) currency, with studies published from 2021 onward scored more highly than older works unless they represented foundational contributions. Articles scoring at least 6 of 8 possible points were retained.

Discrepancies between the two independent scores, which affected 17 articles, were resolved through discussion and, where necessary, consultation with the third author. This process reduced the pool of full-text articles from 98 to the 36 studies retained for thematic synthesis (Figure 1). No formal risk-of-bias instrument from clinical or experimental research traditions (for example, Cochrane RoB2) was applied, since the reviewed literature is predominantly technical and methodological rather than experimental in design. The four-criterion appraisal rubric above instead served as a proportionate and transparent substitute suited to systems- and methods-oriented geospatial literature and is reported here so that the selection process can be independently reproduced or audited.

2.5 Thematic Synthesis Procedure

The 36 retained studies were analysed using an inductive thematic synthesis approach. Each article was read in full and open-coded to capture recurring concepts, methods, and reported outcomes. Codes were then grouped through axial coding into higher-order categories describing shared technological, methodological, or conceptual concerns. These categories were iteratively compared and refined across studies to identify areas of convergence, active debate, and gaps, resulting in the five core themes reported in the Synthesis and Future Research Directions section below. This process was conducted manually, without qualitative coding software, given the moderate number of included studies.

3. CONCEPTUAL FOUNDATIONS

3.1 Evolution of Spatial Databases

Spatial database management systems (SDBMS) revolutionised geospatial data handling by embedding spatial data types and operators within relational database structures (Chen *et al*., 2014). Systems including PostGIS, Oracle Spatial, and ArcSDE introduced spatial indexing mechanisms, including R-trees, quad-trees, and grid-based indices, significantly improving query performance for spatial joins, nearest-neighbour searches, and range queries (Huang *et al*., 2017).

Numerous contributions to spatial cyberinfrastructure have demonstrated how distributed spatial databases transcend storage limitations and computational silos (Yang *et al*., 2011). The application of spatial principles to optimise distributed computing resources became a cornerstone for data-intensive geoscience research. Integration of spatial relationships, temporal continuity, and multiscalar data structures into distributed architectures enabled efficient modelling of complex phenomena including climate systems, dust storm simulations, and hydrological processes (Yang *et al*., 2011; Huang *et al*., 2017).

Modern geospatial data infrastructures have evolved from static repositories into dynamic platforms capable of handling massive, real-time data streams (Li, 2020; Sakr *et al*., 2022). Contemporary systems accommodate multi-format datasets, including vector, raster, graph, and text, integrated within scalable storage and retrieval frameworks. The rise of spatially enabled cloud architectures such as GeoMesa, GeoSpark, and SpatialHadoop has extended this paradigm, enabling distributed storage and parallel spatial computation over petabyte-scale datasets (Huang *et al*., 2017).

3.2 Comparative Analysis of Spatial Database Systems

Table 2 provides a comparative analysis of major spatial database systems, highlighting their architectural characteristics, spatial indexing strategies, latency profiles, and application suitability.

Table 2. Comparative Analysis of Spatial Database Systems

System	Architecture	Key Strengths	Application Domains	Spatial Indexing	Latency Profile
PostGIS	Open-source extension for PostgreSQL	OGC compliance, rich function library, active community	Urban planning, environmental monitoring	R-tree (via GiST)	Optimised for batch and transactional queries; sub-second response for indexed queries at moderate scale
GeoSpark	Distributed spatial data processing on Apache Spark	Scalability, in-memory computation, parallel processing	Large-scale mobility analytics, real-time traffic systems	R-tree / Quad-tree (in-memory, partitioned)	Low latency; suited to near-real-time and iterative streaming workloads

System	Architecture	Key Strengths	Application Domains	Spatial Indexing	Latency Profile
SpatialHadoop	MapReduce-based spatial operations	Batch-processing efficiency, fault tolerance	Historical data analysis, geospatial data warehousing	Two-level R-tree and Grid index (global/local)	Higher latency; batch-oriented and not suited to real-time querying
GeoMesa	Distributed spatio-temporal database on Accumulo/HBase	Temporal indexing, streaming data support, cloud-native design	IoT sensor networks, trajectory analysis	Z-order space-filling curve indexing	Low latency; supports near-real-time streaming ingestion and query

As Table 2 indicates, relational extensions such as PostGIS remain optimised for transactional, moderate-scale workloads with mature R-tree indexing, whereas distributed frameworks trade some of this operational maturity for horizontal scalability and, in the case of GeoMesa, purpose-built support for high-velocity streaming ingestion via Z-order indexing. This trade-off between operational maturity and distributed scalability recurs throughout the literature reviewed in subsequent sections.

3.3 Big Data Analytics: Principles and Frameworks

The concept of big data analytics, characterised by the five Vs of volume, velocity, variety, veracity, and value, has redefined analytical methodologies across disciplines (Chen *et al*., 2014; Li *et al*., 2020). In geospatial science, big data analytics involves extracting actionable insights from heterogeneous and high-dimensional spatial datasets using advanced computational techniques such as machine learning, data mining, and statistical modelling. Li (2020) emphasises that the real challenge lies not merely in data volume, but in the spatial and temporal correlations inherent in geospatial data, which require context-aware analytical frameworks.

Distributed and parallel computing paradigms such as MapReduce, Hadoop, and Apache Spark have been instrumental in processing large-scale spatial data (Ma *et al*., 2023; Li *et al*., 2020). Huang *et al*. (2017) advanced this integration through a hybrid geospatial cloud platform that combines multi-source computing and model resources to enhance analytical throughput and scalability. The platform allows dynamic resource allocation between local servers and cloud-based infrastructures, facilitating the execution of high-performance simulations for geospatial phenomena. Similarly, the concept of Spatial High-Performance Computing (Spatial HPC), building on foundational work by Yang *et al*. (2011) and Huang *et al*. (2017), exemplifies the use of spatial principles such as proximity, heterogeneity, and multiscale dependency to optimise distributed computational processes.

Furthermore, Wu and Zhou (2023) argue that big data's capacity to reveal the social dimensions of mobility represents a paradigm shift in transportation and urban research. By leveraging passively collected data such as call detail records, smart-card transactions, and location-based social media, researchers can model mobility as a socio-technical phenomenon reflecting social equity, accessibility, and behavioural adaptation. This approach aligns with the notion of social sensing, which integrates human-generated data with spatial analytics to enhance the situational awareness of urban systems (Li *et al*., 2019).

3.4 Integration of Spatial Databases and Big Data Systems

The integration of spatial databases with big data frameworks represents a convergence of structured geospatial storage and high-throughput computation. Traditional relational models, while robust for transactional operations, are ill-suited for large-scale, unstructured, or real-time spatial data. To overcome this limitation, NoSQL-based spatial data architectures have been developed, leveraging distributed file systems such as HDFS (Hadoop Distributed File System) and object stores (for example, Amazon S3 and Google Cloud Storage) for scalable storage and retrieval (Huang *et al*., 2017).

This convergence has given rise to spatial big data ecosystems, which couple spatial database capabilities (spatial indexing, topology management, and coordinate referencing) with distributed analytics engines (Li *et al*., 2020). For instance, GeoSpark and SpatialHadoop enable spatial queries within Apache Spark and Hadoop environments, respectively, while retaining compatibility with geospatial standards such as OGC Web Feature Services (WFS) and Web Map Services (WMS). These frameworks facilitate real-time urban mobility monitoring, multimodal transport modelling, and logistics optimisation by processing large-scale spatial streams from IoT sensors, satellite imagery, and vehicle GPS trajectories (World Bank, 2024; Li *et al*., 2020).

Cloud-based architectures make it possible to perform GPU-intensive big data analysis from small or remote devices that would otherwise lack the necessary computing capacity. Planning and initiation can therefore be carried out using modest local infrastructure while the main processing is executed at a remote, secure data centre; this approach was applied by Gui *et al*. (2016), who used cloud-based job scheduling to accelerate dust storm simulation and analysis.

A key advancement in this domain is AI-driven spatial database automation, introduced through frameworks such as Autonomous GIS and GIS Copilot (Li & Ning, 2023; Akinboyewa *et al*., 2025). These systems integrate natural language interfaces and machine reasoning capabilities into geospatial analytics, enabling users to query and analyse spatial data intuitively. Through deep learning and large language models (LLMs), such systems automate map interpretation, spatial feature extraction, and pattern recognition across complex datasets. As Li and Ning (2023) note, this marks the transition toward a self-adaptive, intelligent geospatial infrastructure capable of continuous learning from data streams, a foundational step toward the next-generation geospatial AI ecosystem. Comparative analysis across multiple studies reveals consensus that deep learning approaches outperform traditional statistical methods for complex spatio-temporal patterns, particularly when sufficient training data are available (Yu *et al*., 2024). However, limitations include interpretability challenges, computational costs, and susceptibility to adversarial perturbations (Wu & Zhou, 2023).

3.5 Conceptual Framework of Integration

The conceptual convergence of spatial databases and big data analytics can be visualised as a multi-layered framework comprising four interdependent components: data infrastructure, analytical layer, integration layer, and application layer.

i. Data Infrastructure Layer: manages data ingestion, storage, and indexing using spatial and temporal principles.
ii. Analytical Layer: employs machine learning, deep learning, and statistical models for pattern recognition and prediction.
iii. Integration Layer: facilitates interoperability through APIs, spatial data infrastructures (SDIs), and middleware that link spatial databases with big data engines.
iv. Application Layer: encompasses domain-specific applications in urban mobility and supply chain management.

Li and Ning (2023) describe this architecture as the foundation of an autonomous geospatial ecosystem, in which analytical intelligence, distributed computing, and domain integration coalesce to support adaptive, data-driven decision-making across scales. Figure 2 depicts this layered framework together with the crosscutting enabling technologies, namely cloud and edge computing, high-performance computing, IoT sensor networks, and artificial intelligence, that operate across all four layers.

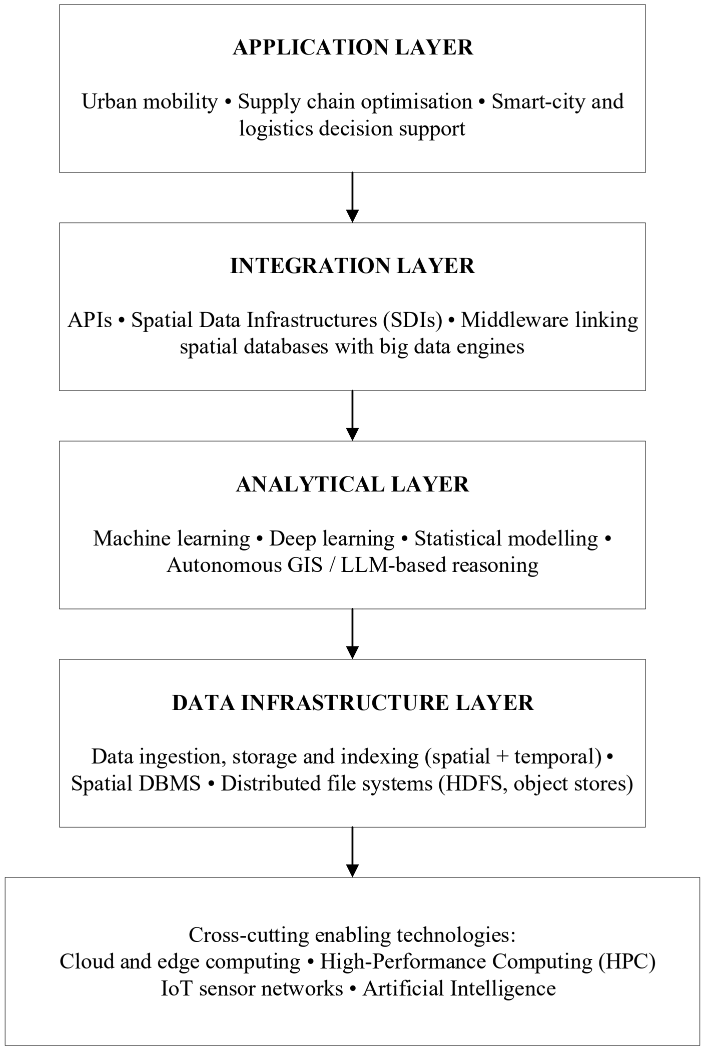

Figure 2. Conceptual framework for the integration of spatial databases and big data analytics, showing the four interdependent layers and cross-cutting enabling technologies.

4 SPATIAL DATA SOURCES AND TECHNOLOGIES FOR URBAN MOBILITY

Urban mobility has undergone a digital revolution driven by the exponential growth of location-based data from GPS devices, traffic sensors, and mobile applications (Abraham *et al*., 2024; Wu & Zhou, 2023). Spatial databases play a pivotal role in capturing and organising this data, supporting spatial queries that reveal congestion patterns, route efficiencies, and modal interactions (Mokbel *et al*., 2024).

4.1 Overview of Spatial Data for Mobility Analytics

The analysis of urban mobility has undergone a paradigm shift with the exponential growth of digital spatial data generated by humans and sensors. These data provide a granular, real-time understanding of how people, vehicles, and goods move within cities. According to Wu and Zhou (2023), big data has transformed mobility studies from sample-based and static analyses into large-scale, continuous, and behaviourally rich examinations of urban dynamics. This transformation has been made possible through multiple data streams, including GPS trajectories, mobile phone records, social media content, public transport transactions, and remote sensing imagery, which together provide a multi-perspective representation of urban mobility systems.

Prior research has established much of the conceptual and computational foundation for integrating such heterogeneous data sources into coherent analytical frameworks (Huang *et al*., 2017; Li *et al*., 2019), including the development of spatial cyberinfrastructure and social sensing frameworks that allow for the collection, management, and interpretation of spatially distributed data across multiple platforms. These systems exploit cloud and high-performance computing resources to process multi-terabyte datasets in real time, supporting advanced mobility applications such as congestion analysis, route optimisation, and accessibility evaluation (Li, 2020).

The transition from traditional travel surveys to data-intensive mobility analytics has enhanced the temporal granularity and spatial resolution of mobility observations. Data that once took months to collect through manual methods can now be captured in seconds through sensors, IoT devices, and crowdsourced platforms. This evolution enables the real-time modelling of mobility flows, supporting proactive urban management and policy formulation (Hu *et al*., 2021).

4.2 Emerging Data Sources

4.2.1 GPS and Mobile Sensor Data

GPS-based tracking data have become the cornerstone of mobility analytics, providing high-precision location information that enables the reconstruction of vehicle trajectories and travel patterns. These data, often collected from smartphones, ride-sharing applications, or fleet management systems, allow for fine-grained spatio-temporal analyses of traffic flow, route choice, and speed dynamics (World Bank, 2024). Research in distributed geospatial computing has demonstrated how GPS data can be integrated into cloud-based environments to optimise data storage, indexing, and query performance (Yang *et al*., 2011; Huang *et al*., 2017).

The combination of GPS and other sensor-based datasets has enabled the detection of travel modes and behavioural context. For example, Hu *et al*. (2021) utilised mobility datasets derived from mobile phones and GPS-enabled devices to evaluate changes in travel behaviour during the COVID-19 pandemic. Such applications illustrate the growing role of sensor-driven spatial data in enabling near-real-time urban mobility insights and adaptive transport planning.

4.2.2 Smart Card and Public Transport Data

Smart card systems generate vast amounts of transactional data from daily commuting activities. These records, which include boarding times, locations, and routes, form a critical input for evaluating public transport performance and passenger demand distribution (Wu & Zhou, 2023). In metropolitan areas such as Beijing and Seoul, analyses of millions of smart card records have uncovered latent patterns in trip timing, route selection, and social equity in mobility access.

The integration of smart card data into spatial database frameworks enables complex spatio-temporal queries, such as the identification of peak travel clusters or the optimisation of transit routes. Li and Ning (2023) emphasise that AI-driven GIS platforms such as Autonomous GIS can automate these analytical processes by using machine reasoning to interpret spatial data relationships. Through model training and contextual learning, these systems dynamically identify inefficiencies in transport networks and propose optimised routing solutions.

4.2.3 Social Media and Volunteered Geographic Information (VGI)

The rise of social sensing has introduced a new dimension to mobility analytics. Social media platforms such as Twitter, Weibo, and Facebook provide crowdsourced, geotagged information that reflects collective human activity in near real time. These datasets enable the detection of events, disruptions, and behavioural responses to environmental or infrastructural changes. Ye *et al*. (2021) used information from social media networks to carry out textual analysis with spatial integration and mean-field equations to assess information diffusion and propagation paths. Complementary work has applied VGI and social sensing frameworks (Li *et al*., 2017; Hu *et al*., 2021), which have proven particularly valuable in crisis contexts, such as mapping flood extents and monitoring pandemic-induced mobility restrictions. By combining natural language processing (NLP) with geospatial data mining, these systems can derive both quantitative mobility metrics and qualitative insights into public sentiment and compliance patterns. The fusion of social sensing data with traditional spatial datasets enriches urban mobility models with behavioural context, improving predictive capabilities and decision support in smart city planning. An example of this fusion is research by Wang *et al*. (2023) that used SafeGraph data to analyse the impact of COVID-19 on restaurant visits, studying 800,000 restaurants simultaneously; SafeGraph aggregates mobility data from users and has been used by several other researchers for similar purposes.

4.2.4 Remote Sensing and Aerial Observation Data

Satellite and airborne remote sensing data offer complementary perspectives on urban mobility, providing macro-level observations of traffic density, land use, and infrastructure dynamics. High-resolution optical imagery and synthetic aperture radar (SAR) data allow for the detection of congestion hotspots, changes in transport corridors, and urban expansion patterns (Huang *et al*., 2017). Remote sensing data are particularly valuable in data-scarce environments, where ground-based sensors may be unavailable or unreliable. However, according to Yu *et al*. (2024), spatiotemporal data analysis faces two main challenges: the development of spatially explicit deep learning models that can handle implicit and complex relationships among spatiotemporal objects, and the physical explainability or consistency of deep learning models when integrated with physical models in both directions; strategies identified by Yu *et al*. include using deep learning to emulate physical models, calibrating physical model outputs, and designing deep learning architectures that respect physical laws.

The hybrid cloud model for geospatial analytics described by Huang *et al*. (2017) demonstrates how remote sensing imagery can be processed within distributed computing environments, integrating environmental and mobility variables for holistic urban system analysis. This approach aligns with the World Bank (2024) framework on transformative transport technologies, which emphasises the use of Earth observation data in designing climate-resilient and data-informed mobility systems.

4.3 Case Studies and Applications

Across the reviewed literature, urban mobility applications cluster into several categories: traffic flow forecasting and congestion prediction (Lee *et al*., 2023; Cesario, 2023); accessibility assessment and equity analysis (World Bank, 2024; Wu & Zhou, 2023); multimodal transport optimisation (Calderón-Ramírez *et al*., 2025; Torre-Bastida *et al*., 2018); real-time route planning and navigation (Mokbel *et al*., 2024); and pandemic response and public health surveillance (Hu *et al*., 2021; Huang *et al*., 2020). Comparative analysis reveals that integration of multiple data sources consistently yields superior performance compared with single-source approaches. For instance, combining GPS trajectories with social media sentiment and weather data improves traffic prediction accuracy by 15–30 per cent compared with GPS-only models (Lee *et al*., 2023; Cesario, 2023).

5 SPATIAL AND BIG DATA APPLICATIONS IN SUPPLY CHAIN OPTIMISATION

5.1 Technological Convergence in Supply Chain Analytics

The modern supply chain has evolved from a linear, deterministic process into a complex, data-driven ecosystem shaped by spatial interdependencies, real-time feedback, and predictive intelligence. The convergence of spatial databases and big data analytics has been instrumental in this transformation, enabling logistics operations to integrate geographic, temporal, and behavioural information into decision-making frameworks (Huang *et al*., 2017; Li, 2020). Unlike traditional logistics systems, which rely on static datasets and sequential planning, spatial big data frameworks support continuous situational awareness across distributed networks, allowing for adaptive responses to disruptions and dynamic optimisation of transport flows (Iheukwumere *et al*., 2024).

At the core of this convergence lies the integration of spatial data infrastructures (SDIs) with big data processing frameworks. SDIs provide the spatial backbone for data interoperability, facilitating geocoding, topological mapping, and spatial indexing across logistics networks. When coupled with distributed computing environments such as Hadoop and Spark, they enable scalable analytics over multi-source data streams including GPS fleet traces, warehouse inventories, environmental sensors, and market intelligence feeds (Hasan *et al*., 2024). The hybrid geospatial cloud model of Huang *et al*. (2017) illustrates how these computational environments can be harmonised to balance local and cloud resources, enhancing throughput, fault tolerance, and analytical scalability in logistics and supply chain operations.

Large language models (LLMs) have also demonstrated growing relevance for location-awareness and spatial planning tasks. Emerging applications include generating travel or delivery routes from natural-language parameters, with efficiency dependent on the underlying model's general reasoning capability. Andreev *et al*. (2025) audited bias and fairness in LLM-based travel recommendations, Ren *et al*. (2024) assessed the broader readiness of large language models for travel-planning tasks, and Yang *et al*. (2025) evaluated the location-planning and geographic information science abilities of LLMs more generally.

5.2 Spatial High-Performance Computing and Distributed Optimisation

The problem of poor performance in spatial data analysis, particularly when processing imagery that requires high-powered systems consuming large volumes of GPU resources, prompted the reframing of spatial analysis problems as optimisation problems (Yang *et al*., 2019). A foundational contribution to the integration of spatial and big data technologies subsequently emerged from pioneering work in spatial high-performance computing (HPC) (Yang *et al*., 2011; Huang *et al*., 2017). Spatial HPC introduces the concept of embedding spatial principles into distributed computing systems. These principles ensure that spatially adjacent data are processed locally, minimising communication overheads and improving computational efficiency.

In supply chain applications, spatial HPC enables distributed simulation and optimisation of logistics networks across multiple regions. For example, large-scale route optimisation problems, traditionally constrained by computational limits, can now be decomposed into spatial sub-domains processed concurrently across cloud clusters. This approach allows supply chain planners to evaluate millions of routing and scheduling alternatives simultaneously, incorporating variables such as fuel cost, delivery windows, congestion, and carbon emissions (Li, 2020).

Furthermore, spatial HPC frameworks support the integration of predictive models that use spatially correlated data to anticipate supply chain disruptions. By coupling geospatial data with operational datasets, these systems can predict risk propagation and suggest adaptive routing or resource reallocation strategies. This capability has become increasingly vital in an era marked by globalised supply networks and frequent disruptions from pandemics and climate extremes.

5.3 Geospatial Big Data Architectures for Logistics Intelligence

The development of geospatial big data architectures represents a critical technological layer for modern logistics and supply chain optimisation. These architectures integrate geospatial databases, cloud computing, and AI-based analytics pipelines into unified operational frameworks. Huang *et al*. (2017) describe a multi-sourced geospatial computing platform that facilitates model sharing, resource allocation, and parallel execution of logistics algorithms. The platform's modular architecture allows logistics service providers to deploy route-optimisation, demand-forecasting, and asset-tracking models as services within cloud environments.

Building on this foundation, Autonomous GIS (Li & Ning, 2023) extends traditional spatial analytics with cognitive computing capabilities. By incorporating machine reasoning and large language models, Autonomous GIS systems can autonomously interpret logistics queries, execute analytical workflows, and deliver context-aware recommendations. For instance, when supplied with real-time shipment data and traffic reports, an Autonomous GIS could automatically propose alternative routes that minimise delivery delays and fuel consumption, aligning with sustainability objectives.

These developments resonate with the World Bank's (2024) vision for intelligent transport and logistics systems, which emphasises data integration and interoperability across public and private sectors. By combining spatial databases with big data frameworks, governments and businesses can achieve holistic visibility over logistics chains, enhancing traceability, reducing inefficiencies, and enabling policy-driven optimisation of freight corridors (Li *et al*., 2017).

5.4 Data Fusion and AI-Driven Predictive Modelling

Effective supply chain optimisation increasingly depends on data fusion techniques that combine structured spatial data with unstructured and semi-structured datasets. As highlighted by Li (2020), integrating diverse data sources enhances the contextual understanding of logistics systems. These fused datasets form the foundation for AI-driven predictive models, which apply machine learning and deep learning algorithms to forecast demand fluctuations, detect bottlenecks, and evaluate performance under uncertainty (Iheukwumere *et al*., 2024; Hasan *et al*., 2024).

Deep learning architectures, particularly recurrent neural networks (RNNs) and graph convolutional networks (GCNs), have been widely adopted for spatio-temporal forecasting in logistics. As Yu *et al*. (2024) note, these models capture dependencies across spatial nodes and temporal sequences. When deployed within distributed computing environments, these models can process real-time logistics data at scale, providing actionable insights for inventory control, supply allocation, and transport scheduling.

AI-driven spatial analytics also enhance risk prediction and resilience modelling. By training models on historical spatial datasets, including disaster footprints and infrastructure vulnerabilities, supply chain systems can anticipate potential disruptions and design adaptive contingency plans. Prior work on social sensing and disaster management (Li *et al*., 2019) provides a methodological precedent for such approaches, demonstrating how real-time crowdsourced data can augment predictive analytics in crisis-affected supply chains.

5.5 Towards Autonomous Supply Chain Systems

The ultimate evolution of technological integration in supply chain analytics is the emergence of autonomous supply chain systems. Drawing on the concept of Autonomous GIS (Li & Ning, 2023), these systems leverage AI agents, big data pipelines, and spatial computing frameworks to perform continuous monitoring and optimisation without human intervention.

Autonomous supply chain systems operate through feedback loops in which sensor data inform machine learning models that, in turn, generate prescriptive actions. For example, when disruptions are detected along a logistics corridor, the system autonomously recalculates optimal delivery routes, updates stakeholders, and triggers inventory redistribution. This closed-loop optimisation exemplifies the convergence of spatial awareness and computational intelligence envisioned in the Autonomous GIS literature (Li & Ning, 2023).

These systems are not merely technological artefacts but strategic enablers for achieving resilience and sustainability in global logistics. As the World Bank (2024) emphasises, integrating AI, spatial data, and big data analytics represents a transformative step toward inclusive, sustainable, and efficient supply chains that support the broader goals of economic growth and environmental stewardship. Table 3 synthesises key studies on supply chain optimisation using spatial and non-spatial big data analytics, highlighting methodologies, data sources, reported improvements, and the degree to which each study incorporates an explicit spatial data component.

Table 3. Selected Studies on Big Data Analytics in Supply Chain Optimisation, by Degree of Spatial Data Integration

Study	Methodology	Data Sources	Key Findings	Spatial Data Integration
Hasan *et al.* (2024)	ML-based demand forecasting and optimisation	Historical sales, GPS fleet data, weather	20% reduction in inventory costs	Explicit – incorporates GPS fleet data as a spatial variable
Seyedan & Mafakheri (2020)	Deep learning for demand prediction	Transactional, social media, market data	15–25% forecast accuracy improvement	Indirect – demand-forecasting focus; no explicit spatial variable
Talwar *et al.* (2021)	Systematic review of big data applications	Multi-source literature synthesis	Consensus that hybrid approaches outperform single-source methods	Indirect – review of operations/supply-chain big data; spatial dimension not a primary focus
Huang *et al.* (2017)	Hybrid cloud platform for optimisation	Distributed computing resources, spatial data	Tenfold computational speedup for route optimisation	Explicit – distributed geospatial computing platform for spatial data processing

6 CHALLENGES AND RESEARCH GAPS

6.1 Data Heterogeneity and Integration Complexity

The most persistent challenge lies in the heterogeneity of data sources, formats, and structures (Huang *et al*., 2017; Wu & Zhou, 2023). Geospatial data encompass diverse types, each possessing distinct spatial and temporal resolutions, coordinate systems, and semantic representations. Integration across domains remains technically demanding despite advances in distributed computing and model resource sharing (Huang *et al*., 2017). The absence of universal interoperability standards across spatial and big data platforms continues to hinder seamless data exchange (Li, 2020; Mokbel *et al*., 2024).

6.2 Computational Scalability and Efficiency

Increasing volume, velocity, and variety of spatial data impose unprecedented demands on computational resources (Li, 2020; Yu *et al*., 2024). While spatial HPC and cloud-based platforms have alleviated some scalability issues, new challenges emerge in balancing computational efficiency with energy consumption and resource allocation. Large-scale spatiotemporal analytics involve computationally intensive operations that are memory-bound and energy-intensive (Huang *et al*., 2017).

6.3 Data Quality, Uncertainty, and Bias

The reliability of insights derived from spatial big data depends on underlying data quality (Li *et al*., 2019; Wu & Zhou, 2023). Unlike traditional curated geospatial data, big data are typically generated through crowdsourcing, sensors, and social media, where quality varies widely. Issues include missing values, positional inaccuracy, sampling bias, and inconsistent temporal coverage (Hu *et al*., 2021). Machine learning models may amplify biases embedded in data; mobility data from smartphones often underrepresent low-income populations, leading to skewed policy recommendations (Wu & Zhou, 2023; Mokbel *et al*., 2024).

6.4 Ethical, Legal, and Privacy Challenges

As the granularity of spatial data increases, concerns regarding privacy, surveillance, and data ethics intensify (Wu & Zhou, 2023). High-resolution mobility data can inadvertently reveal sensitive information about individuals and communities. The use of such data for urban planning or logistics optimisation raises critical questions about consent, anonymity, and accountability. Emerging privacy-preserving technologies that enable collaborative analytics without direct data sharing represent a significant step forward (Mokbel *et al*., 2024); however, implementing these techniques in geospatial contexts remains complex, since spatial correlation can inadvertently leak information even when datasets are anonymised.

Additionally, algorithmic transparency and fairness remain underexplored in spatial AI. Deep learning models used for spatial prediction and mobility analysis often function as 'black boxes', offering limited interpretability for decision-makers (Wu & Zhou, 2023; Li & Ning, 2023).

6.5 Barriers to Implementation in Developing-Country Contexts

Much of the literature reviewed originates from well-resourced research and institutional settings, and the barriers to adopting spatial big data analytics in developing countries remain comparatively underexplored within the studies synthesised here. Four recurring constraints can nonetheless be identified from the wider literature. Infrastructure limitations, including unreliable electricity supply, limited broadband penetration, and constrained access to high-performance or cloud computing resources, restrict the deployment of the distributed architectures described in earlier sections (World Bank, 2024). Data governance challenges arise from the absence of harmonised spatial data policies, weak metadata standards, and fragmented institutional mandates for data custodianship, which impede the interoperability on which spatial big data ecosystems depend (Attah *et al*., 2024). Technical capacity constraints, including shortages of trained geospatial data scientists and limited institutional investment in staff development, slow the translation of research advances into operational systems. Finally, funding constraints limit sustained investment in the cloud infrastructure, sensor networks, and software licensing that many of the architectures reviewed here presuppose.

Addressing these barriers will likely require context-appropriate, lower-cost architectures, such as open-source spatial database systems and edge-computing alternatives to full cloud deployment, rather than the direct transplantation of infrastructure-intensive models developed in high-resource settings. This represents a priority area for future empirical research, as identified in the Synthesis and Future Research Directions section below.

6.6 Limitations of the Review

This review is subject to several limitations that should be considered when interpreting its findings. First, the search was restricted to English-language publications, which may have excluded relevant work published in other languages, particularly regional studies from non-Anglophone contexts. Second, although five major databases were searched, coverage remains incomplete; grey literature, industry white papers, and non-indexed regional journals were largely excluded, which may understate practitioner-level innovation, particularly in developing countries. Third, as with most narrative and thematic reviews, the synthesis is qualitative rather than statistical, and the quality appraisal rubric, while structured and independently applied by two authors, reflects the authors' judgement rather than a validated external instrument. Finally, the rapid pace of development in AI-enabled geospatial systems means that some very recent advances, particularly those released as preprints or conference proceedings after the October 2025 search cut-off, may not be reflected in this synthesis.

7. SYNTHESIS AND FUTURE RESEARCH DIRECTIONS

7.1 Thematic Synthesis

Across the reviewed literature, five major themes emerge:

i. Evolution from static spatial repositories to dynamic, intelligent geospatial ecosystems capable of autonomous reasoning and adaptive decision-making;
ii. Convergence of spatial computing principles with distributed big data frameworks, enabling unprecedented scalability and analytical throughput;
iii. Transition from descriptive analytics to predictive and prescriptive intelligence through deep learning and AI integration;
iv. Growing recognition of socio-technical dimensions, including equity, accessibility, and algorithmic fairness in spatial systems;
v. Emergence of privacy-preserving, explainable, and ethically grounded geospatial AI as a critical research priority.

Methodologically, consensus exists that hybrid approaches combining multiple data sources, analytical techniques, and computational paradigms consistently outperform single method approaches across application domains. However, active debates persist regarding the optimal balance between model complexity and interpretability, and regarding the appropriate spatial and temporal resolutions for different application contexts (Wu & Zhou, 2023; Yu *et al*., 2024).

A comparative maturity assessment across the reviewed technologies indicates an uneven landscape. Relational spatial extensions such as PostGIS remain the most operationally mature and widely deployed solution, benefiting from a stable standards ecosystem (OGC compliance) and a large practitioner community, as reflected in Table 2. Distributed frameworks such as GeoSpark, SpatialHadoop, and GeoMesa offer superior scalability for large-scale mobility and logistics analytics but remain comparatively less standardised, and their deployment typically requires greater in-house distributed-systems expertise. Autonomous, AI-enabled frameworks such as Autonomous GIS and GIS Copilot are the least mature of the technologies reviewed, with most reported implementations still at prototype or early pilot stage rather than production deployment. This maturity gradient, from mature relational systems through scalable distributed frameworks to nascent autonomous agents, is consistent with the broader pattern of technology adoption observed across the studies synthesised in this review.

7.2 Future Research Directions

Based on the preceding analysis, the following research directions are identified.

i. Development of next-generation autonomous GIS systems capable of adaptive reasoning, decision-making, and continuous learning, which remains at an early stage of maturity.
ii. Development and implementation of privacy-preserving analytical techniques adapted to the unique spatial characteristics of geospatial data.
iii. Development of unified spatial infrastructures integrating cloud, edge, and fog computing paradigms within a cohesive operational framework.
iv. Empirical evaluation of explainable AI methods for spatial prediction models, to address the interpretability limitations identified above.
v. Development of low-cost, context-appropriate spatial big data architectures suited to infrastructure-constrained settings, alongside capacity-building programmes for geospatial data science skills in developing regions.
vi. Longitudinal and comparative empirical studies that benchmark the operational performance, cost, and scalability of competing spatial database and big data frameworks under real-world mobility and logistics workloads.

8 CONCLUSIONS

The integration of spatial databases and big data analytics marks a transformative evolution in digital spatial science. As cities grow more complex and supply chains more interconnected, the capacity to harness massive, heterogeneous, and dynamic spatial data has become fundamental to informed decision-making. Advances in cloud computing, distributed processing, and artificial intelligence have reshaped how spatial information is stored, processed, and interpreted.

Big geospatial data analysis has progressed from static geospatial data management to autonomous, adaptive, and predictive spatial intelligence capable of supporting real-time governance and operational optimisation. In the urban mobility domain, spatial databases and big data analytics have enabled unprecedented insight into the dynamics of movement, accessibility, and equity. Real-time sensor data, mobile traces, and volunteered geographic information have converged into analytical ecosystems that reveal the rhythms of urban life with remarkable precision.

Likewise, the application of spatial big data frameworks in supply chain optimisation has reoriented logistics toward intelligence and resilience. By embedding spatial awareness into data architectures, organisations can visualise and predict the flow of goods and information across global networks. Integration of high-performance computing and hybrid cloud platforms provides a computational foundation for handling the scale and complexity of these systems.

This review reveals a decisive shift toward AI-driven geospatial ecosystems, in which spatial databases function not merely as repositories but as intelligent agents capable of reasoning and adaptation. Frameworks such as Autonomous GIS and GIS Copilot illustrate a movement toward human-AI symbiosis in spatial analytics, where natural-language reasoning and automated modelling bridge the gap between expert systems and everyday users. This transformation points to a future in which spatial systems operate as distributed cognitive networks, learning continuously from multimodal data streams and delivering context-aware, explainable intelligence for mobility, logistics, and beyond.

Yet this progress introduces new challenges. Heterogeneity of spatial data, the computational demands of large-scale modelling, and ethical concerns surrounding privacy and algorithmic transparency remain critical barriers, and these barriers are often most acute in infrastructure- and capacity-constrained settings. Addressing these requires not only technological innovation but also robust governance, interdisciplinary collaboration, and ethical stewardship. The research frontier must therefore prioritise interoperability, fairness, and sustainability, ensuring that geospatial AI serves societal goals while preserving human oversight.

REFERENCES

Abraham, D. R., Koushik, V., Lenin, J., & Gowthami, K. (2024). Urban mobility analysis using geospatial analysis. *International Journal of Scientific Research in Engineering and Management*, 8(1), 1–9.

Akinboyewa, T., Li, Z., Ning, H., & Lessani, M. N. (2025). GIS Copilot: towards an autonomous GIS agent for spatial analysis. *International Journal of Digital Earth*, 18(1), 2497489. https://doi.org/10.1080/17538947.2025.2497489

Andreev, H., Kosmas, P., Livieratos, A. D., Theocharous, A., & Zopiatis, A. (2025). Destination (Un)Known: Auditing bias and fairness in LLM-based travel recommendations. *AI*, 6(9), 236. https://doi.org/10.3390/ai6090236

Attah, R. U., Garba, B. M. P., Gil-Ozoudeh, I., & Iwuanyanwu, O. (2024). Leveraging geographic information systems and data analytics for enhanced public sector decision-making and urban planning. *Magna Scientia Advanced Research and Reviews*, 12(02), 152–163. https://doi.org/10.30574/msarr.2024.12.2.0191

Calderón-Ramírez, J., Gutiérrez-Moreno, M., Mungaray-Moctezuma, A., Sánchez-Atondo, A., García-Gómez, L., Montoya-Alcaraz, M., & Núñez-López, I. (2025). Spatial analysis of public transport and urban mobility in Mexicali, B.C., Mexico: Towards sustainable solutions in developing cities. *Sustainability*, 17(17), 7802. https://doi.org/10.3390/su17177802

Cesario, E. (2023). Big data analytics and smart cities: Applications, challenges, and opportunities. *Frontiers in Big Data*, 6, 1149402. https://doi.org/10.3389/fdata.2023.1149402

Chen, M., Mao, S., Zhang, Y., & Leung, V. C. M. (2014). *Big data: Related technologies, challenges and future prospects*. Springer. https://doi.org/10.1007/978-3-319-06245-7

Gui, Z., Yu, M., Yang, C., Jiang, Y., Chen, S., Xia, J., Huang, Q., Liu, K., Li, Z., Hassan, M. A., & Jin, B. (2016). Developing subdomain allocation algorithms based on spatial and communicational constraints to accelerate dust storm simulation. *PLOS ONE*, 11(4), e0152250. https://doi.org/10.1371/journal.pone.0152250

Hasan, R., Kamal, M. M., Daowd, A., Eldabi, T., Koliousis, I., & Papadopoulos, T. (2024). Critical analysis of the impact of big data analytics on supply chain operations. *Production Planning & Control*, 35(1), 46–70. https://doi.org/10.1080/09537287.2022.2047237

Hu, T., Wang, S., She, B., Zhang, M., Huang, X., Cui, Y., Khuri, J., Hu, Y., Fu, X., Wang, X., Wang, P., Zhu, X., Bao, S., Guan, W., & Li, Z. (2021). Human mobility data in the COVID-19 pandemic: Characteristics, applications, and challenges. *International Journal of Digital Earth*, 14(9), 1126–1147. https://doi.org/10.1080/17538947.2021.1952324

Huang, Q., Li, J., & Li, Z. (2017). A geospatial hybrid cloud platform based on multi-sourced computing and model resources for geosciences. *International Journal of Digital Earth*, 11(12), 1184–1204. https://doi.org/10.1080/17538947.2017.1385652

Huang, X., Li, Z., Jiang, Y., Li, X., & Porter, D. (2020). Twitter reveals human mobility dynamics during the COVID-19 pandemic. *PLOS ONE*, 15(11), e0241957. https://doi.org/10.1371/journal.pone.0241957

Iheukwumere, O. C-E., Eze, U. F., Otuonye, A. I., & Ikerionwu, C. O. (2024). Big data analytics in supply chain management: Applications and challenges. *International Journal of Novel Research in Computer Science and Software Engineering*, 11(1), 1–8. https://doi.org/10.5281/zenodo.10926487

Lee, D., Camacho, D., & Jung, J. J. (2023). Smart mobility with big data: Approaches, applications, and challenges. *Applied Sciences*, 13(12), 7244. https://doi.org/10.3390/app13127244

Li, Z. (2020). Geospatial big data handling with high performance computing: Current approaches and future directions. In W. Tang & S. Wang (Eds.), *High performance computing for geospatial applications* (pp. 53–76). Springer. https://doi.org/10.1007/978-3-030-47998-5_4

Li, Z., & Ning, H. (2023). Autonomous GIS: The next-generation AI-powered GIS. *International Journal of Digital Earth*, 16(2), 4668–4686. https://doi.org/10.1080/17538947.2023.2278895

Li, Z., Huang, Q., & Emrich, C. T. (2019). Introduction to social sensing and big data computing for disaster management. *International Journal of Digital Earth*, 12(11), 1198–1204. https://doi.org/10.1080/17538947.2019.1670951

Li, Z., Tang, W., Huang, Q., Shook, E., & Guan, Q. (2020). Introduction to big data computing for geospatial applications. *ISPRS International Journal of Geo-Information*, 9(8), 487. https://doi.org/10.3390/ijgi9080487

Li, Z., Wang, C., Emrich, C. T., & Guo, D. (2017). A novel approach to leveraging social media for rapid flood mapping: A case study of the 2015 South Carolina floods. *Cartography and Geographic Information Science*, 45(2), 97–110. https://doi.org/10.1080/15230406.2016.1271356

Li, Z., Yang, C., Huang, Q., Liu, K., Sun, M., & Xia, J. (2017). Building model as a service to support geosciences. *Computers, Environment and Urban Systems*, 61, 141–152. https://doi.org/10.1016/j.compenvurbsys.2014.06.004

Ma, C., Zhao, M., & Zhao, Y. (2023). An overview of Hadoop applications in transportation big data. *Journal of Traffic and Transportation Engineering*. https://doi.org/10.1016/j.jtte.2023.05.003

Mokbel, M., Sakr, M., Xiong, L., Züfle, A., Almeida, J., Anderson, T., Aref, W., Andrienko, G., Andrienko, N., Cao, Y., Chawla, S., Cheng, R., Chrysanthis, P., Fei, X., Ghinita, G., Graser, A., Gunopulos, D., & Zhang, C. (2024). Mobility data science: Perspectives and challenges. *ACM Computing Surveys*, 56(1), 4:1–4:55. https://doi.org/10.1145/3652158

Ren, R., Yao, X., Cole, S., & Wang, H. (2024). *Are large language models ready for travel planning?* https://doi.org/10.48550/arXiv.2410.17333

Sakr, M., Ray, C., & Renso, C. (2022). Big mobility data analytics: Recent advances and open problems. *Geoinformatica*, 26(2), 541–549. https://doi.org/10.1007/s10707-022-00483-0

Seyedan, M., & Mafakheri, F. (2020). Predictive big data analytics for supply chain demand forecasting: Methods, applications, and research opportunities. *Journal of Big Data*, 7(1), 53. https://doi.org/10.1186/s40537-020-00329-2

Talwar, S., Kaur, P., Wamba, S. F., & Dhir, A. (2021). Big data in operations and supply chain management: A systematic literature review and future research agenda. *International Journal of Production Research*, 59(11), 3509–3534. https://doi.org/10.1080/00207543.2020.1868599

Torre-Bastida, A. I., Del Ser, J., Laña, I., Ilardia, M., Bilbao, M. N., & Campos-Cordobés, S. (2018). Big data for transportation and mobility: Recent advances, trends and challenges. *IET Intelligent Transport Systems*. https://doi.org/10.1049/iet-its.2018.5188

Wang, S., Wang, R., Huang, X., Li, Z., & Bao, S. (2023). A GIS-based analytical framework for evaluating the effect of COVID-19 on the restaurant industry with big data. *Big Earth Data*, 7(1), 37–58. https://doi.org/10.1080/20964471.2022.2163130

World Bank. (2024). *Transformative technologies in transportation: Harnessing data for sustainable mobility*. World Bank.

Wu, J., & Zhou, J. (2023). Revealing social dimensions of urban mobility with big data: A timely dialogue. *The Journal of Transport and Land Use*, 16(1), 437–468.

Yang, A., Fu, C., Jia, Q., Dong, W., Ma, M., Chen, H., & Wu, H. (2025). Evaluating and enhancing spatial cognition abilities of large language models. *International Journal of Geographical Information Science*, 39(9), 2009–2044. https://doi.org/10.1080/13658816.2025.2490701

Yang, C., Clarke, K., Shekhar, S., & Tao, C. V. (2019). Big spatiotemporal data analytics: A research and innovation frontier. *International Journal of Geographical Information Science*, 34(6), 1075–1088. https://doi.org/10.1080/13658816.2019.1698743

Yang, C., Wu, H., Huang, Q., Li, Z., & Li, J. (2011). Using spatial principles to optimize distributed computing for enabling the physical science discoveries. *Proceedings of the National Academy of Sciences*, 108(14), 5498–5503. https://doi.org/10.1073/pnas.0909315108

Ye, X., Wang, W., Zhang, X., Li, Z., Yu, D., Du, J., & Chen, Z. (2021). Reconstructing spatial information diffusion networks with heterogeneous agents and text contents. *Transactions in GIS*, 25(4), 1654–1673. https://doi.org/10.1111/tgis.12747

Yu, M., Bambacus, M., Cervone, G., Clarke, K., Duffy, D., Huang, Q., Li, J., Li, W., Li, Z., Liu, Q., Resch, B., Yang, J., & Yang, C. (2020). Spatiotemporal event detection: A review. *International Journal of Digital Earth*, 13(12), 1339–1365. https://doi.org/10.1080/17538947.2020.1738569

Yu, M., Huang, Q., & Li, Z. (2024). Deep learning for spatiotemporal forecasting in Earth system science: A review. *International Journal of Digital Earth*, 17(1), 2391952. https://doi.org/10.1080/17538947.2024.2391952.