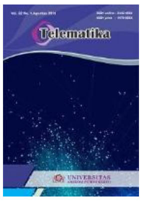
Telematika – Vol. 19, No. 2, August (2026) pp. 156-169 e-ISSN 2442-4528 | p-ISSN 1979-925X



Available online at :
http://ejournal.amikompurwokerto.ac.id/index.php/telematika/

**Telematika**

Accreditated SINTA "2" Kemdiktisaintek, No.10/C/C3/DT.05.00/2025

UNIVERSITAS AMIKOM PURWOKERTO

# Predictive Modeling for ETA and Delivery Delay Prediction in Logistics and Transportation: A Systematic Literature Review

**Syafrial Fachri Pane[1,*], Muhammad Qinthar Sabilla Almaliki[2]**

[1,2]School Information Technology: Applied Bachelor Program of Informatics Engineering, Universitas Logistik dan Bisnis Internasional, Bandung, Indonesia

**ARTICLE INFO**

***History of the article:***
Received February 1, 2026
Revised March 10, 2026
Accepted April 22, 2026

**Keywords:**
Predicting Delivery Delay
Estimated Time of Arrival (ETA)
Logistics & Transportation
Machine & Deep Learning
Systematic Literature Review

**Correspondence(*):**

E-mail:
syafrial.fachri@ulbi.ac.id

**ABSTRACT**

This study presents a Systematic Literature Review (SLR) of data-driven predictive models for delivery delay prediction in logistics and transportation systems. A total of 508 articles were initially retrieved from the Scopus database (IEEE, Elsevier, MDPI, Springer), which served as the primary source of literature, and were systematically evaluated using the PRISMA framework through the identification, screening, eligibility, and inclusion stages, resulting in 55 selected studies published between 2016 and 2026. This review addresses two research questions: (RQ1) how data-driven predictive models, including machine learning and modern statistical approaches, are developed and applied to predict delivery delays; and (RQ2) how these models perform under varying operational conditions. The findings reveal the dominance of machine learning, deep learning, and hybrid models, which leverage heterogeneous data sources such as GPS, AIS, traffic, weather, and IoT data to capture complex spatio-temporal dependencies. Hybrid and deep learning approaches generally demonstrate superior predictive performance in dynamic and nonlinear environments, whereas conventional methods offer advantages in interpretability and computational efficiency. Model performance is strongly influenced by operational conditions, including congestion, weather variability, prediction horizon, and data quality. Inconsistent evaluation metrics limit cross-study comparability, while context-specific datasets reduce model generalizability. Dependence on historical data further constrains adaptability in real-time and disruption-prone environments. This study provides a structured synthesis of predictive modeling approaches, performance trends, and research gaps, offering guidance for researchers and practitioners in selecting and developing delivery delay prediction models. Future research should focus on integrating underutilized contextual and human-related variables, real-time multi-source data, and multi-objective optimization techniques to improve model robustness, scalability, and real-world applicability in intelligent logistics systems.

## 1. INTRODUCTION

In increasingly complex and disruption-prone logistics and transportation systems**,** reliable prediction of delivery delays has become critical for both research and practice (Qiao *et al.*, 2016; Bauer and Tulic, 2018; Noman *et al.*, 2025) . In time-sensitive supply chains, even minor delays can lead to operational inefficiencies, increased costs, and reduced customer satisfaction (Jenelius and Koutsopoulos, 2018; Gabellini *et al.*, 2024; Silvestre *et al.*, 2024). Therefore, accurate estimation of Estimated Time of Arrival (ETA) and early detection of delays are essential to support real-time decision-making and ensure supply chain reliability (Wang, Liang and Delahaye, 2020; Li *et al.*, 2021). This challenge is intensified by increasing system interconnectivity and exposure to disruptions such as traffic congestion, weather variability, and operational uncertainties (Zhang *et al.*, 2018; Alaoua and Karim, 2025).

In this context, studies on Travel Time Prediction (TTP) and ETA provide a key methodological foundation for delay prediction. Despite its importance, delivery delay prediction remains challenging due to the dynamic, uncertain, and non-linear nature of logistics systems (Oh et al., 2018). Traditional model-based approaches are limited in capturing complex interactions among traffic, environmental, and operational factors (Ting *et al.*, 2020; Kwak and Geroliminis, 2021). Although data-driven approaches have gained traction, existing research remains fragmented, with heterogeneous methods, data sources, and application contexts, limiting a comprehensive understanding of predictive performance (Balster *et al.*, 2020; Chen and Fan, 2021). In addition, challenges persist in integrating real-time and multi-source data, and many models lack adaptability to dynamic disruptions (Kim, 2016; Abdelhalim and Zhao, 2025).

Most studies focus on specific transportation modes or localized scenarios, limiting model generalizability across broader logistics systems (Lin, Chen and Chou, 2023; Arbabkhah et al., 2024). The diversity of predictive techniques, including machine learning, deep learning, and ensemble models, combined with heterogeneous data sources such as IoT, AIS, and computer vision, complicates model evaluation and comparison (Yuan, Huang and Zhang, 2020; Li *et al.*, 2021; Huang *et al.*, 2026) . This lack of standardization makes it difficult to assess model effectiveness, scalability, and real-world applicability.

Recent advancements in machine learning and deep learning, including attention-based and spatio-temporal architectures, have improved prediction accuracy and the ability to capture complex patterns (Ting *et al.*, 2020; Jawad-Ur-Rehman, Ul Haq and Muneeb, 2022). The integration of multimodal and real-time data further enhances predictive performance (Li et al., 2017; Alaoua and Karim, 2025). Although several reviews have examined travel time prediction, ETA estimation, and transportation analytics, most focus on specific modes, techniques, or application domains. (Zhang *et al.*, 2020; Wani, 2025).

Previous review studies remain fragmented in scope. Model-based travel-time prediction approaches for highway transportation systems have been reviewed extensively, with a primary focus on traffic simulation and analytical models (Oh *et al.*, 2018b). Machine learning methods for delivery delay prediction have also been reviewed within supply chain contexts, emphasizing prediction algorithms, data sources, and implementation readiness (Müller *et al.*, 2025). In addition, payment delays have been examined from a supply chain coordination perspective, focusing on inventory, cash flow, financial decisions, and supply chain performance (Eddine, Saikouk and Berrado, 2021).

Existing reviews primarily focus on algorithm development and prediction accuracy, with limited synthesis of how model types, data sources, and operational conditions influence predictive performance across logistics and transportation systems. To address this gap, this SLR reviews 55 peer-reviewed studies to examine the development, application, and performance of data-driven predictive models for delivery delay prediction, while identifying research gaps related to data, contextual factors, and model evaluation.

## 2. Recents Studies

Delivery delay prediction is closely related to TTP and ETA, which support travel time estimation and decision-making in logistics systems (Qiao et al., 2016; Jenelius and Koutsopoulos, 2018; Čelan and

Lep, 2020). Early approaches relied on physics-based and mathematical models but were limited in capturing real-world complexity, including data heterogeneity and nonlinear relationships (Kwak and Geroliminis, 2021; Li et al., 2021). With the availability of large-scale and real-time data, research has shifted toward data-driven approaches that improve adaptability and predictive accuracy (Wang, Liang and Delahaye, 2020; Chen and Fan, 2021) . However, variations in models, data sources, and application contexts lead to inconsistent performance and limited comparability across studies (Fan et al., 2018; Ting et al., 2020) While many studies remain context-specific, reducing generalizability (Li et al., 2022; Arbabkhah et al., 2024). Existing research also lacks a systematic linkage between model development and performance under varying operational conditions, resulting in fragmented insights and limited practical applicability. Therefore, a structured synthesis integrating model types, data sources, and operational conditions is required to better understand predictive performance.

### 2.1. Evolution of Predictive Modeling

The development of delivery delay prediction has evolved from traditional model-based approaches to data-driven techniques. Early mathematical and statistical models showed limited adaptability to dynamic and uncertain conditions (Zhang et al., 2018; Cheng, Li and Chen, 2019). With increasing data availability, machine learning approaches were introduced to capture nonlinear relationships and improve predictive performance (Cristóbal et al., 2019; Balster et al., 2020). More recently, deep learning and hybrid models have been adopted to capture complex spatio-temporal dependencies and enhance prediction accuracy (Bauer and Tulic, 2018; Li et al., 2021; Noman et al., 2025). This evolution reflects increasing model complexity, accompanied by challenges in interpretability, scalability, and performance consistency.

### 2.2. Data-Driven Models

Current studies emphasize data-driven approaches that leverage historical and real-time data for prediction in complex logistics environments (Qiao et al., 2016; Zhao et al., 2019; Fan et al., 2023). Machine learning and deep learning models are widely adopted for their ability to capture complex patterns and dependencies (Ting et al., 2020; Sun et al., 2021; Lin, Chen and Chou, 2023). A key trend is the integration of heterogeneous data sources, including GPS, AIS, IoT, weather, and contextual data, to enhance prediction accuracy (Abdelhalim and Zhao, 2025; Alaoua and Karim, 2025).Challenges related to data quality, compatibility, and computational complexity, along with variations in data characteristics and modeling approaches, hinder performance comparability and generalizability across studies (Alessandrini, Mazzarella and Vespe, 2019; Rezki and Mansouri, 2024).

### 2.3. Applications in Logistics and Transportation Systems

Predictive models have been applied across various transportation and logistics domains. In road transportation, they are used for traffic and travel time prediction, while in maritime and logistics contexts, they support ETA estimation and delivery delay prediction (Qiao et al., 2016; Basturk and Cetek, 2021). Furthermore, predictive models are increasingly integrated into supply chain operations to improve planning, reduce uncertainty, and enhance decision-making processes (Li et al., 2022; Gabellini et al., 2024; Silvestre et al., 2024). Despite these applications, many studies focus on specific scenarios or single transportation modes, limiting the generalizability of their findings across different operational contexts

(Cheng, Li and Chen, 2019; Arbabkhah et al., 2024). This indicates that current applications remain fragmented and context-dependent, limiting model transferability across logistics systems.

## 3. RESEARCH METHODS

The study adopts a Systematic Literature Review (SLR) to identify and analyze data-driven predictive models for delivery delay prediction in logistics and transportation systems. To ensure transparency and reproducibility, the review follows the PRISMA framework, which provides a structured process for literature identification, screening, selection, and analysis.

### 3.1. Scope of the Literature Review

The scope of this study focuses on data-driven predictive models for delivery delay prediction in logistics and transportation systems. The review examines how predictive models are developed, applied, and evaluated across different operational conditions. To guide the review, two research questions are defined. The first research question (RQ1) is: How are data-driven predictive models, including machine learning and modern statistical approaches, developed and applied to predict delivery delays in logistics and transportation systems? The second research question (RQ2) is: How do different predictive models perform and compare in terms of effectiveness under various operational conditions? Both research questions are addressed using a systematic literature review approach.

### 3.2. Literature Search Strategy

The literature search strategy was developed using defined keywords expanded with wildcards and combined into a Boolean search string: ("predict*" OR "forecast*" OR "estimated*" OR "model*") AND ("estimated time of arrival" OR "ETA" OR "arrival time" OR "delivery time" OR "transit time" OR "travel time") AND ("delay*" OR "supply chain delay" OR "delivery delay"). The search was applied to titles, abstracts, and keywords, and limited to a defined publication period and relevant subject areas. Only English-language, peer-reviewed journal articles and conference papers were included to ensure quality and relevance. No restrictions were imposed on logistics or transportation modes. Studies covering road transportation, maritime logistics, aviation, public transportation, and urban or last-mile delivery systems were included, provided they addressed delivery delay prediction using data-driven predictive approaches.

### 3.3. Screening and Selection Process

The study selection followed the PRISMA framework for publications from 2016–2026. An initial 508 articles were identified, resulting in 302 articles after duplicate removal and preliminary filtering. Subsequent title, abstract, and full-text screening reduced the dataset to 55 eligible studies for review. These studies were analyzed based on publication year and journal quartiles (Q1–Q3), as presented in Figure 1, while the PRISMA selection process is illustrated in Figure 2.

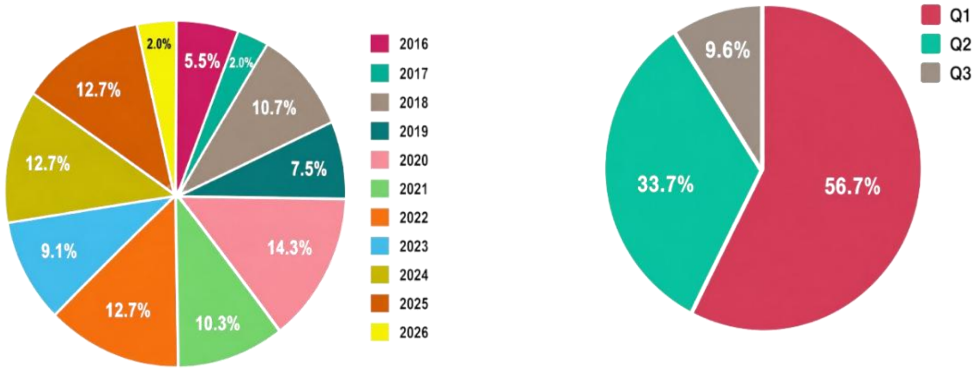



**Figure 1**. Year article classification (a) and Tier journal classification (b)

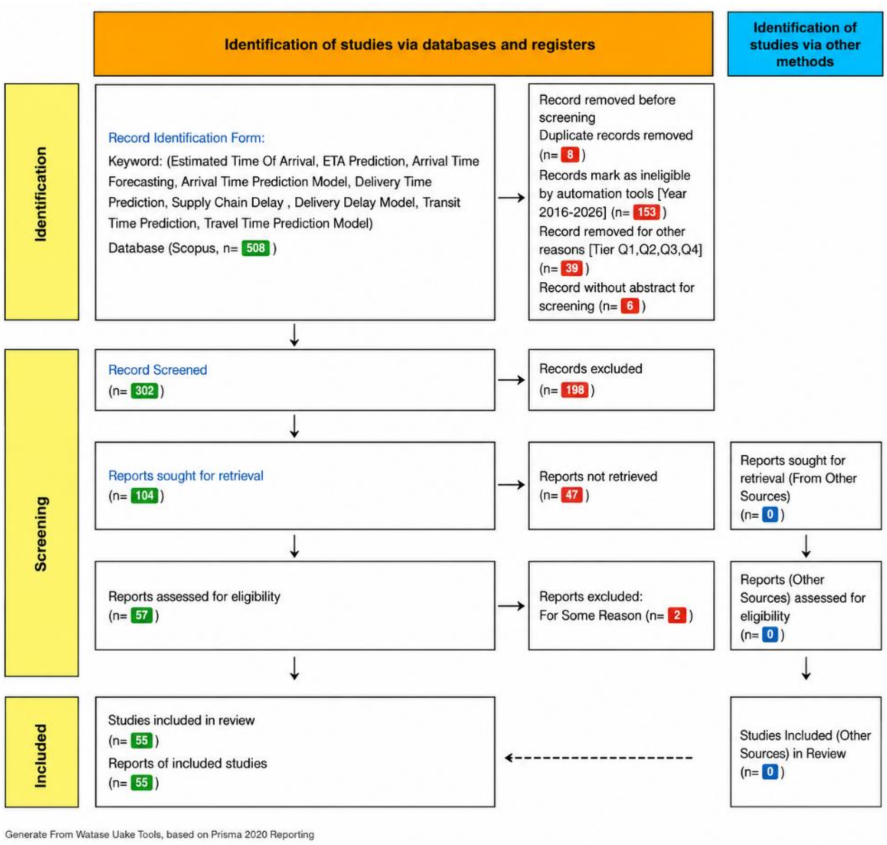



**Figure 2**. PRISMA Flow Diagram for Article Selection Process

### 3.4. Data Extraction and Analysis

Data from the selected studies were systematically extracted and analyzed using a thematic and comparative approach, focusing on predictive models, data sources, performance metrics, and application domains. This analysis enables the identification of research trends, methodological differences, and gaps in the literature. It also supports the investigation of how predictive models are developed and applied (RQ1) and how their performance compares across different operational conditions (RQ2). In addition, word cloud visualizations were used to highlight frequently occurring terms. Figure 3(a) presents the word cloud of extracted variables, while Figure 3(b) shows the word cloud of metadata keywords from Scopus, where term size indicates frequency.

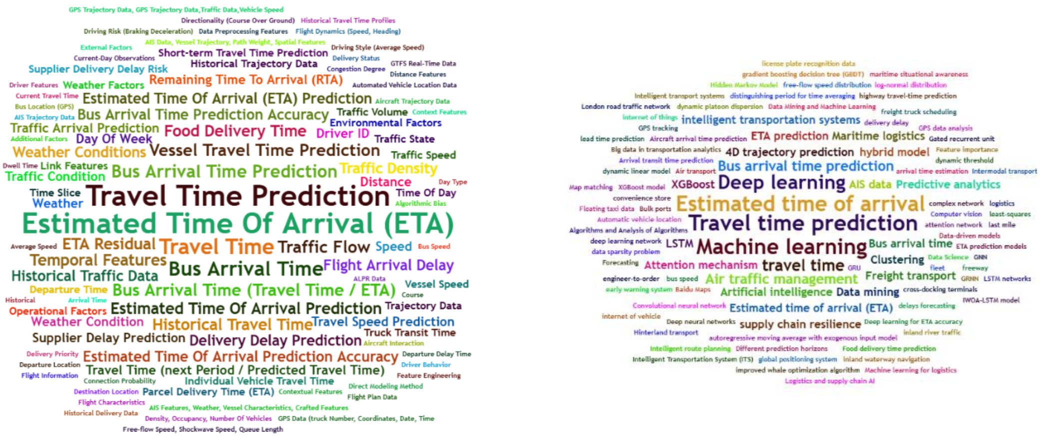



**Figure 3**. Word clouds of extracted variables (a) and Scopus metadata keywords (b)

### 3.5. Quality Assessment (QA)

To ensure the quality and relevance of the selected studies, a quality assessment was conducted after the screening process. Each article was evaluated based on four criteria: (1) relevance to delivery delay prediction, (2) clarity of methodology, (3) adequacy of data and experimental design, and (4) completeness of performance evaluation. Studies that did not provide sufficient methodological details or clear evaluation results were excluded during the eligibility assessment stage. In addition, potential sources of bias were minimized through predefined inclusion and exclusion criteria, the use of a structured search strategy, and the application of the PRISMA framework. The assessment process focused on ensuring that the selected studies provided reliable evidence for addressing the research questions and supporting comparative analysis across different predictive approaches.

## 4. RESULT

The results indicate the dominance of data-driven approaches in delivery delay prediction, particularly deep learning and hybrid models, with travel time prediction and estimated time of arrival (ETA) as the primary application domains (Nguyen and Liem, 2025). The reviewed studies also show a growing publication trend, reflecting increasing interest in data-driven transportation analytics. GPS, AIS, traffic, and weather data emerge as the most frequently utilized data sources, often integrated with contextual and real-time information (El Mekkaoui, Benabbou and Berrado, 2023; Abdelhalim and Zhao, 2025). Model performance varies across model types, data characteristics, and operational conditions, while differences in evaluation metrics limit direct comparability across studies (Yuan, Huang and Zhang, 2020; Wani, 2025). To provide a structured comparison, the selected studies are synthesized based on model type, data source, operational condition, and performance metrics, as summarized in Table 1. Figure 4 presents the percentage distribution of predictive model categories identified across the reviewed studies, highlighting the dominance of deep learning and machine learning approaches.

**Table 1**. Mapping of models, data sources, conditions, and performance across selected studies.

| Source | QA | Prediction Model | Data Source | Experimental Condition | Performance |
|---|---|---|---|---|---|
| Huang et al. (2026) | 4 | Hybrid Ensemble (LightGBM + XGBoost + Random Forest + Linear Regression) | AIS data | Long-distance tramp bulk shipping, port-to-port routing, dynamic maritime conditions | MAE = 3.30 h; 74.7% improvement over baseline; 11.3% improvement over best single model |
| Kaya & Utku Kalay (2025) | 4 | Hybrid (LSTM + GRU + Transformer) | GPS, weather, transit data | Weather conditions, time blocks, peak vs. off-peak | MAE = 2.9; MAPE = 14.79%; $R^2$ = 0.927 |

| | | | | | |
|---|---|---|---|---|---|
| Su et al. (2025) | 3 | Hybrid (GRNN + LSTM + GNN) | Traffic and external-factor data | Traffic conditions and environmental variability | — |
| Noman et al. (2025) | 3 | Hybrid (CNN + LSTM + Transformer + MLP) | AIS and meteorological data | Upstream/downstream navigation, segment length, external factors | — |
| Nguyen & Liem (2025) | 4 | Deep Learning (Attention-based) | Flight trajectory and airspace data | Airspace density, aircraft interaction, queue conditions | MAE improved by 25%; RMSE improved by 20% |
| Qin et al. (2025) | 4 | Statistical (Hidden Markov Model) | ALPR data | Traffic demand level, signal delay, free-flow traffic | RMSE improved by 57–75%; MAE improved by 53–73% |
| Alaoua & Karim (2025) | 4 | Statistical Simulation (Monte Carlo) | IoT and synthetic lead-time data | Supply risk level, IoT availability, high variability | Accuracy = 99.62%; TPR = 100%; MAE; RMSE; FPR |
| Zhang et al. (2024) | 4 | Hybrid (IWOA + LSTM) | GPS data | Traffic conditions and time periods | MAPE = 9.47%; RMSE = 12.77%; MAE = 8.93%; $R^2$ improved by 10.65% |
| Arbabkhah et al. (2024) | 4 | XGBoost | AIS data | Short vs. long trips, congestion level | MAPE = 5%; $R^2$ = 98% |
| Ghazikhani et al. (2024) | 3 | XGBoost | GPS data | Mixed traffic conditions, time-series window size | — |
| Abdelhalim & Zhao (2025) | 4 | Deep Learning (Vision Transformer) | Camera images and transit data | Traffic conditions and scene variability | Accuracy = 80–85%; Precision = 95% |
| Chu et al. (2023) | 4 | Deep Learning (DECN) | Traffic, weather, and road-speed data | Congestion, traffic composition, late vs. early arrival | Overall error = 11%; Late-arrival error = 49% |
| Huang, Jiang & Chen (2024) | 3 | Deep Learning (DMAN with Attention) | Order, driver, and contextual data | Peak lunch hours, order density | — |
| Ma et al. (2023) | 3 | Hybrid (Bi-LSTM + CNN + Attention) | Flight trajectory and weather data | Congestion, convective weather, distance to destination | — |
| Sun, Fu et al. (2022) | 4 | Deep Learning (WDR + LSTM, Multi-task) | Ride-hailing and driver data | Data sparsity, driver variability | MAPE = 1.1%; 1.3% |
| Jawad-Ur-Rehman et al. (2022) | 4 | Deep Learning (Attention-GRU) | Traffic flow data | Prediction horizon (15–60 min), noisy data | RMSE = 0.50%; MAE = 2.20%; MAPE = 1.87%; $R^2$ = 0.87 |
| Shanthi et al. (2022) | 4 | Machine Learning (SVR) | GPS and weather data | Traffic congestion and weather conditions | RMSE = 20 s; Accuracy = 92% |
| Sun, Hu et al. (2022) | 3 | Deep Learning (WDR + LSTM + MLP, Metric Learning) | Ride-hailing, road network, and driver data | Road network and driver sparsity | — |
| Lu et al. (2022) | 4 | Analytical (Kinematic Wave Model) | Traffic signal and vehicle trajectory data | Signal timing, undersaturated traffic | RMSE = 5.2 s and 6.60 s; MAPE = 11.33% and 9.98%; $R^2$ = 0.96 |
| Kwak & Geroliminis (2021) | 3 | Statistical (Dynamic Linear Model) | Loop detector traffic data | Prediction horizon (0–60 min), congestion level | — |
| Li et al. (2021) | 4 | Hybrid (CNN + LSTM + Attention) | Taxi trajectory data | Rush hours, short- vs. long-term sequence | Accuracy = 90%; Error = 4.66%; RMSE = 18.6%; $R^2$ improved by 22.46% |

### 4.1. Development and Application of Data-Driven Predictive Models (RQ1)

This section addresses RQ1 by examining how data-driven predictive models are developed and applied in delivery delay prediction. Delivery delay prediction is predominantly addressed using machine learning, deep learning, and hybrid approaches, with deep learning and hybrid models dominating the reviewed studies due to their ability to capture complex spatio-temporal relationships (Balster et al., 2020; Sharmila, Velaga and Choudhary, 2020; Noman et al., 2025). Deep learning and hybrid approaches dominate the reviewed studies, particularly LSTM, CNN, Transformer, and attention-based models. Their ability to integrate convolutional, recurrent, and attention mechanisms enables the capture of complex spatial and temporal dependencies (Li et al., 2017; Ma et al., 2023; Zhang et al., 2024). Advanced architectures such as attention-based and multitask learning models enhance the representation of traffic dynamics and multi-agent interactions (Sharmila, Velaga and Choudhary, 2020; Huang, Jiang and Chen, 2024; Nguyen and Liem, 2025).

These developments indicate a shift toward more adaptive modeling strategies, supporting RQ1 by demonstrating how predictive models capture dynamic system behavior. Figure 4 presents the classification of predictive methods, while Table 1 summarizes model types, data sources, operational conditions, and performance metrics across the selected studies. This suggests that hybrid models generally outperform conventional machine learning and statistical approaches in complex environments, while deep learning models require large-scale data for optimal performance (De Araujo and Etemad, 2021; Noman et al., 2025). Conversely, conventional approaches offer greater interpretability and efficiency but are less effective in dynamic and nonlinear conditions (Wang, Tsapakis and Zhong, 2016; Shanthi et al., 2022).

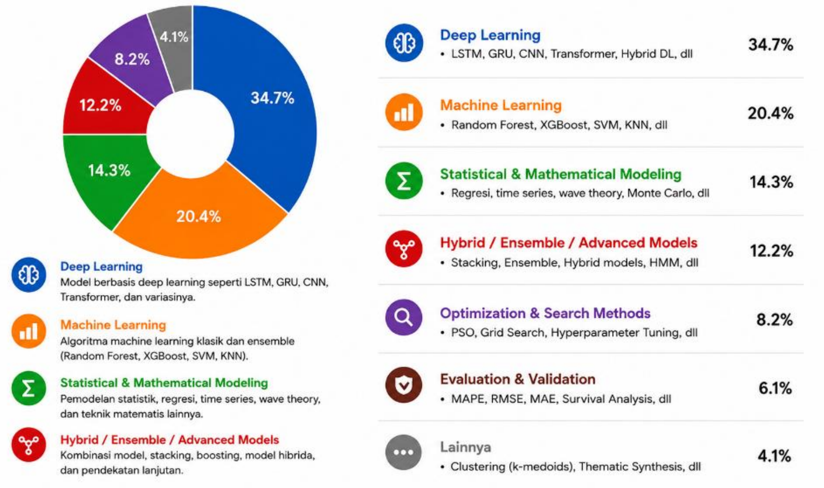



**Figure 4**. Percentage Distribution of Predictive Model Categories Across the Reviewed Studies

Conventional machine learning models, such as XGBoost and Support Vector Regression, remain effective for structured data and achieve competitive performance with feature engineering (Arbabkhah et al., 2024; Ghazikhani et al., 2024). Statistical and analytical approaches, including Dynamic Linear Models and kinematic wave-based models, are applied in specific contexts, particularly for short-term prediction and controlled traffic conditions, offering advantages in interpretability and computational efficiency (Kwak and Geroliminis, 2021; Lu et al., 2022). This diversity indicates that model development is strongly influenced by data characteristics and operational context. Several limitations remain. Many models rely on context-specific datasets, limiting generalizability, while increasing model complexity reduces interpretability and increases computational cost (Sun, Hu, et al., 2022; Alaoua and Karim, 2025). In addition, reliance on historical data constrains adaptability to dynamic disruptions (Zhang et al., 2020; Noman et al., 2025). These limitations highlight the need for more generalizable and interpretable predictive models. In addition to model development, model performance under varying operational conditions is further examined in Section 4.2.

### 4.2. Performance of Predictive Models under Operational Conditions (RQ2)

This section addresses RQ2 by examining how data sources and operational conditions influence model evaluation. Predictive models utilize diverse data, including GPS, AIS, traffic flow, and sensor-based inputs such as ALPR and computer vision, often combined with contextual variables such as weather, traffic density, and temporal features (Jawad-Ur-Rehman, Ul Haq and Muneeb, 2022; Shanthi et al., 2022). GPS, AIS, traffic, and weather data emerge as the most frequently utilized sources across the reviewed studies. The integration of heterogeneous data, including IoT and driver-related information, improves predictive capability but also introduces challenges related to data quality, compatibility, and availability (Alaoua and Karim, 2025; Sun et al., 2022). These challenges may affect data consistency and model reliability, particularly in dynamic and real-time operational environments. Figure 5 illustrates the classification of data sources used to support predictive modeling.

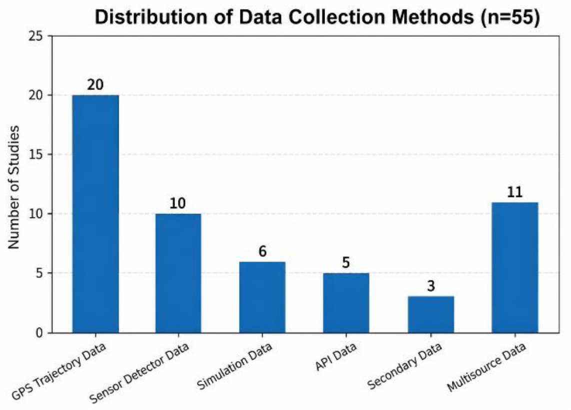



**Figure 5.** Data Collection Classification

In addition to data sources, operational conditions also play a critical role in shaping model performance. Factors such as congestion, weather variability, and peak periods are associated with increased prediction errors (Kim, 2016; Ma et al., 2023; Nguyen and Liem, 2025). Additional conditions, including prediction horizon, route characteristics, navigation direction, and signal timing, further influence evaluation outcomes, with short-term predictions generally achieving higher accuracy than long-term (Kwak and Geroliminis, 2021; Noman et al., 2025). Data-related factors such as sparsity and noise also reduce model performance (Čelan and Lep, 2020; Ghazikhani et al., 2024).

As summarized in Table 1, model performance is highly dependent on operational conditions. Congestion, adverse weather, and longer prediction horizons generally increase prediction errors, whereas more stable conditions improve accuracy. This supports RQ2 by demonstrating that predictive performance is influenced by both data characteristics and operational conditions.

## 5. DISCUSSION

The findings show that TTP and ETA dominate the literature, primarily supported by historical variables such as historical traffic data and historical travel time (Wang, Tsapakis and Zhong, 2016; Yuan, Huang and Zhang, 2020). In contrast, contextual variables including weather conditions, traffic density, and temporal features remain underutilized despite their potential to improve predictive performance (Kang et al., 2020; Wang, Liang and Delahaye, 2020). A notable gap is the limited incorporation of human-related variables, particularly driver behavior and algorithmic bias, which are rarely considered in existing models (Chai, Zhang and Yang, 2020; Chen and Fan, 2021). This indicates that current approaches remain largely data-centric and may not fully capture real-world complexity. Methodologically, machine learning and deep learning models, especially LSTM, XGBoost, and Transformer, dominate, with hybrid and ensemble approaches showing superior performance (Balster et al., 2020; Huang et al., 2026).

Despite the strong performance of machine learning, deep learning, and hybrid models, metaheuristic-based optimization remains underexplored and offers potential to improve parameter tuning, model robustness, and predictive performance. Most studies rely on secondary data sources such as GPS, AIS, and traffic sensors (Zhang et al., 2018; El Mekkaoui, Benabbou and Berrado, 2023). While integration of real-time and multi-source data, including IoT and V2X, is still limited (Alessandrini, Mazzarella and Vespe, 2019; Kaya and Utku Kalay, 2025). This highlights opportunities for heterogeneous data integration to enhance model robustness. As illustrated in Figure 6, the relationships between variables, models, data sources, and limitations highlight key research gaps and future directions in delivery delay prediction.

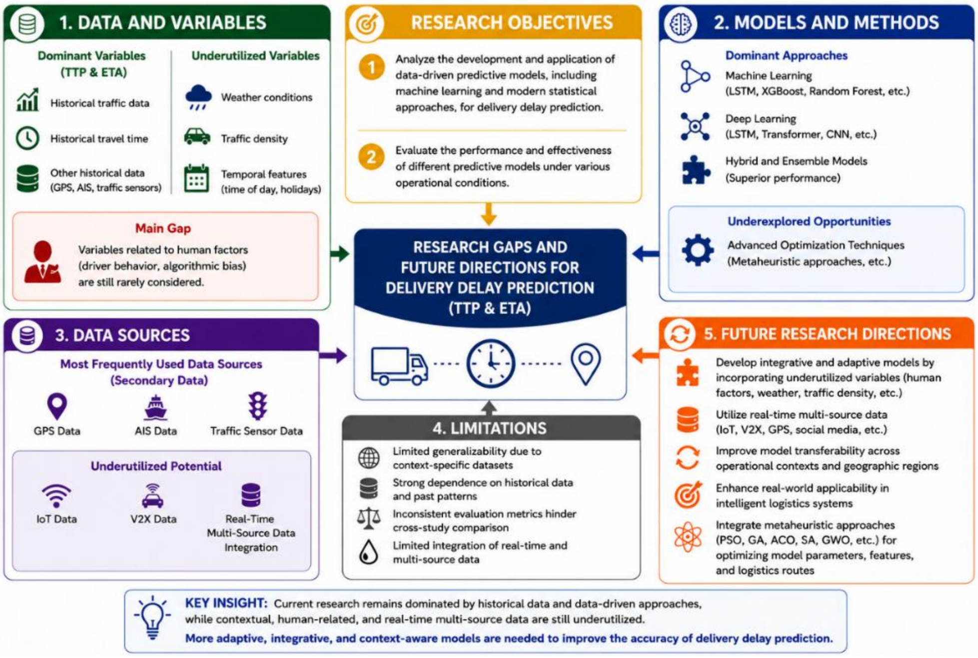



**Figure 6**. Overview of delivery delay prediction: variables, methods, gaps, and future needs.

Several limitations persist, including limited generalizability due to context-specific datasets (Zhao et al., 2019; Sun, Hu, et al., 2022) , strong dependence on historical data, and inconsistent evaluation metrics that hinder comparison across studies (Chen and Fan, 2021; Nguyen and Liem, 2025). These findings suggest the need for more integrative and adaptive models that incorporate underutilized variables, leverage real-time multi-source data, improve transferability across operational contexts, and explore advanced optimization techniques such as metaheuristic-based approaches. These gaps also highlight a disconnect between current model development and real-world applicability in logistics systems.

## 6. CONCLUSION

This study reviews data-driven predictive models for delivery delay prediction in logistics and transportation systems. Regarding RQ1, the findings indicate a transition from conventional statistical methods to machine learning, deep learning, and hybrid approaches that leverage heterogeneous data sources, including GPS, AIS, traffic, weather, and IoT data. These models have been widely applied across road transportation, maritime logistics, aviation, and intelligent supply chain systems, reflecting the growing adoption of data-driven approaches for delay prediction. Regarding RQ2, predictive performance is strongly influenced by data characteristics and operational conditions, including congestion, weather variability, prediction horizon, data sparsity, and noise. Although deep learning and hybrid models generally perform better in complex environments, no single approach consistently outperforms others across all operational contexts. This highlights the importance of aligning model selection with data availability and operational requirements. From a practical perspective, logistics and transportation practitioners should consider both operational conditions and data characteristics when implementing predictive models. The integration of heterogeneous and real-time data can improve prediction reliability and support more effective decision-making. Future research should focus on integrating real-time multi-source data and incorporating underexplored variables, such as weather conditions, traffic density, temporal features, and human-related factors. In addition, metaheuristic-based optimization offers promising opportunities to improve model robustness, adaptability, and practical applicability.

## REFERENCES


Abdelhalim, A. and Zhao, J. (2025) "Computer vision for transit travel time prediction: an end-to-end framework using roadside urban imagery," *Public Transport*, 17(1), pp. 221–246. Available at: https://doi.org/10.1007/s12469-023-00346-3.

Alaoua, A. and Karim, M. (2025) "Intelligent Early Warning System for Supplier Delays Using Dynamic IoT-Calibrated Probabilistic Modeling in Smart Engineer-to-Order Supply Chains," *Applied System Innovation*, 8(5). Available at: https://doi.org/10.3390/asi8050124.

Alessandrini, A., Mazzarella, F. and Vespe, M. (2019) "Estimated Time of Arrival Using Historical Vessel Tracking Data," *IEEE Transactions on Intelligent Transportation Systems*, 20(1), pp. 7–15. Available at: https://doi.org/10.1109/TITS.2017.2789279.

De Araujo, A.C. and Etemad, A. (2021) "End-to-End Prediction of Parcel Delivery Time with Deep Learning for Smart-City Applications," *IEEE Internet of Things Journal*, 8(23), pp. 17043–17056. Available at: https://doi.org/10.1109/JIOT.2021.3077007.

Arbabkhah, H. *et al.* (2024) "Automatic Identification System-Based Prediction of Tanker and Cargo Estimated Time of Arrival in Narrow Waterways," *Journal of Marine Science and Engineering*, 12(2). Available at: https://doi.org/10.3390/jmse12020215.

Balster, A. *et al.* (2020) "An ETA Prediction Model for Intermodal Transport Networks Based on Machine Learning," *Business and Information Systems Engineering*, 62(5), pp. 403–416. Available at: https://doi.org/10.1007/s12599-020-00653-0.

Basturk, O. and Cetek, C. (2021) "Prediction of aircraft estimated time of arrival using machine learning methods," *Aeronautical Journal*, 125(1289), pp. 1245–1259. Available at: https://doi.org/10.1017/aer.2021.13.

Bauer, D. and Tulic, M. (2018) "Travel time predictions: should one model speeds or travel times?," *European Transport Research Review*, 10(2). Available at: https://doi.org/10.1186/s12544-018-0315-7.

Čelan, M. and Lep, M. (2020) "Bus-arrival time prediction using bus network data model and time periods," *Future Generation Computer Systems*, 110, pp. 364–371. Available at: https://doi.org/10.1016/j.future.2018.04.077.

Chai, G., Zhang, L. and Yang, M. (2020) "Prediction of transit time on urban roads based on particle filtering," *Revue d'Intelligence Artificielle*, 34(2), pp. 189–194. Available at: https://doi.org/10.18280/ria.340209.

Chen, Z. and Fan, W. (2021) "A freeway travel time prediction method based on an xgboost model," *Sustainability (Switzerland)*, 13(15). Available at: https://doi.org/10.3390/su13158577.

Cheng, J., Li, G. and Chen, X. (2019) "Research on travel time prediction model of freeway based on gradient boosting decision tree," *IEEE Access*, 7, pp. 7466–7480. Available at: https://doi.org/10.1109/ACCESS.2018.2886549.

Chu, K.F. *et al.* (2023) "Deep Encoder Cross Network for Estimated Time of Arrival," *IEEE Access*, 11, pp. 76095–76107. Available at: https://doi.org/10.1109/ACCESS.2023.3294345.

Cristóbal, T. *et al.* (2019) "Bus travel time prediction model based on profile similarity," *Sensors (Switzerland)*, 19(13). Available at: https://doi.org/10.3390/s19132869.

Eddine, M.H.S., Saikouk, T. and Berrado, A. (2021) "Payment delay in supply chain management: a systematic literature review," *International Journal of Procurement Management*, 14(4), pp. 441–458.

El Mekkaoui, S., Benabbou, L. and Berrado, A. (2023) “Deep learning models for vessel’s ETA prediction: bulk ports perspective,” *Flexible Services and Manufacturing Journal*, 35(1), pp. 5–28. Available at: https://doi.org/10.1007/s10696-022-09471-w.

Fan, Q. *et al.* (2018) “Space-Time Hybrid Model for Short-Time Travel Speed Prediction,” *Discrete Dynamics in Nature and Society*, 2018. Available at: https://doi.org/10.1155/2018/7696592.

Fan, T. *et al.* (2023) “Inland Vessel Travel Time Prediction via a Context-Aware Deep Learning Model,” *Journal of Marine Science and Engineering*, 11(6). Available at: https://doi.org/10.3390/jmse11061146.

Gabellini, M. *et al.* (2024) “A Deep Learning Approach to Predict Supply Chain Delivery Delay Risk Based on Macroeconomic Indicators,” *Applied Sciences*, 14(11), p. 4688. Available at: https://doi.org/10.3390/app14114688.

Ghazikhani, A. *et al.* (2024) “Robust Truck Transit Time Prediction through GPS Data and Regression Algorithms in Mixed Traffic Scenarios,” *Mathematics*, 12(13). Available at: https://doi.org/10.3390/math12132004.

Huang, F., Jiang, W. and Chen, S. (2024) “A Multitask Attention Network for Food Delivery Time Prediction,” *Journal of Circuits, Systems and Computers*, 33(2). Available at: https://doi.org/10.1142/S0218126624500257.

Huang, P. *et al.* (2026) “High-Accuracy ETA Prediction for Long-Distance Tramp Shipping: A Stacked Ensemble Approach,” *Journal of Marine Science and Engineering*, 14(2), p. 177. Available at: https://doi.org/10.3390/jmse14020177.

Jawad-Ur-Rehman, C., Ul Haq, I. and Muneeb, M. (2022) “An attention-based recurrent learning model for short-term travel time prediction,” *PLoS ONE*, 17(12 December). Available at: https://doi.org/10.1371/journal.pone.0278064.

Jenelius, E. and Koutsopoulos, H.N. (2018) “Urban Network Travel Time Prediction Based on a Probabilistic Principal Component Analysis Model of Probe Data,” *IEEE Transactions on Intelligent Transportation Systems*, 19(2), pp. 436–445. Available at: https://doi.org/10.1109/TITS.2017.2703652.

Kang, L. *et al.* (2020) “Urban Traffic Travel Time Short-Term Prediction Model Based on Spatio-Temporal Feature Extraction,” *Journal of Advanced Transportation*, 2020, p. 1DUMMMY. Available at: https://doi.org/10.1155/2020/3247847.

Kaya, O. and Utku Kalay, M. (2025) “Spatio-Temporal Forecasting of Bus Arrival Times Using Context-Aware Deep Learning Models in Urban Transit Systems,” *IEEE Access*, 13, pp. 161423–161435. Available at: https://doi.org/10.1109/ACCESS.2025.3609530.

Kim, M.S. (2016) “Analysis of short-term forecasting for flight arrival time,” *Journal of Air Transport Management*, 52, pp. 35–41. Available at: https://doi.org/10.1016/j.jairtraman.2015.12.002.

Kwak, S. and Geroliminis, N. (2021) “Travel Time Prediction for Congested Freeways with a Dynamic Linear Model,” *IEEE Transactions on Intelligent Transportation Systems*, 22(12), pp. 7667–7677. Available at: https://doi.org/10.1109/TITS.2020.3006910.

Li, J. *et al.* (2017) “Bus arrival time prediction based on mixed model,” *China Communications*, 14(5), pp. 38–47. Available at: https://doi.org/10.1109/CC.2017.7942193.

Li, X. *et al.* (2021) “Spatiotemporal features-extracted travel time prediction leveraging deep-learning-enabled graph convolutional neural network model,” *Sustainability (Switzerland)*, 13(3), pp. 1–15. Available at: https://doi.org/10.3390/su13031253.

Li, Y. *et al.* (2022) “Real-Time Travel Time Prediction Based on Evolving Fuzzy Participatory Learning Model,” *Journal of Advanced Transportation*, 2022. Available at: https://doi.org/10.1155/2022/2578480.

Lin, Y.K., Chen, C.F. and Chou, T.Y. (2023) “Developing Prediction Model of Travel Times of the Logistics Fleets of Large Convenience Store Chains Using Machine Learning,” *Algorithms*, 16(6). Available at: https://doi.org/10.3390/a16060286.

Lu, L. *et al.* (2022) “A Real-Time Prediction Model for Individual Vehicle Travel Time on an Undersaturated Signalized Arterial Roadway,” *IEEE Intelligent Transportation Systems Magazine*, 14(5), pp. 72–87. Available at: https://doi.org/10.1109/MITS.2021.3068416.

Ma, Y. *et al.* (2023) “A Spatiotemporal Neural Network Model for Estimated-Time-of-Arrival Prediction of Flights in a Terminal Maneuvering Area,” *IEEE Intelligent Transportation Systems Magazine*, 15(1), pp. 285–299. Available at: https://doi.org/10.1109/MITS.2021.3132766.

Müller, N. *et al.* (2025) “An analytical review of predictive methods for delivery delays in supply chains,” *Supply Chain Analytics*, 11. Available at: https://doi.org/10.1016/j.sca.2025.100130.

Nguyen, C.H.C. and Liem, R.P. (2025) “Multi-aircraft attention-based model for perceptive arrival transit time prediction,” *Advanced Engineering Informatics*, 64. Available at: https://doi.org/10.1016/j.aei.2024.103067.

Noman, A. Al *et al.* (2025) “Multi-model learning for vessel ETA prediction in inland waterways using multi-attribute data,” *Systems Science and Control Engineering*, 13(1). Available at: https://doi.org/10.1080/21642583.2025.2546833.

Oh, S. *et al.* (2018a) “Short-term travel-time prediction on highway: A review on model-based approach,” *KSCE Journal of Civil Engineering*, 22(1), pp. 298–310. Available at: https://doi.org/10.1007/s12205-017-0535-8.

Qiao, W. *et al.* (2016) “Freeway path travel time prediction based on heterogeneous traffic data through nonparametric model,” *Journal of Intelligent Transportation Systems: Technology, Planning, and Operations*, 20(5), pp. 438–448. Available at: https://doi.org/10.1080/15472450.2016.1149700.

Qin, H. *et al.* (2025) “Real-Time Traffic Arrival Prediction for Intelligent Signal Control Using a Hidden Markov Model-Filtered Dynamic Platoon Dispersion Model and Automatic License Plate Recognition Data,” *Applied Sciences (Switzerland)*, 15(21). Available at: https://doi.org/10.3390/app152111537.

Rezki, N. and Mansouri, M. (2024) “Machine Learning For Proactive Supply Chain Risk Management: Predicting Delays And Enhancing Operational Efficiency,” *Management Systems in Production Engineering*, 32(3), pp. 345–356. Available at: https://doi.org/10.2478/mspe-2024-0033.

Shanthi, N. *et al.* (2022) “Analysis on the Bus Arrival Time Prediction Model for Human-Centric Services Using Data Mining Techniques,” *Computational Intelligence and Neuroscience*, 2022. Available at: https://doi.org/10.1155/2022/7094654.

Sharmila, R.B., Velaga, N.R. and Choudhary, P. (2020) “Bus arrival time prediction and measure of uncertainties using survival models,” *IET Intelligent Transport Systems*, 14(8), pp. 900–907. Available at: https://doi.org/10.1049/iet-its.2019.0584.

Silvestre, J. *et al.* (2024) “A deep learning-based approach for predicting in-flight estimated time of arrival,” *Journal of Supercomputing*, 80(12), pp. 17212–17246. Available at: https://doi.org/10.1007/s11227-024-06060-6.

Su, X., Alatas, B. and Sohaib, O. (2025) “An Express Management System With Graph Recurrent Neural Network for Estimated Time of Arrival,” *Journal of Organizational and End User Computing*, 37(1). Available at: https://doi.org/10.4018/JOEUC.370912.

Sun, Y. *et al.* (2021) “Interactive gated recurrent unit and its application for estimated time of arrival,” *Scientia Sinica Informationis*, 51(5), pp. 822–833. Available at: https://doi.org/10.1360/SSI-2020-0147.

Sun, Y., Hu, W., *et al.* (2022) “Alleviating Data Sparsity Problems in Estimated Time of Arrival via Auxiliary Metric Learning,” *IEEE Transactions on Intelligent Transportation Systems*, 23(12), pp. 23231–23243. Available at: https://doi.org/10.1109/TITS.2022.3200445.

Sun, Y., Fu, K., *et al.* (2022) “CoDriver ETA: Combine Driver Information in Estimated Time of Arrival by Driving Style Learning Auxiliary Task,” *IEEE Transactions on Intelligent Transportation Systems*, 23(5), pp. 4037–4048. Available at: https://doi.org/10.1109/TITS.2020.3040386.

Ting, P.Y. *et al.* (2020) "Freeway Travel Time Prediction Using Deep Hybrid Model-Taking Sun Yat-Sen Freeway as an Example," *IEEE Transactions on Vehicular Technology*, 69(8), pp. 8257–8266. Available at: https://doi.org/10.1109/TVT.2020.2999358.

Wang, J., Tsapakis, I. and Zhong, C. (2016) "A space-time delay neural network model for travel time prediction," *Engineering Applications of Artificial Intelligence*, 52, pp. 145–160. Available at: https://doi.org/10.1016/j.engappai.2016.02.012.

Wang, Z., Liang, M. and Delahaye, D. (2018) "A hybrid machine learning model for short-term estimated time of arrival prediction in terminal manoeuvring area," *Transportation Research Part C: Emerging Technologies*, 95, pp. 280–294. Available at: https://doi.org/10.1016/j.trc.2018.07.019.

Wang, Z., Liang, M. and Delahaye, D. (2020) "Automated data-driven prediction on aircraft Estimated Time of Arrival," *Journal of Air Transport Management*, 88. Available at: https://doi.org/10.1016/j.jairtraman.2020.101840.

Wani, A.A. (2025) "Ten quick tips for improving estimated time of arrival predictions using machine learning in logistics and transportation systems," *PeerJ Computer Science*, 11. Available at: https://doi.org/10.7717/peerj-cs.3259.

Yuan, H., Huang, Z. and Zhang, H. (2020) "Travel Time Prediction Model of Freeway Corridor Based on Real-Time Safety Reliability," *Journal of Advanced Transportation*, 2020. Available at: https://doi.org/10.1155/2020/8882011.

Zhang, B. *et al.* (2024) "Bus Arrival Time Prediction Based on the Optimized Long Short-Term Memory Neural Network Model With the Improved Whale Algorithm," *Journal of Advanced Transportation*, 2024(1). Available at: https://doi.org/10.1155/2024/6997338.

Zhang, H. *et al.* (2020) "A Prediction Model for Bus Arrival Time at Bus Stop Considering Signal Control and Surrounding Traffic Flow," *IEEE Access*, 8, pp. 127672–127681. Available at: https://doi.org/10.1109/ACCESS.2020.3004856.

Zhang, Y. *et al.* (2018) "Travel time prediction with viscoelastic traffic model," *Applied Mathematics and Mechanics (English Edition)*, 39(12), pp. 1769–1788. Available at: https://doi.org/10.1007/s10483-018-2400-9.

Zhao, B. *et al.* (2019) "Data-driven next destination prediction and ETA improvement for urban delivery fleets," *IET Intelligent Transport Systems*. Institution of Engineering and Technology, pp. 1624–1635. Available at: https://doi.org/10.1049/iet-its.2019.0148.

Zhao, J. *et al.* (2018) "Travel time prediction of expressway based on multi-dimensional data and the particle swarm optimization–autoregressive moving average with exogenous input model," *Advances in Mechanical Engineering*, 10(2). Available at: https://doi.org/10.1177/1687814018760932.