Ng *et al. European Transport Research Review* (2026) 18:27
https://doi.org/10.1186/s12544-026-00788-3

European Transport Research Review

**REVIEW** **Open Access**

# Application of license plate recognition data in intelligent transportation systems: a review

Jiehui Ng[1], Baichuan Mo[2], Lyuchao Liao[3] and Ruimin Li[4*]

**Abstract**

Massive and detailed data collected by the license plate recognition (LPR) systems installed at intersections can offer opportunities and support for a wide range of research directions in the transportation field. This paper summarizes the literatures on LPR-based research and corresponding methods from 2006 to the present, demonstrating the significant role of LPR data across multiple transportation research domains. The study surveys the literature and highlights how LPR data serves as a foundation for estimating and predicting traffic state, analyzing travel demand, optimizing traffic control, and enhancing traffic safety management, owing to its unique data advantages. By reviewing data modelling methods and models employed across different research directions, this study proposes new opportunities for future innovations in applications of LPR data. Specifically, studies on LPR-based traffic control strategies, proactive traffic management, intelligent vehicle infrastructure cooperative systems, and other related areas remains relatively underdeveloped. The application of more advanced techniques, such as reinforcement learning, hybrid learning, and large models (i.e. large language models and pre-trained foundation models), as well as the development of LPR-oriented simulation environment, are needed for further breakthroughs in LPR data modelling.

**Keywords** LPR, Traffic state, Travel demand, Traffic safety and security, Traffic control, Data mining

## 1 Introduction

The advancement of intelligent transportation systems (ITS) has provided extensive traffic data for various aspects of traffic analysis, traffic operations, and transportation planning. Motor vehicle detection data generated from ITS can be divided into two categories: (1) fixed detector-based data and (2) probe vehicle data. The former data mainly come from fixed devices such as inductive loops, magnetometers, microwave detectors, and auto vehicle identification (AVI) systems. The later ones are collected from movable devices such as global positioning system (GPS), on-board diagnostics (OBD) recorders, and mobile phone sensors.

License plate recognition (LPR) system, as a class of the AVI systems, are deployed in various traffic management scenarios, including intersections, ramps, and parking facilities. Although a few studies have explored the use of LPR data collected at ramps [49] or parking entrances [67], such cases remain relatively uncommon because LPR devices installed in these locations provide limited spatial coverage and discontinuous observations, making them insufficient to support in-depth analyses across diverse research dimensions. In contrast,

---

*Correspondence:
Ruimin Li
lrmin@tsinghua.edu.cn
[1]Department of Civil Engineering, Tsinghua University, Room 321, New Civil Engineering Building, Beijing 100084, China
[2]Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[3]Fujian Provincial Universities Engineering Research Center for Intelligent Driving Technology, Fujian University of Technology, Fuzhou, Fujian 350118, China
[4]Department of Civil Engineering, Tsinghua University, Room 316, New Civil Engineering Building, Beijing 100084, China

intersection-based LPR installations constitute the most widely deployed and consistently utilized infrastructure worldwide, offering broader coverage, higher continuity, and richer spatiotemporal information. Therefore, the present study focuses on LPR data collected at intersections, which represents the predominant and most methodologically scalable application scenario in existing research.

The LPR data, derived from LPR system, is also known by various terms such as automatic license plate recognition data [19], automatic number plate recognition data [71], and vehicle license plate recognition data [111]. With rapid expansions of LPR systems in many regions, it has captured many research attentions in recent years. The first large-scale implementation of LPR system for megacities was done by Transport for London, where 700 cameras are installed on all roads that cross the congestion charging zone. Apart from the UK, the application and technological development of LPR systems in Europe have been longstanding, and the early deployment of these systems enabled European researchers to explore LPR data at an earlier stage. This early start provided a demonstrative and foundational role for subsequent LPR-based studies conducted worldwide. Hadavi et al. [25] reported wide coverage of LPR cameras in European countries such as Belgium, Denmark, and the Netherlands. From the perspective of vehicle detection technology, countries like Germany, Italy, and France have accumulated substantial results in vision-based vehicle detection methods [75]. Studies on the application and modeling of LPR data have also been ongoing in Spain [4, 6, 7, 72], Sweden [36, 60], the UK [14, 47, 71, 89], and other European countries.

Over the past two decades, while some cities in the USA have achieved high coverage of LPR cameras [69], the installation of LPR devices at signalized intersections in Chinese cities, both large and small, has been even more extensive [82, 94, 109]. Based on the geographic distribution of the literature reviewed in this study, research communities in Asia, particularly China, have been considerably more active in LPR-based studies in recent years. This trend appears to be driven primarily by differences in data availability rather than methodological discrepancies. In Europe and North America, increasing attention to data-privacy regulations has made access to LPR data more challenging, and many studies have shifted toward Bluetooth-based data, which share fundamental information attributes with LPR data such as device-level identifiers, timestamps, and movement patterns [34, 39]. In contrast, the broader accessibility of LPR data in China has enabled a larger volume of LPR-based research, although the underlying data characteristics used in these studies remain consistent with those documented in earlier European deployments. With the advancement of LPR system technologies and their increasing coverage globally, they are no longer limited to ad-hoc traffic monitoring, parking management, and public security applications. Instead, they may provide a reliable data foundation for data-driven traffic management, transportation planning and so on.

Incorporating multi-dimensional information such as vehicle ID, passing timestamps, detection location, and other vehicle attributes, LPR data offers advantages in both macroscopic and microscopic traffic data analysis. In general, LPR-based traffic datasets exhibits three unique features: (1) accurate and complete timestamp sequences for vehicles passing fixed detection points across individual lanes, (2) easy acquisition of link travel times enabled by vehicle tracking between upstream and downstream LPR locations, and (3) lane-level or location-specific traffic state estimation enabled by high-resolution monitoring at each detection point [25, 108].

Research on LPR data has been progressively developing over the past 20 years. Reviews on this research can be categorized into two aspects: (1) reviews focusing on the development of plate recognition algorithms and (2) reviews on application studies based on LPR data (LPR-based studies). For instance, Anagnostopoulos et al. [3] and Du et al. [19] reviewed the development of plate recognition techniques and algorithms in LPR systems, while He et al. [29] summarized various aspects of LPR data application, with a particular emphasis on traffic state representation, including traffic state estimation and prediction, network flow reconstruction, etc. Tang et al. [84] conducted a review dealing with the technological development of LPR system, robustness of LPR performance under different conditions, application scenarios of LPR data, and relevant policies. However, existing reviews lack an in-depth and comprehensive investigation into the topics and methodologies of LPR-based research in the transportation field. Furthermore, the overall landscape and evolutionary directions of these studies are missing.

To address the research gaps, this study provides a systematic review for the applications and methodologies of LPR-based studies. Figure 1 illustrates the overall framework of this study. The contributions of this study are threefold.

- First, from the topic point of view, LPR-based studies are summarized into four categories: (1) traffic state estimation and prediction, (2) travel demand analysis, (3) traffic control, and (4) traffic safety and security. Each category is further subdivided into various research branches, detailing the pros and cons of LPR data and current research state within these branches.

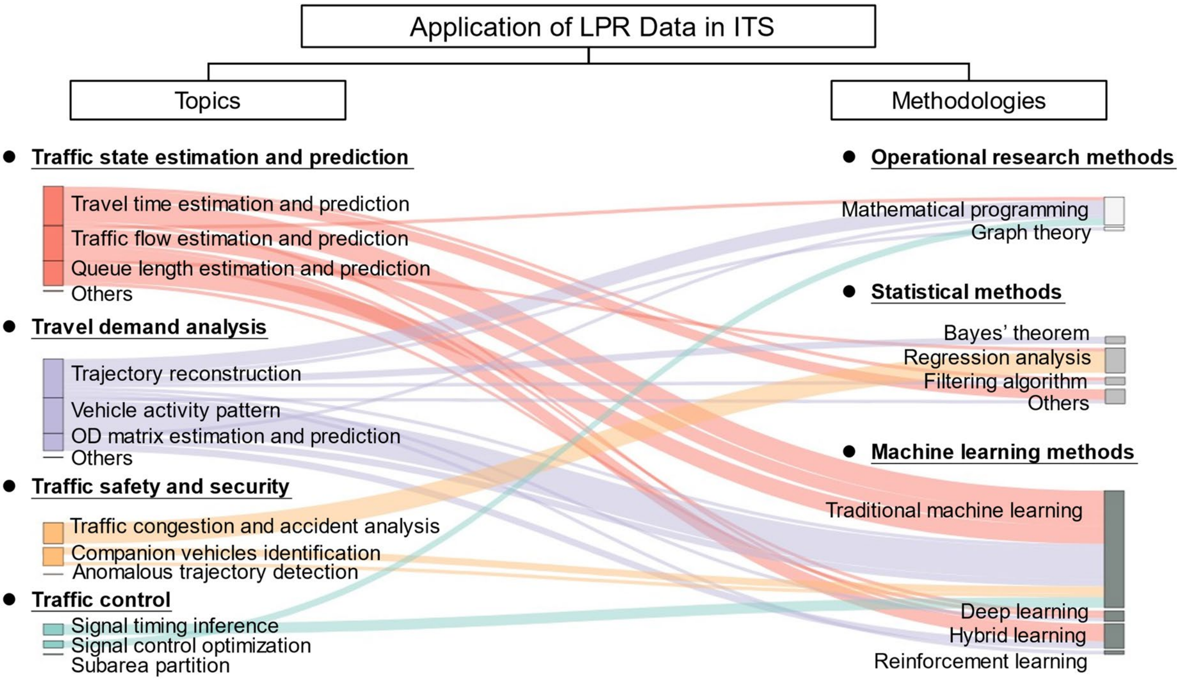

**Fig. 1** Overall framework of the review

- Second, from the methodology point of view, the methods used in LPR-based research are investigated and categorized, with an analysis of the advantages and inherent issues of each method in specific research directions.
- Lastly, this study outlines future directions for LPR-based research and identifies potential methodologies for future exploration.

The rest of this paper is structured as follows. Section 2 demonstrates the methodology used to conduct the review. Section 3 classifies LPR-based studies from the topic point of view and analyzes the research trend for the last ten years. Following this, methods and models based on LPR data are introduced in Sect. 4. Section 5 discusses open research issues and future opportunities. Finally, Sect. 6 provides the conclusions of this review.

## 2 Methodology

To address the research gaps outlined in this study, relevant works were identified using a few stages. First, relevant literature was collected from Web of Science, Scopus, and ProQuest databases using combinations of keywords such as LPR, NPR, ITS, along with terms representing various aspects of the transportation field, including traffic signal control, traffic safety, travel demand, etc. Based on the overall research trends of LPR-based studies, the time span for the relevant literature was set to the past 20 years (from 2006 to the present).

Besides, the abstracts and conclusions of relevant works were reviewed and filtered through a screening process with three criteria to meet the purpose of this study:

- Studies that explore the processing or mining of LPR data, with a focus on its contributions to key areas in the transportation field such as traffic control, transportation planning, and traffic safety.
- For innovative LPR data mining research, a clear methodological framework or model should be established to complete tasks such as data preprocessing, information extraction, and data analysis.
- In data fusion research, LPR data should serve as the primary source of information extraction, rather than as an auxiliary task.

Additionally, the snowball search method was applied to find more related studies by cross-referencing the studies discovered from previous stages. Finally, following the literature filtering and selection stages, 103 works were used to conduct the systematic literature review.

## 3 Topics in LPR-based studies

Over the last decade, research on LPR data has rapidly increased, establishing a distinct research trend that can be categorized into four major fields: (1) traffic state estimation and prediction, (2) travel demand analysis, (3) traffic control, and (4) traffic safety and security. Figure 2 outlines the popularity of the four fields from 2016 to the present, where y-axis is the number of relevant studies for each topic.

The trend indicates that studies concerning traffic state estimation and travel demand prediction has always been popular. Conversely, research initiatives in the realm of traffic safety and security emerged at an earlier juncture, while recent years have witnessed a shift in research focus towards the domain of traffic control. Also, each category can further be divided, where we will elaborate in the following subsections.

### 3.1 Traffic state estimation and prediction

Given the rich information of large LPR dataset, it is widely applied in various studies concerning the estimation and prediction of traffic state. The main research directions are as follows.

- Travel time estimation and prediction: one of the most prevalent research domains within LPR data is travel time estimation [17, 59, 105], prediction [20, 21], and distribution estimation [22, 43]. Timestamp details of LPR data inherently include both vehicle-level and traffic-flow-level information, such as vehicle queues and delays at the intersections. By inferring traffic state parameters from different dimensions, further in-depth studies can be conducted, such as the modelling of travel time distribution. In this context, a notable feature of LPR data is its high penetration rate among passing vehicles, which allows travel time distributions to be estimated more reliably and enables researchers to better capture the inherent variability and patterns for travel times. Research has revealed considerable variations in travel time distribution and reliability among different granularities such as urban arterials [11, 36, 64], expressways [41], and networks [113].
- Traffic flow (mainly traffic volume and speed) estimation and prediction: In addition to timestamp information, LPR data also includes the intersection IDs and passing direction details of the vehicles. With the support of road network topology data, insights such as traffic volume and speed on specific road segments can be further inferred. The robust traffic flow patterns that can be extracted from the large-scale LPR data, which is oriented toward various types of motor vehicles, are irreplaceable by other types of sparse data. As a consistently popular topic, estimation and prediction of traffic flow-related parameters encompasses various aspects such as traffic volume [23, 73, 78, 85, 86, 107, 110] and average traffic speed [70, 96]. To capture the spatiotemporal heterogeneity of traffic flow, data fusion frameworks that integrate data from media check-ins, point of interest (POIs), and GPS sources have also been used [74, 91, 96].
- Queue length estimation and prediction: embedded information like stop line passing times and discharge headways can be leveraged to estimate queue lengths through time-space diagrams by mining large volumes of LPR data [51, 58, 76, 1, 88,

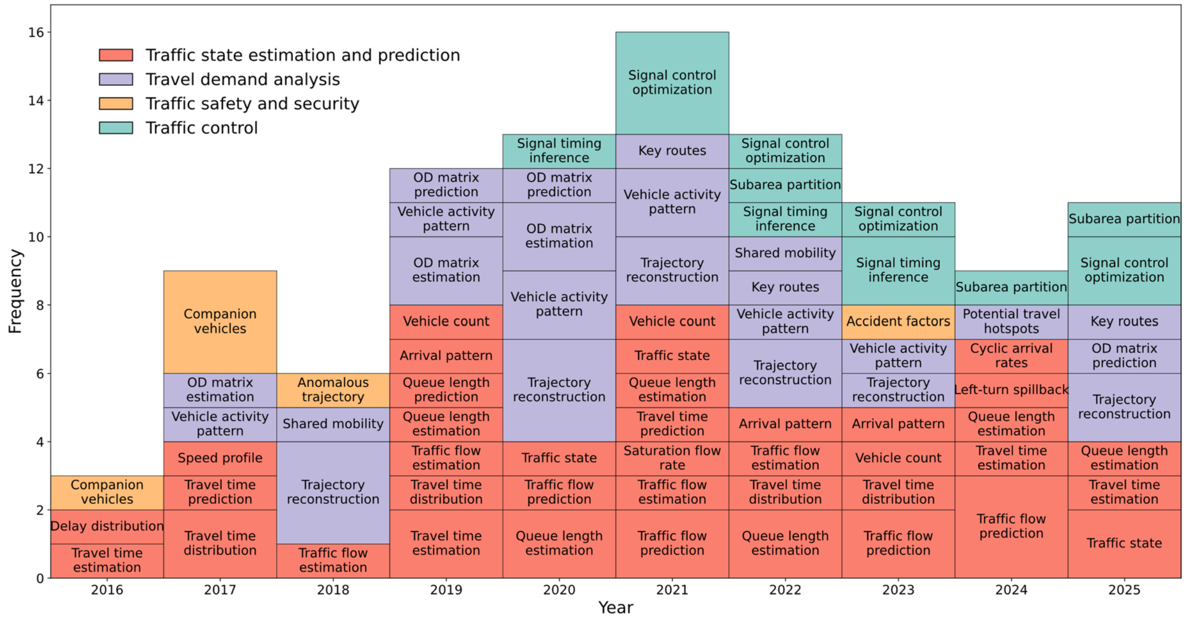

**Fig. 2** Research topics distribution from 2016 to the present

93, 108]. Additionally, the lane-level information recorded for individual vehicles enables the forecasting of travel demand at downstream intersections, making the prediction of queue lengths at these intersections feasible [42]. To further leverage the advantages of different data types and improve experimental accuracy, Tan et al. [81] constructed a data fusion framework that leverages the queueing attributes of LPR data and the continuity attributes of GPS data to estimate lane-based queue lengths at signalized intersections.

- Others: as mentioned above, LPR data allows for the extracting of arrival patterns at the intersection level, providing valuable information such as arrival rate and time headway within both upstream and downstream detection areas [1, 42, 43]. Besides the conventional research areas described earlier, studies also use LPR data for estimating delay distribution at signalized intersections [12], speed profile and emissions [63], and saturation flow rate of mixed traffic flows [90]. With the high coverage of LPR cameras, studies on link dynamic vehicle counting, which indicates road occupancy rates and offers useful inputs for signal control optimization, have also become feasible [26, 27]. Additionally, LPR-based research covers the classification of traffic states, aiming to categorize the severity of traffic congestion in urban networks based on traffic state parameters such as travel time and traffic flow [18, 77, 83, 109]. Recently, some studies have also explored cyclic arrival rates [2] and the identification of left-turn spillback [94] at signalized intersections using LPR data.

Compared to other types of traffic data, LPR systems can provide a rich and robust traffic state information without explicitly accounting for data penetration rate, covering spatial scales from the macro network level and meso link level to the micro intersection level. However, it is important to note that although LPR records can be linked through timestamp order and intersection connectivity, the data may not be truly continuous. In fact, LPR data reveals its limitations in certain micro-level traffic state parameter extraction tasks that rely on continuous data characteristics. Therefore, a trade-off must be made between the high-coverage LPR data and the highly continuous probe vehicle data to improve research accuracy.

### 3.2 Travel demand analysis

The high coverage rates of LPR data provides valuable insights for understanding detailed vehicle-based travel demand. The main research directions are as follows.

- Trajectory reconstruction: vehicle trajectories are important for understanding traffic conditions and travel demand patterns. In this context, LPR-based trajectory reconstruction data typically offers high coverage and representativeness, providing strong support for various subsequent tasks, such as travel pattern and demand analysis. On one hand, LPR systems installed at intersections may not capture complete vehicle trajectories when there are missing data or just no cameras on some road sections. To address the issue of LPR detection not providing complete coverage of all intersections within a road network, many ongoing research focused on reconstructing vehicle trajectories utilizing LPR data [9, 28, 46, 68, 100, 106]. On the other hand, unlike other types of traffic data, LPR systems are able to provide lane information of vehicle's movement (i.e., straight or turning). This information facilitates the direct reconstruction of trajectories when missing information for a specific vehicle spans less than two consecutive intersections. In order to enhance the accuracy and robustness of trajectory reconstruction, many studies also proposed data fusion frameworks incorporating probe vehicle-based data, predominantly GPS data [5, 16, 61].
- Origin-destination (OD) matrix estimation and prediction: after preprocessing LPR data to separate trip chains, the vast amount of resulting sub-trip data offers a significant advantage in representing transportation network vehicle-based travel demand compared to other datasets. Vehicle trajectories extracted from LPR data have been widely used for the estimation [62, 66, 87] and prediction [33, 53, 112] of network-level OD matrices, providing more precise and representative results at lower cost compared to traditional vehicle OD survey methods. However, it is worth noting that using LPR data to generate OD information captures only the travel patterns of licensed-plated motor vehicles, excluding other modes of transport like non-motorized or rail-based vehicles.
- Vehicle activity pattern: with the trip chains extracted from LPR data, another important research branch is to use them to identify and analyze regional vehicle activity patterns based on clustering or other machine learning techniques [8, 10, 30, 31, 79, 95]. Research in this area spans different levels of travel pattern including microscopic commuting [104] and non-commuting [89] travel demand for car commuters, and macroscopic analyzes of travel patterns within regional urban network [56].
- Others: studies related to travel demand analysis also includes the identification of key routes, prediction of potential travel hotspots, and analysis of shared

mobility. For example, Yao et al. [103] and Liu et al. [52] recognized routes with high traffic demand within the road network using LPR data. Since LPR data provides fined-grained trajectory-based activity patterns in urban areas, research also explored the feasibility and benefits of shared mobility [44, 55] based on LPR data. Mining and processing the latent travel demand embedded within LPR data can also facilitate predictive analyzes for potential travel hotspots [99].

Despite the vast amount of license plate-oriented trip information provided by LPR data, which strongly supports various branches of travel demand analysis research, its limitations are evident. Specifically, the data recorded by LPR systems is restricted to motor vehicles traveling on the road, and thus cannot objectively or accurately capture the macro-scale urban mobility patterns across all travel modes, unlike data from sources such as cellular probe data.

### 3.3 Traffic control

As LPR data contains information on both traffic conditions and travel demand patterns, it holds significant potential for application in traffic control. The focuses are as follows.

- Signal control optimization: since LPR data can be used to predict operational status and traffic demand within transportation networks, it facilitates research endeavors in optimizing signal control strategies. Notably, the real time generation of LPR data offers a dynamic and closed loop feedback mechanism for traffic signal control, thereby enhancing the operational efficiency of traffic management initiatives. However, current research based on LPR data is predominantly confined to the optimizing signal control for isolated intersection [65, 92] or arterial coordination [13, 45, 83, 98]. There is a lack of studies on network-level control [57].
- Signal timing inference: considering the limited transparency and periodic updates of signal timing schemes in different regions, the inference of signal timing is crucial for several application fields like real-time navigation guidance. Since LPR data can be used to detect vehicle queueing status and the passage of vehicles through stop lines, signal timing schemes for each intersection can be inferred from the integrated statistics of vehicle queueing or movement statuses [43, 48, 102, 109].
- Subarea partition: the main objective of subarea partition is to determine the optimal segmentation of urban arterials [38] or networks [37] to facilitate traffic demand-based coordinated signal control. Utilizing the OD flow information extracted from LPR data, a correlation analysis of OD flow matrices is conducted for subarea partition using graph segmentation algorithms. Recently, Hu et al. [32] have introduced a multi-step hybrid subarea partitioning approach, offering a new perspective for LPR-based subarea partitioning research.

Although LPR data demonstrates superior performance in supporting applications within the traffic control domain, LPR-based studies in this field remains limited. This is primarily because establishing real-time LPR data-driven decision-feedback mechanisms for control optimization (mainly traffic signal control and subarea partitioning) remains highly challenging. In fact, existing studies on signal control optimization and subarea partitioning based on LPR data have so far only validated their models at the simulation level.

### 3.4 Traffic safety and security

Given the rich traffic state information embedded in LPR data, the extensive data generated by high-coverage LPR systems provides significant advantage for multifactorial analyzes of traffic safety issues, such as traffic congestion and incident analysis, thereby offering a foundational basis for decision-making in traffic safety management. Additionally, leveraging LPR data for spatiotemporal information mining has also facilitated research on identifying companion vehicles and anomalous trajectories.

- Traffic congestion and accident analysis: LPR data harbors various traffic-related information (traffic volume, average traffic speed, travel time, congestion levels, etc.), making it a robust foundation for studies on traffic congestion and incidents, especially for understanding causes and factors for traffic accidents. As a pioneering city in large-scale LPR deployment, London provided the LPR data used by Chow et al. [14] to conduct an empirical analysis of traffic congestion in its central area. Xu et al. [97], on the other hand, integrated LPR data with accident records, network topology data, and POI data to conduct a multifactorial analysis of how traffic state information, road network structural characteristics, and built environment factors influence traffic safety.
- Companion vehicles identification: companion vehicles are defined as vehicles that coexist within a specific temporal and spatial window. Identifying companion vehicles is helpful for various areas like transportation management and military surveillance. With the spatial and temporal information recorded by LPR systems, techniques

for identifying companion vehicles find broad applications across diverse domains such as traffic management and safeguarding specialized vehicles [50]. Existing studies have relied on sequence mining as an analytical framework [40, 111, 114].

- Anomalous trajectory detection: the LPR system plays a dominant role in traffic supervision conducted by Chinese traffic police. It also supports anomalous trajectory identification research conducted by domestic scholars in China, utilizing the LPR data collected by the system [24, 80].

Figure 2 shows a significant decline in traffic safety and security research in recent years. Although LPR data can sometimes be supportive of traffic safety analysis, it does not directly include safety information. This limitation makes it challenging to conduct more in-depth, insightful, and extensive applications using LPR data in this domain.

## 4 Methodologies of LPR-based modelling

While many research directly employ sequence mining and analytical techniques for LPR data implementation [50, 83, 88, 106], the methods or models employed in LPR-based research can be broadly categorized into three domains: operational research (OR) methods, statistical methods, and machine learning (ML) methods. These methods are applied to different research topics that are summarized in Table 1.

### 4.1 Operational research methods

OR methods involved in LPR-based studies consist of two major branches, namely graph theory and mathematical programming.

- Graph theory: if intersections in a road network are considered as graph nodes and the links between them as graph edges, graph theory offers an unparalleled analytical perspective for analyzing LPR data integrated with road network topology information. Graph theory has been widely employed in LPR data applications to address problems within the domain of travel demand analysis. For instance, in LPR-based trajectory reconstruction, K-shortest paths (KSP, mainly based on Dijkstra algorithm) has been used to generate potential routes between two intersections, complementing the missing data in LPR records [28, 68]. However, conventional graph theory not only faces the dimensionality curse in complex network systems but also encounters significant limitations in extracting and representing graph embeddings.
- Mathematical programming (MP): many real-world transportation issues can be formulated as optimization problems, enabling the application of MP models for solutions. As demonstrated in Table 1, weighted quadratic function, ordinary least square (OLS) model, integer programming, mixed integer linear programming (MILP), model predictive control (MPC), etc., have been employed in research issues such as traffic flow (including traffic volume and average traffic speed) estimation, trajectory reconstruction, OD estimation, and signal control optimization. Some studies, like the staircase vehicle order-changing model proposed by Deng et al. [16] and the stochastic user equilibrium (SUE) principle-based path-flow estimator proposed by Yang and Sun [100], proposed MP models directly oriented by traffic theories. Notably, for complex problems involving multiple objectives, constraints, and non-convexity, MP models pose considerable challenges and limitations in terms of modelling difficulty and model convergence.

### 4.2 Statistical methods

Statistical methods provide a systematic framework for collecting, organizing, analyzing, and interpreting data, and play an irreplaceable role in inferring general patterns and regularities from traffic data. The diverse array of statistical methods implicated in LPR-based studies can be summarized as follows.

- Bayes' theorem: as a fundamental statistical principle that allows for the calculation of conditional probabilities by incorporating prior knowledge and new empirical evidence, Bayes' theorem finds widespread application in state inference of stochastic traffic processes. Specifically, Bayes' theory not only serves as a theoretical underpinning for inferring complex relationships among diverse factors in hierarchical and multi-level analysis regarding traffic congestion and accidents issues, but also provides a principled basis for trajectory reconstruction tasks based on traffic information mined from LPR data, as well as providing theoretical foundations for certain machine learning models (Table 1). Notably, due to its reliance on the prior distribution of data, Bayes' theorem has limited robustness when handling missing or anomalous data.
- Markov chain: in transportation research, Markov chains are frequently utilized to represent state transitions in stochastic traffic systems characterized by memoryless properties. Table 1 demonstrates

**Table 1** Utilization of various methods or models across different topics

| Topics | OR methods | Statistical methods | ML methods |
|---|---|---|---|
| Travel time estimation | - | Dion's Algorithm [17], Cycle-based travel-time filtering [105] | CFSDP [59] |
| Travel time prediction | - | DLM [47], DLM with Markov Switching [21], Extended KF [60] | LSTM-CNN [20] |
| Travel time distribution | - | - | LMM [36], FGM Copula [11], Bayesian +Shannon's Information Entropy [22], GMM [43], Modified GMM [41] |
| Traffic flow estimation | Longest Common Subsequence +Broyden-Fletcher-Goldfarb-Shanno Algorithm [91] | - | CTD [74], JS Divergence +Ensemble SVR [96] |
| Traffic flow prediction | - | - | GA-LSTM [86], Independent XGBoost with Spatial Lag [78], STGGAT [85], Trajectory-based RNN [73], STMGG [107], Bayesian LSTM-CNN [23], k-Nearest Neighbor [110], GCN-GRU Federated Learning [15] |
| Queue length estimation | - | Queue Length in Immediate Past Cycle Model + LR [58] | Gaussian Process Interpolation + Markov Chain Monte Carlo +Boundary Constrained Car-following Model [108], Bayesian + KDE [81], Light-weighted Gaussian Process Model [109], RF [51], GMM [82] |
| Queue length prediction | - | - | GRU [42] |
| Link dynamic vehicle count | - | KDE [26] | GRU +Bayesian + Improved Complete Ensemble Empirical Mode Decomposition with Adaptive Noise model [27] |
| Trajectory reconstruction | Weighted Quadratic Function [7], Kinematic Wave +Variational Theory [61], PF-Path Flow Estimator Satisfying SUE principle [100], KSP +Analytic Hierarchy Process with Entropy Weight Method [28], Staircase Vehicle Order-changing Model [16], KSP +extremely randomized trees +TOPSIS [54] | Bayesian Network [72], Bayes' Theorem [9], PF [69] | Semi-supervised Route Choice Model [5], Improved KSP + AE [68], Unsupervised Bayesian Inverse RL [46], Spatio-temporal residual networks + Path Flow Estimator [101] |
| OD estimation | Two-step OLS Model [62] | - | Time-Varying Mixture Model of Discrete-Time Markov Chains [40], GA-CNN [66], Res3D [87] |
| OD prediction | - | - | PCA + KF [53, 112], k-means +Transformer [33] |
| Vehicle activity pattern | - | PCM [95] | k-means [4, 8, 10, 31, 79], Ward's Hierarchical Clustering [30], DBSCAN +CFSDP +Decision Tree [104], k-means + PCA +TD [56] |
| Shared mobility | Integer Programming [44, 55] | - | - |
| Potential travel hotspots | - | - | KSC-TD [99] |
| Traffic congestion and accident analysis | - | LR [14], Geographical and Temporal Weighted Regression [97] | - |
| Companion vehicles | - | - | MCM Model [40] |
| Signal timing inference | - | - | Customized WSMM [109], DBSCAN +Phase WSMM with Average Phase Duration [48], k-means + OLS [102] |
| Subarea partition | Correlation-degree Model +Newman Fast algorithm [37], Similarity analyzes +Aggregation-based Principle [38] | - | - |
| Signal control optimization | MILP Model Named Multipath-Band [98], MPC [45], MILP [13], Multi-objective dynamic programming [57] | - | - |

their usage not only in dynamic linear models (DLM) for predicting travel times but also as the theoretical basis for certain machine learning models. However, the assumption that future states depend solely on the current state may limit its effectiveness in modelling complex traffic scenarios, particularly in prediction tasks.

- Filtering algorithm: in stochastic traffic processes, filtering algorithms help to determine the state of a particular system based on an incomplete and possibly noisy set of observations. In this context, particle filters (PF) are employed for reconstructing trajectories from defective LPR data [69, 100], while Kalman filters (KF) have been proposed for predicting travel time on urban arterials [60] and network-level OD [53, 112]. Also, Dion and Rakha , [17] introduced a dynamic low-pass adaptive filtering algorithm, which effectively extracts travel times from low-sampling-rate LPR data. Dion's Algorithm, known for its robustness and precision in extracting travel time from sparse LPR data, has become a benchmark model for many subsequent studies in this field. The main limitation of filtering algorithms lies in their reliance on several empirical assumptions (or prior knowledge) to handle uncertainty and noise in data sources. However, these assumptions often deviate from real-world conditions, which limits their performance.
- Regression analysis: regression analysis serves as a statistical method for investigating the association between one or multiple independent (explanatory) variables and a dependent (response) variable. As evidenced by Table 1, regression models are extensively utilized in studying factors contributing to traffic congestion and accidents. However, a notable critique of regression analysis is that statistically significant results do not always correspond to meaningful or physically interpretable relationships in real-world phenomena. Consequently, the reliability of analytical outcomes is contingent upon, and confined to the traffic-related explanatory variables extracted through LPR data mining efforts.
- Others: other than the methods mentioned above, DLMs have been proposed in LPR-based travel time estimation and prediction works, while the personalized choice model (PCM) has also shown promising results in characterizing vehicle activity patterns in urban networks (Table 1).

### 4.3 Machine learning methods

As a subfield of artificial intelligence (AI), ML methods, while largely grounded in statistics, are characterized by their capacity to empower models to autonomously discern patterns and regularities from data. In contrast to statistical methods, ML obviates the need for manually crafting all rules. The realm of ML-based research using LPR data can be divided into four categories based on the learning paradigms of ML, including (1) traditional ML, (2) deep learning (DL), (3) reinforcement learning (RL), and (4) hybrid learning (HL).

- Traditional ML: traditional ML models are commonly divided into supervised and unsupervised learning, each demonstrating distinct capabilities in handling different types of tasks. Unsupervised learning models aim to train on large volumes of unlabeled data to uncover inherent patterns, classify data, and perform dimensionality reduction, etc. When combined with algorithms like expectation maximization (EM), models such as lognormal mixture model (LMM), Farlie-Gumbel-Morgenstern (FGM) copula, Gaussian mixture model (GMM), and kernel density estimation (KDE) are extensively used for probability density estimation due to their ability to accurately estimate the parameters of probability distributions. These models find wide applications in estimating travel time distributions and queue lengths using large-scale LPR data. Markov chain mixture (MCM) models, inspired by Markov chain theory, have been applied in companion vehicle identification tasks. Clustering models like adaptive clustering by fast search and find of density peaks (CFSDP) are employed not only in travel time estimation, but also in categorizing and analyzing network-level vehicle activity patterns, along with other clustering methods like k-means, Ward's hierarchical clustering, and density-based spatial clustering of applications with noise (DBSCAN). Furthermore, collaborative tensor decomposition (CTD) and principal component analysis (PCA) are utilized for data dimensionality reduction in tasks such as traffic flow estimation, OD prediction, and vehicle activity pattern recognition. In supervised learning, methods including support vector regression (SVR), linear regression (LR), XGBoost, and Random Forest (RF) serve various purposes in different research directions, encompassing both regression and classification tasks, by extracting features and fitting model parameters on labeled data. Notably, the weighted soft margin maximization (WSMM) method, a variant of support vector machines (SVM), exhibits unique applications in signal timing inference tasks.

- DL: by employing various mechanism to associate and modify layers of neurons within neural networks, DL has made a significant advancement over traditional neural networks, thereby facilitating the extraction of intricate patterns and higher-level features from multidimensional data sources such as LPR data. Different forms for DL are applied across various LPR-based research domains, as indicated in Table 1. For instance, deep neural networks (DNN) like autoencoders (AE) have demonstrated robust and accurate results in trajectory reconstruction works. In network-level studies such like OD estimation, convolutional neural network (CNN), typically used for image recognition and analysis has been proposed. Recurrent neural network (RNN), long short-term memory (LSTM), and gated recurrent unit (GRU), are adept at capturing temporal dependencies within data sequences, making them suitable for tasks like traffic flow and queue length prediction. Recently, Dai et al. [15] integrated a graph convolutional network (GCN) and GRU into a federated learning framework for the task of traffic flow prediction, addressing the data security and privacy issues of different data-providing participants. Nevertheless, the lack of interpretability in DL and the trial-and-error nature of its hyperparameter tuning have been widely criticized.
- RL: RL is designed to facilitate agents' acquisition of optimal policies for a given task by interacting with a specific traffic environment. While referred to as "approximate dynamic programming", RL operates without the need for a complete model of the environment, thus accommodating ongoing adjustments and optimization of decision-making processes. The aforementioned characteristic makes RL advantageous in addressing challenges pertaining to model representation and convergence that are encountered in certain MP approaches. Nevertheless, apart from the approach suggested by Li et al. [46] that combines inverse RL method with Bayesian inference for trajectory reconstruction, RL methods are rarely utilized in LPR-based research.
- HL: by integrating multiple ML models or other techniques, HL enhances the performance of learning models in capturing complex data features and spatiotemporal patterns, leveraging the strengths of diverse methodologies to achieve higher generalization, accuracy, and robustness. Table 1 indicates the prevalent utilization of HL methodologies in capturing spatiotemporal heterogeneity in LPR data for studying traffic states and travel demand. These methods encompass various applications, including LSTM-CNN for travel time and traffic flow prediction, generic algorithm with attention-based LSTM (GA-LSTM) for traffic flow prediction, as well as Spatiotemporal Gated Graph Attention Network (STGGAT) and Spatiotemporal Multigraph Gated Network (STMGG). Additionally, HL methods extend to OD estimation, featuring models such as GA-CNN and 3D convolution-based deep neural network (Res3D). Recently, K-shape clustering transformer-decoder (KSC-TD) has also been utilized for predicting potential travel hotspots. Similar to DL methods, the interpretability issues arising from the "black-box" characteristic of HL models during the training process remain a significant challenge.

### 4.4 Discussion

Figure 3 demonstrates the proportions of three types of methods utilized in LPR-based studies for the last 10 years, as referenced in the literature listed in Table 1. The trend shows that ML has become the dominant method in LPR-based research, while OR and statistical methods are used less frequently. Complex ML frameworks can deliver more precise and robust predictive results by extracting higher-level features, incorporating attention mechanisms, and leveraging temporal analysis capabilities. However, challenges such as limited interpretability, restricted transferability, high computational requirements, and overfitting issues remain significant hurdles to the advancement of these techniques. In the meanwhile, though less popular in recent years, both OR and statistical methods retain distinct modeling logics, characteristics, and advantages for addressing specific transportation issues, supported by their substantial theoretical foundations. Thus, selecting techniques aligned with core transportation principles is crucial for effectively resolving key challenges while minimizing issues of interpretability and transferability of specific methodology.

A synthesis of the reviewed literature suggests that certain LPR-oriented data processing and modeling tasks exhibit clear methodological preferences, indicating that specific techniques have become widely regarded as well-suited for particular research directions. Specifically, the use of statistical distribution-fitting models for estimating travel time reliability has become an established research paradigm. Similar to predictive studies on spatiotemporal series in other domains, various neural network models and their variants have been widely adopted for forecasting traffic states, including travel time, traffic flow, and queue length. In addition, clustering models have consistently proven to be effective tools for classifying and analyzing motor-vehicle travel behavior. Graph

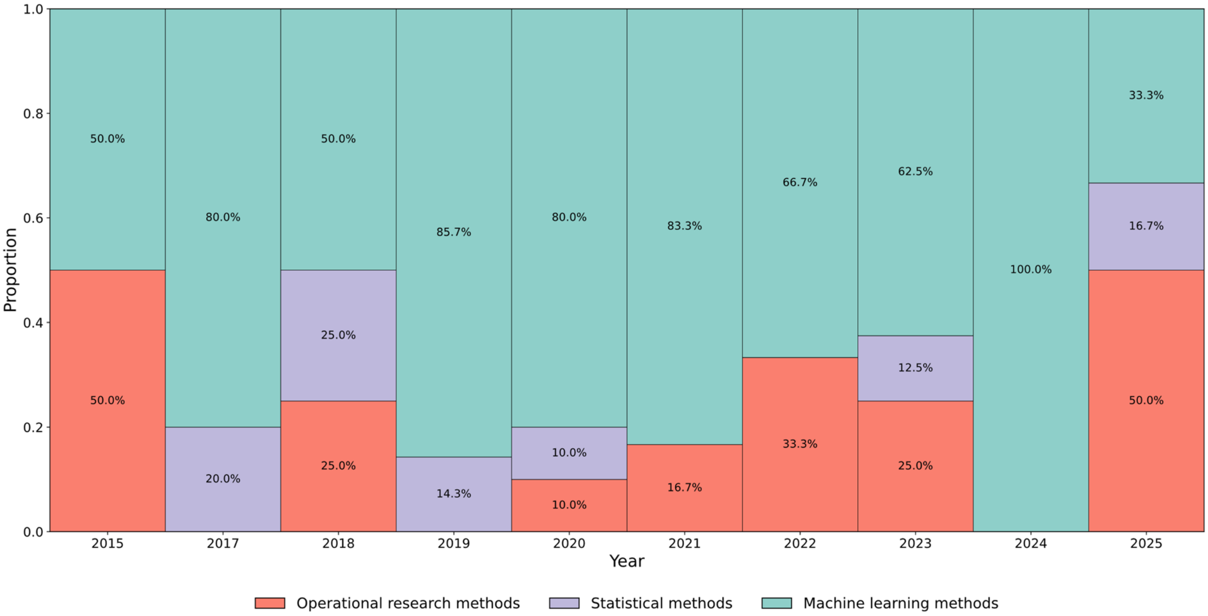

**Fig. 3** Proportions of each method for the last 10 years

partitioning methods are frequently applied in studies on subarea partitioning. Furthermore, WSMM-type models have emerged as benchmark approaches for inferring intersection signal timing plans. These methodological patterns highlight that different approaches inherently exhibit varying levels of suitability for distinct tasks, reflecting their alignment with the underlying nature of the problems they aim to address.

## 5 Open research issues and future opportunities

Building on the synthesis of trends and methodologies in LPR-based research discussed in Sects. 3 and 4, this section highlights several the open research issues and future opportunities for using LPR data.

### 5.1 Explorations in research directions

The challenges and opportunities of LPR data applications lie in the following aspects.

- Full trajectory reconstruction: Although there is considerable research on LPR-based trajectory reconstruction, the majority of studies strongly relies on empirical assumptions or prior knowledge during the reconstruction process (e.g., people tend to choose shorter paths with fewer intersections). These assumptions, although reasonable to some extent, may not necessarily align with all real-world scenarios, limiting the accuracy and completeness of the reconstructed trajectories. Additionally, some intersections or specific entry directions without LPR devices are underrepresented in the reconstructed data, resulting in quasi-complete trajectory reconstruction. Therefore, further investigation into path selection factors in urban mobility and the exploration of optimization methods for trajectory reconstruction are essential for advancing LPR-based studies and achieving full trajectory reconstruction.
- Traffic control optimization: LPR data can be interpreted as trajectory data from a macroscopic perspective, while at the micro level, it contains extensive traffic state information at both network and intersection levels. Therefore, it is evident that LPR data not only provide environmental and state-related information for RL methods, but also offer substantial parameter support for MPC methods in signal control optimization research. Furthermore, the substantial volume of data generated by high-coverage LPR systems is instrumental in developing highly accurate signal timing plans. It also facilitates the segmentation of network subareas according to traffic demand. Specifically, LPR data implementation within this domain reveals several notable research gaps: (1) Current traffic signal control models primarily focus on optimizing control delay. In contrast, the ability of LPR data to provide dynamic, reliable travel time and actual delay information offers a strong foundation for continuously optimizing control strategies through online learning methods with accurate real-time feedback within a closed-loop system. Nevertheless, this area remains underdeveloped. (2) Also, network-level coordination controls are

rarely explored in existing LPR-based studies. (3) In control subarea partition domain, LPR data can provide a new perspective for subarea partitioning optimization given its capability in capturing accurate and comprehensive traffic flow relationships between upstream and downstream intersections compared to floating car data. This capability also offers the potential to develop advanced evaluation metrics and computational methods for subarea partitioning. (4) Research on trajectory-based signal control optimization and the combination of route and signal control optimization based on these data characteristics also holds great potential for exploration.

- Development of proactive traffic management: as a cutting edge and future-oriented traffic operation approach, proactive traffic management requires accurate support in predicting traffic conditions, travel demand, and other related factors. LPR data, with its high coverage and rich embedded information, serves as a highly valuable data source for such prediction tasks. However, Fig. 2 and Table 1 show that the majority of LPR-based studies focus on estimating traffic scenarios (traffic state, OD, accident, etc.), as opposed to prediction. Predictive modelling of traffic scenarios constitutes a crucial aspect in transitioning from passive management to active optimization within ITS, thereby facilitating a significant advancement in transportation system management and operations, transportation emergency management, and even autonomous driving technologies. Specifically, studies on traffic state and demand prediction, OD prediction, travel time distribution prediction, and parking demand prediction, are highly anticipated for future exploration. With the advancement of the aforementioned work, in addition to signal control, related research could also be conducted in areas such as dynamic speed limits, dynamic hard shoulder running, and accurate ramp control.
- Calibration of regional traffic flow model parameters: LPR data encompass rich traffic flow-related information. Recognizing that traffic flow characteristics vary across regions, and the relationships among the fundamental parameters of volume, speed, and density can deviate significantly from classical traffic flow theory. Mining large-scale LPR data, which integrates characteristics of fixed detectors and probe vehicles, enables researches to extract insights into these parameters. This process facilitates the calibration of traffic flow models at the network or arterial level within a specific area, uncovering the interrelationships among these key parameters. Such calibration signifies a deeper understanding of the interactions between traffic demand, traffic conditions, and roadway capacity, providing a strong basis for subsequent research and applications at both macro levels, such as transportation system management and operations, and micro levels, including the identification of traffic anomalies.
- Intelligent vehicle infrastructure cooperative systems (IVICS): the LPR system collects specific vehicle information for individual license plate numbers. Its wide coverage and fine-grained data characteristics give it a significant advantage in the field of IVICS. Based on this, in addition to the aforementioned areas, LPR data also holds potential for research directions such as driving route optimization, cooperative adaptive cruise control, platooning and fleet management, and connected autonomous vehicles.

### 5.2 Innovations in methodologies

Future innovations for methodologies with LPR data could focus not only on emerging techniques such as HL, RL, or large models but also on exploring new simulation modeling paradigms for LPR data.

- HL and RL methods: with the ability to extract complex features from multi-dimensional data and the integration of powerful, scalable training-feedback mechanisms, HL and RL, built upon traditional ML methods, have emerged as increasingly robust, accurate and flexible tools for big data analysis in recent years. Compared to other research areas within ITS, there is still noticeable potential for the application and development of these emerging techniques in LPR-based research. However, it is important to note that the interpretability of learning models tends to decrease as model complexity increases. The selection of suitable "black-box" models for specific issues, or the application of XAI (Explainable AI) techniques, may serve as key driving forces in advancing LPR data applications.
- The utilization of large models: compared to conventional ML methods, large models' parameter scale ranges from billions to trillions, allowing them to handle vast amounts of diverse data. This enables them to tackle more complex tasks and possess superior generalization capabilities. Jin et al. [35] highlighted the introduction of large language models (LLMs) and pre-trained foundation models (PFMs) for time series and spatiotemporal data in transportation field. In this context, LLMs contribute methodological advancements in handling unstructured data and facilitating interactions

between users and transportation systems. Meanwhile, PFMs integrate multi-source data and leverage the extensive knowledge embedded within them to efficiently and accurately handle various transportation tasks such as traffic flow prediction, travel demand analysis, and incident detection. These models integrate datasets from various sources, including loop detectors, video detectors, etc. for traffic flow modelling and prediction. With the development in LLMs and PFMs, coupled with the resolution of data privacy concerns, LPR data has the potential to become a reliable training dataset for large models. Utilizing large models for automated, proactive mining and analysis of multidimensional traffic information from LPR data will elevate the development of ITS to a whole new level.

- LPR-based simulation modeling: current LPR-based studies use real-world LPR data as input for simulation modeling, and the simulation process has been implemented using software such as VISSIM, as applied by Li et al. [45], and Synchro [38]. Since traffic conditions in these simulated environments may differ from the actual scenarios under which LPR data is collected, future efforts could focus on developing traffic simulation platforms that are oriented around real LPR data to test newly developed methodologies. In other words, applying and validating LPR data-driven methods in real-world settings, or developing more realistic LPR-based simulation environments, is crucial for advancing LPR-based studies.

## 6 Conclusions

Given the high penetration rate and cost-effectiveness of LPR data in ITS, its multidimensional information provides valuable support for a wide range of transportation applications, offering both essential data foundations and opportunities for advanced data mining. Although there is still a long way to go before LPR devices are widely used and obtaining LPR data remains challenging due to data privacy and security concerns, the development of vehicle-to-everything (V2X) technology increases the likelihood to obtain data similar to the one obtained by LPR system (i.e., data providing unique vehicle identifiers, timestamps, and widespread detection locations). As a result, LPR-based research is becoming increasingly important.

This paper presents an in-depth survey on the current state of LPR data applications. Specifically, we categorized LPR-based studies into four major topics, each of which is further divided into various research directions. This framework outlines the pros and cons of using LPR data in each area. The technical features, advantages, and limitations of data-driven methods on OR, statistical, and ML methods are summarized within diverse research directions. Finally, an in-depth analysis of the latest trends in LPR-based studies directions and techniques are conducted, providing insights into future opportunities based on research directions and methodologies. This paper aims to provide readers with a professional and systematic reference for understanding the current status of LPR data modelling and implementation, serving as a solid foundation for guiding future research work in this field.

**Abbreviations**

| | |
|---|---|
| LPR | License plate recognition |
| ITS | Intelligent transportation systems |
| AVI | Auto vehicle identification |
| GPS | Global positioning system |
| OBD | On-board diagnostics |
| OD | Origin-destination |
| OR | Operational research |
| ML | Machine learning |
| KSP | K-shortest paths |
| MP | Mathematical programming |
| OLS | Ordinary least square |
| MILP | Mixed integer linear programming |
| MPC | Model predictive control |
| SUE | Stochastic user equilibrium |
| DLM | Dynamic linear model |
| PF | Particle filter |
| KF | Kalman filter |
| PCM | Personalized choice model |
| AI | Artificial intelligence |
| DL | Deep learning |
| RL | Reinforcement learning |
| HL | Hybrid learning |
| LMM | Lognormal mixture model |
| FGM | Farlie-Gumbel-Morgenstern |
| GMM | Gaussian mixture model |
| KDE | Kernel density estimation |
| EM | Expectation maximization |
| MCM | Markov chain mixture |
| CFSDP | Adaptive clustering by fast search and find of density peaks |
| DBSCAN | Density-based spatial clustering of applications with noise |
| CTD | Collaborative tensor decomposition |
| PCA | Principal component analysis |
| SVR | Support vector regression |
| LR | Linear regression |
| RF | Random forest |
| WSMM | Weighted soft margin maximization |
| SVM | Support vector machine |
| DNN | Deep neural networks |
| AE | Autoencoder |
| CNN | Convolutional neural networks |
| RNN | Recurrent neural networks |
| LSTM | Long short-term memory |
| GRU | Gated recurrent unit |
| GCN | Graph convolution network |
| GA-LSTM | Generic algorithm with attention-based LSTM |
| STGGAT | Spatiotemporal gated graph attention network |
| STMGG | Spatiotemporal multigraph gated network |
| Res3D | 3D convolution-based deep neural network |
| KSC-TD | K-shape clustering transformer-decoder |
| IVICS | Intelligent vehicle infrastructure cooperative systems |
| V2X | Vehicle-to-everything |

**Acknowledgments**
Not applicable.

### Author contributions

Ng Jiehui developed the review framework, conducted the analysis and synthesis and wrote the paper. Mo Baichuan contributed in developing introduction, presentation of the results, and provided comments and revisions on the paper. Liao Lyuchao provided comments and revisions on the paper. Li Ruimin contributed in developing future opportunities, presentation of the results, and provided comments and revisions on the paper.

### Funding

This work was supported by The Key Program of National Natural Science Foundation of China [grant number U21B2089], National Natural Science Foundation of China [grant number 62376059], and Beijing Natural Science Foundation [grant number 8242011].

### Data availability

Not applicable.

## Declarations

### Competing interests

No potential conflict of interest was reported by author(s).

Received: 17 January 2025 / Accepted: 24 March 2026

### References

1. An, C., Guo, X., Hong, R., Lu, Z., & Xia, J. (2022). Lane-based traffic arrival pattern estimation using License Plate Recognition data. *IEEE Intelligent Transportation Systems Magazine*, *14*(4), 133–144. https://doi.org/10.1109/mits.2021.3051489
2. An, C., He, Y., Lu, J., Lu, Z., & Xia, J. (2024). One-stage estimation of cyclic arrival rates using license plate recognition data. *Journal of Intelligent Transportation Systems*, 1–14. https://doi.org/10.1080/15472450.2024.2392720
3. Anagnostopoulos, C. N. E., Anagnostopoulos, I. E., Psoroulas, I. D., Loumos, V., & Kayafas, E. (2008). License plate recognition from still images and video sequences: a survey. *IEEE Transactions on Intelligent Transportation Systems*, *9*(3), 377–391. https://doi.org/10.1109/tits.2008.922938
4. Bolaños-Martinez, D., Bermudez-Edo, M., & Garrido, J. L. (2023). Clustering study of vehicle behaviors using License Plate Recognition. In: *Proceedings of the International Conference on Ubiquitous Computing & Ambient Intelligence (UCAmI 2022)* (pp. 784–795). https://doi.org/10.1007/978-3-031-21333-5_77
5. Cao, Q., Ren, G., Li, D., Ma, J., & Li, H. (2020). Semi-supervised route choice modeling with sparse Automatic Vehicle Identification data. *Transportation Research Part C: Emerging Technologies*, *121*. https://doi.org/10.1016/j.trc.2020.102857
6. Castillo, E., Jiménez, P., Menéndez, J. M., & Nogal, M. (2012). A Bayesian method for estimating traffic flows based on plate scanning. *Transportation*, *40*(1), 173–201. https://doi.org/10.1007/s11116-012-9443-4
7. Castillo, E., Menéndez, J. M., & Jiménez, P. (2008). Trip matrix and path flow reconstruction and estimation based on plate scanning and link observations. *Transportation Research Part B: Methodological*, *42*(5), 455–481. https://doi.org/10.1016/j.trb.2007.09.004
8. Chang, Y., Duan, Z., & Yang, D. (2018). Using ALPR data to understand the vehicle use behaviour under TDM measures. *IET Intelligent Transport Systems*, *12*(10), 1264–1270. https://doi.org/10.1049/iet-its.2018.5233
9. Chen, H., Xie, X., Feng, Y., & Zhou, J. (2013). A novel method of trip route estimation based on Vehicle License Plate Recognition system. *Procedia - Social and Behavioral Sciences*, *96*, 643–652. https://doi.org/10.1016/j.sbspro.2013.08.075
10. Chen, H., Yang, C., & Xu, X. (2017). Clustering vehicle temporal and spatial travel behavior using License Plate Recognition data. *Journal of Advanced Transportation*, *2017*, 1–14. https://doi.org/10.1155/2017/1738085
11. Chen, M., Yu, G., Chen, P., & Wang, Y. (2017). A copula-based approach for estimating the travel time reliability of urban arterial. *Transportation Research Part C: Emerging Technologies*, *82*, 1–23. https://doi.org/10.1016/j.trc.2017.06.007
12. Chen, P., Sun, J., & Qi, H. (2016). Estimation of delay variability at signalized intersections for urban arterial performance evaluation. *Journal of Intelligent Transportation Systems*, *21*(2), 94–110. https://doi.org/10.1080/15472450.2016.1216319
13. Chen, P., Xu, J., Mei, Y., & Wei, L. (2025). Network-wide critical routes identification and coordinated control based on automatic vehicle identification data. *Transportation Research Part C: Emerging Technologies*, *172*. https://doi.org/10.1016/j.trc.2025.105019
14. Chow, A. H. F., Santacreu, A., Tsapakis, I., Tanasaranond, G., & Cheng, T. (2013). Empirical assessment of urban traffic congestion. *Journal of Advanced Transportation*, *48*(8), 1000–1016. https://doi.org/10.1002/atr.1241
15. Dai, G., Tang, J., Zeng, J., Hu, C., & Zhao, C. (2024). Road network traffic flow prediction: A personalized federated learning method based on client reputation. *Computers and Electrical Engineering*, *120*. https://doi.org/10.1016/j.compeleceng.2024.109678
16. Deng, Y., Cao, Q., Ren, G., Ma, J., & Zhu, S. (2023). Vehicle trajectory reconstruction incorporating probe and fixed sensor data. *Journal of Transportation Engineering Part A: Systems*, *149*(9). https://doi.org/10.1061/jtepbs.Teeng-7788
17. Dion, F., & Rakha, H. (2006). Estimating dynamic roadway travel times using Automatic Vehicle Identification data for low sampling rates. *Transportation Research Part B: Methodological*, *40*(9), 745–766. https://doi.org/10.1016/j.trb.2005.10.002
18. Du, J., Cui, J., Ren, G., & Thompson, R. G. (2025). Recognizing the traffic state of urban road networks: a resilience-based data-driven approach. *Transportation Research Record: Journal of the Transportation Research Board*, *2679*(6), 183–202. https://doi.org/10.1177/03611981241312914
19. Du, S., Ibrahim, M., Shehata, M., & Badawy, W. (2013). Automatic license plate recognition (ALPR): a state-of-the-art review. *IEEE Transactions on Circuits and Systems for Video Technology*, *23*(2), 311–325. https://doi.org/10.1109/tcsvt.2012.2203741
20. Du, W., Sun, B., Kuai, J., Xie, J., Yu, J., Sun, T., & Hu, Q. (2021). Highway travel time prediction of segments based on ANPR data considering traffic diversion. *Journal of Advanced Transportation*, *2021*, 1–16. https://doi.org/10.1155/2021/9512501
21. Fei, X., Zhang, Y., Liu, K., & Guo, M. (2013). Bayesian dynamic linear model with switching for real-time short-term freeway travel time prediction with License Plate Recognition data. *Journal of Transportation Engineering*, *139*(11), 1058–1067. https://doi.org/10.1061/(asce)te.1943-5436.0000538
22. Fu, F., Qian, W., & Dong, H. (2019). Estimation of route travel time distribution with information fusion from automatic number plate recognition data. *Journal of Transportation Engineering Part A: Systems*, *145*(7). https://doi.org/10.1061/jtepbs.0000242
23. Fu, F., Wang, D., Sun, M., Xie, R., & Cai, Z. (2024). Urban traffic flow prediction based on bayesian deep learning considering optimal aggregation time interval. *Sustainability*, *16*(5). https://doi.org/10.3390/su16051818
24. Gao, Y., Nie, Y., Yang, W., Xu, H., & Huang, L. (2018). Anomalous trajectory detection between regions of interest based on ANPR system. In: *Computational Science – ICCS 2018* (pp. 631–643). https://doi.org/10.1007/978-3-319-93701-4_50
25. Hadavi, S., Rai, H. B., Verlinde, S., Huang, H., Macharis, C., & Guns, T. (2020). Analyzing passenger and freight vehicle movements from automatic-Number plate recognition camera data. *European Transport Research Review*, *12*(1). https://doi.org/10.1186/s12544-020-00405-x
26. He, C., Wang, D., Chen, M., Qian, G., & Cai, Z. (2021). Link dynamic vehicle count estimation based on travel time distribution using License Plate Recognition data. *Transportmetrica A: Transport Science*, *19*(2). https://doi.org/10.1080/23249935.2021.2012299
27. He, C., Wang, D., Yu, Y., Cai, Z., & Lee, Y. J. (2023). A hybrid deep learning model for link dynamic vehicle count forecasting with Bayesian optimization. *Journal of Advanced Transportation*, *2023*, 1–15. https://doi.org/10.1155/2023/5070504
28. He, M., Wu, X., Zhan, Z., Guo, J., & Yao, Z. (2022). Vehicle path recognition approach based on incomplete automatic vehicle identification. *Journal of Advanced Transportation*, *2022*, 1–12. https://doi.org/10.1155/2022/3462267
29. He, Z., Chen, K., Chen, X., & Sun, W. (2018). Introduction and analysis of a large-scale benchmark Automatic Vehicle Identification dataset. *International Conference on Transportation and Development 2018*, 35–43. https://doi.org/10.1061/9780784481530.004
30. Hong, R., Rao, W., Zhou, D., An, C., Lu, Z., & Xia, J. (2020). Commuting pattern recognition using a systematic cluster framework. *Sustainability*, *12*(5). https://doi.org/10.3390/su12051764
31. Hong, R., Zhou, D., An, C., Rao, W., & Xia, J. (2019). *Demand structure snalysis for urban traffic using Automatic License Plate Recognition data*. https://doi.org/10.1061/9780784482292.441

32. Hu, C., Tang, J., Wang, Y., Li, Z., & Dai, G. (2025). A flexible road network partitioning framework for traffic management via graph contrastive learning and multi-objective optimization. *Computer-Aided Civil and Infrastructure Engineering*, *40*(18), 2671–2694. https://doi.org/10.1111/mice.13454
33. Huang, S., Zhang, C., Zhao, J., & Han, Y. (2025). Traffic origin–destination flow prediction considering individual travel frequency: a classification-based approach. *Journal of Transportation Engineering Part A: Systems*, *151*(2). https://doi.org/10.1061/jtepbs.Teeng-8704
34. Jenelius, E., & Chen, M. (2022). Traveller recurrence and inter- versus intra-traveller speed variability: analysis with bluetooth data. *Journal of Advanced Transportation*, *2022*, 1–10. https://doi.org/10.1155/2022/6650310
35. Jin, M., Wen, Q., Liang, Y., Zhang, C., Xue, S., Wang, X., Zhang, J., Wang, Y., Chen, H., & Li, X. (2023). Large models for time series and spatio-temporal data: A survey and outlook. *arXiv preprint arXiv:2310.10196*.
36. Kazagli, E., & Koutsopoulos, H. N. (2013). Estimation of arterial travel time from Automatic Number Plate Recognition data. *Transportation Research Record: Journal of the Transportation Research Board*, *2391*(1), 22–31. https://doi.org/10.3141/2391-03
37. Ke, S., Liu, W., Lu, Z., Rao, W., An, C., & Xia, J. (2022). Subarea partition based on correlation analysis with edge-elimination strategy using Automatic License Plate Recognition data. *Transportation Research Record: Journal of the Transportation Research Board*, *2676*(12), 641–652. https://doi.org/10.1177/03611981221096439
38. Ke, S., Wang, Y., An, C., Lu, Z., & Xia, J. (2024). Traffic origin-destination flow-inspired dynamic urban arterial partition for coordinated signal control using Automatic License Plate Recognition data. *IEEE Intelligent Transportation Systems Magazine*, *16*(1), 132–147. https://doi.org/10.1109/mits.2023.3313621
39. Laharotte, P. A., Billot, R., Come, E., Oukhellou, L., Nantes, A., & Faouzi, E., N.-E (2015). Spatiotemporal analysis of bluetooth data: application to a large urban network. *IEEE Transactions on Intelligent Transportation Systems*, *16*(3), 1439–1448. https://doi.org/10.1109/tits.2014.2367165
40. Lawlor, S. (2017). *Traffic estimation and detection methods utilizing Automatic Vehicle Identification systems*. McGill University (Canada).
41. Li, A., Lam, W. H. K., Ma, W., Wong, S. C., Chow, A. H. F., & Tam, L., M (2024). Real-time estimation of multi-class path travel times using multi-source traffic data. *Expert Systems with Applications*, 237. https://doi.org/10.1016/j.eswa.2023.121613
42. Li, B., Yu, Z., Huang, L., & Guo, B. (2019). Vehicle departure pattern and queue length prediction at an isolated intersection with Automatic Vehicle Identity detection. *IET Intelligent Transport Systems*, *13*(12), 1804–1813. https://doi.org/10.1049/iet-its.2019.0117
43. Li, M., Tang, J., Chen, Q., & Liu, Y. (2023). Traffic arrival pattern estimation at urban intersection using License Plate Recognition data. *Physica A: Statistical Mechanics and its Applications*, *625*. https://doi.org/10.1016/j.physa.2023.128995
44. Li, R., Liu, Z., & Zhang, R. (2018). Studying the benefits of carpooling in an urban area using Automatic Vehicle Identification data. *Transportation Research Part C: Emerging Technologies*, *93*, 367–380. https://doi.org/10.1016/j.trc.2018.06.012
45. Li, R., Wang, S., Jiao, P., & Lin, S. (2023). Traffic control optimization strategy based on License Plate Recognition data. *Journal of Traffic and Transportation Engineering (English Edition)*, *10*(1), 45–57. https://doi.org/10.1016/j.jtte.2021.12.004
46. Li, W., Liu, C., Wang, T., & Ji, Y. (2022). An innovative supervised learning structure for trajectory reconstruction of sparse LPR data. *Transportation*, *51*(1), 73–97. https://doi.org/10.1007/s11116-022-10320-w
47. Li, Y. (2008). Short-term prediction of motorway travel time using ANPR and loop data. *Journal of Forecasting*, *27*(6), 507–517. https://doi.org/10.1002/for.1070
48. Lin, Q., Chen, J., Li, G., & He, Z. (2022). Signal timing parameters inference method at intersections using License Plate Recognition data. *IET Intelligent Transport Systems*, *16*(8), 1092–1107. https://doi.org/10.1049/itr2.12198
49. Lin, Q., Huang, W., Wu, Z., Zhang, M., & He, Z. (2025). Multi-agent game theory-based coordinated ramp metering method for urban expressways with multi-bottleneck. *IEEE Transactions on Intelligent Transportation Systems*, *26*(3), 3643–3658. https://doi.org/10.1109/tits.2024.3521460
50. Liu, C., Wang, X., Zhu, M., & Han, Y. (2016). Discovering companion vehicles from live streaming traffic data. In *Web Technologies and Applications* (pp. 116–128). https://doi.org/10.1007/978-3-319-45814-4_10
51. Liu, D., An, C., Yasir, M., Lu, J., & Xia, J. (2022). A machine learning based method for real-time queue length estimation using License Plate Recognition and GPS trajectory data. *KSCE Journal of Civil Engineering*, *26*(5), 2408–2419. https://doi.org/10.1007/s12205-022-0451-4
52. Liu, F., Zeng, J., Tang, J., & Yu, T. (2025). Travel frequent-route identification based on the snake algorithm using license plate recognition data. *Mathematics*, *13*(15). https://doi.org/10.3390/math13152536
53. Liu, J., Zheng, F., van Zuylen, H. J., & Li, J. (2020). A dynamic OD prediction approach for urban networks based on Automatic Number Plate Recognition data. *Transportation Research Procedia*, *47*, 601–608. https://doi.org/10.1016/j.trpro.2020.03.137
54. Liu, Y., Chen, X., Hu, R., Huang, W., Li, L., Xie, J., & Yao, D. (2025). Vehicle trajectory reconstruction with sparse automatic license plate recognition data for urban road networks. *Transportmetrica A: Transport Science*, 1–24. https://doi.org/10.1080/23249935.2024.2449389
55. Liu, Z., Li, R., & Dai, J. (2022). Effects and feasibility of shared mobility with shared autonomous vehicles: An investigation based on data-driven modeling approach. *Transportation Research Part A: Policy and Practice*, *156*, 206–226. https://doi.org/10.1016/j.tra.2022.01.001
56. Liu, Z., Wang, X., Bi, Y., Kong, J., Xu, R., Chen, Y., Tang, J., & Zhang, G. (2022). Travel patterns analysis using tensor-based model from large-scale License Plate Recognition data. *Journal of Advanced Transportation*, *2022*, 1–14. https://doi.org/10.1155/2022/3930795
57. Luo, L., Tan, C., Tang, K., & Zhu, H. (2025). A distributed model predictive approach for network traffic signal control using multi-objective dynamic programming. *Computer-Aided Civil and Infrastructure Engineering*, *40*(24), 3953–3978. https://doi.org/10.1111/mice.70035
58. Luo, X., Ma, D., Jin, S., Gong, Y., & Wang, D. (2019). Queue length estimation for signalized intersections using License Plate Recognition data. *IEEE Intelligent Transportation Systems Magazine*, *11*(3), 209–220. https://doi.org/10.1109/mits.2019.2919541
59. Luo, X., Wang, D., Ma, D., & Jin, S. (2019). Grouped travel time estimation in signalized arterials using point-to-point detectors. *Transportation Research Part B: Methodological*, *130*, 130–151. https://doi.org/10.1016/j.trb.2019.10.007
60. Ma, X., Khoury, A., F., & Jin, J. (2017). Prediction of arterial travel time considering delay in vehicle re-identification. *Transportation Research Procedia*, *22*, 625–634. https://doi.org/10.1016/j.trpro.2017.03.056
61. Mehran, B., Kuwahara, M., & Naznin, F. (2011). Implementing kinematic wave theory to reconstruct vehicle trajectories from fixed and probe sensor data. *Procedia - Social and Behavioral Sciences*, *17*, 247–268. https://doi.org/10.1016/j.sbspro.2011.04.517
62. Mo, B., Li, R., & Dai, J. (2020). Estimating dynamic origin–destination demand: A hybrid framework using License Plate Recognition Data. *Computer-Aided Civil and Infrastructure Engineering*, *35*(7), 734–752. https://doi.org/10.1111/mice.12526
63. Mo, B., Li, R., & Zhan, X. (2017). Speed profile estimation using License Plate Recognition data. *Transportation Research Part C: Emerging Technologies*, *82*, 358–378. https://doi.org/10.1016/j.trc.2017.07.006
64. Nantes, A., Ngoduy, D., Miska, M., & Chung, E. (2015). Probabilistic travel time progression and its application to Automatic Vehicle Identification data. *Transportation Research Part B: Methodological*, *81*, 131–145. https://doi.org/10.1016/j.trb.2015.09.001
65. Nie, C., Wei, H., Shi, J., & Zhang, M. (2021). Optimizing actuated traffic signal control using License Plate Recognition data: Methods for modeling and algorithm development. *Transportation Research Interdisciplinary Perspectives*, *9*. https://doi.org/10.1016/j.trip.2021.100319
66. Ou, J., Lu, J., Xia, J., An, C., & Lu, Z. (2019). Learn, assign, and search: real-time estimation of dynamic origin-destination flows using machine learning algorithms. *IEEE Access : Practical Innovations, Open Solutions*, *7*, 26967–26983. https://doi.org/10.1109/access.2019.2901289
67. Pereira, G., & Raetzsch, C. (2022). From banal surveillance to function creep: automated license plate recognition (ALPR) in Denmark. *Surveillance & Society*, *20*(3), 265–280.
68. Qi, X., Ji, Y., Li, W., & Zhang, S. (2021). Vehicle trajectory reconstruction on urban traffic network using Automatic License Plate Recognition data. *IEEE Access : Practical Innovations, Open Solutions*, *9*, 49110–49120. https://doi.org/10.1109/access.2021.3068964
69. Rao, W., Wu, Y. J., Xia, J., Ou, J., & Kluger, R. (2018). Origin-destination pattern estimation based on trajectory reconstruction using Automatic License Plate Recognition data. *Transportation Research Part C: Emerging Technologies*, *95*, 29–46. https://doi.org/10.1016/j.trc.2018.07.002
70. Robinson, A. (2019). Validating traffic models using large-scale Automatic Number Plate Recognition (ANPR) data. *Journal of the South African Institution of Civil Engineering*, *61*(3), 45–57.

71. Robinson, S., & Polak, J. (2006). Overtaking rule method for the cleaning of matched license-plate data. *Journal of Transportation Engineering*, *132*(8), 609–617. https://doi.org/10.1061/(ASCE)0733-947X(2006)132:8(609).
72. Sánchez-Cambronero, S., Castillo, E., Menéndez, J. M., & Jiménez, P. (2011). Dealing with error recovery in traffic flow prediction using bayesian networks based on license plate scanning data. *Journal of Transportation Engineering*, *137*(9), 615–629. https://doi.org/10.1061/(asce)te.1943-5436.0000249
73. Shan, X., Yu, W., Li, Z., Wang, C., Ren, Y., & Zhang, J. (2023). Vehicle trajectory-based traffic volume prediction on urban roads with fast-communication License Plate Recognition data. *IEEE Transactions on Intelligent Transportation Systems*, 1–11. https://doi.org/10.1109/tits.2023.3279332
74. Shao, W., & Chen, L. (2018). License Plate Recognition data-based traffic volume estimation using collaborative tensor decomposition. *IEEE Transactions on Intelligent Transportation Systems*, *19*(11), 3439–3448. https://doi.org/10.1109/tits.2017.2780093
75. Song, H., Liang, H., Li, H., Dai, Z., & Yun, X. (2019). Vision-based vehicle detection and counting system using deep learning in highway scenes. *European Transport Research Review*, *11*(1). https://doi.org/10.1186/s12544-019-0390-4
76. Song, J., Hellinga, B., Cao, Q., & Ren, G. (2025). Queue length estimation on urban signalized intersection combining automatic vehicle identification and vehicle trajectory data. *Journal of Transportation Engineering Part A: Systems*, *151*(1). https://doi.org/10.1061/jtepbs.Teeng-8541
77. Song, J., Hellinga, B., Ren, G., Yuan, J., Cao, Q., Ma, J., & Deng, Y. (2025). Traffic state estimation and prediction based on Bayesian approach in urban road networks using AVI and floating vehicle data. *Transportmetrica A: Transport Science*, 1–37. https://doi.org/10.1080/23249935.2025.2517306
78. Sun, B., Sun, T., Jiao, P., & Tang, J. (2021). Spatio-temporal segmented traffic flow prediction with ANPRS data based on improved XGBoost. *Journal of Advanced Transportation*, *2021*, 1–24. https://doi.org/10.1155/2021/5559562
79. Sun, Y., Zhou, X., Sun, L., & Chen, S. (2014). Vehicle activity analysis based on ANPR system. *2014 12th IEEE International Conference on Embedded and Ubiquitous Computing*, 89–96. https://doi.org/10.1109/euc.2014.22
80. Sun, Y., Zhu, H., Liao, Y., & Sun, L. (2015). 6–10 Dec. 2015). Vehicle anomaly detection based on trajectory data of ANPR system. 2015 IEEE Global Communications Conference (GLOBECOM), 1–6. https://doi.org/10.1109/GLOCOM.2015.7417520
81. Tan,C., Liu, L., Wu, H., Cao, Y., & Tang, K. (2020). Fuzing License Plate Recognition data and vehicle trajectory data for lane-based queue length estimation at signalized intersections. *Journal of Intelligent Transportation Systems*, *24*(5), 449–466. https://doi.org/10.1080/15472450.2020.1732217
82. Tan, C., Wu, H., Tang, K., Tan, C., & Jin, S. (2022). An extendable gaussian mixture model for lane-based queue length estimation based on License Plate Recognition data. *Journal of Advanced Transportation*, *2022*, 1–14. https://doi.org/10.1155/2022/5119209
83. Tan, W. (2021). *Development of a cloud-based traffic diagnosis and management laboratory based on high-coverage ALPR* (Publication Number 28540465) [Ph.D., The University of Wisconsin - Milwaukee]. ProQuest Dissertations & Theses Global. United States -- Wisconsin. https://www.proquest.com/dissertations-theses/development-cloud-based-traffic-diagnosis/docview/2546645007/se-2?accountid=14426
84. Tang, J., Wan, L., Schooling, J., Zhao, P., Chen, J., & Wei, S. (2022). Automatic Number Plate Recognition (ANPR) in smart cities: A systematic review on technological advancements and application cases. *Cities*, *129*. https://doi.org/10.1016/j.cities.2022.103833
85. Tang, J., & Zeng, J. (2021). Spatiotemporal gated graph attention network for urban traffic flow prediction based on License Plate Recognition data. *Computer-Aided Civil and Infrastructure Engineering*, *37*(1), 3–23. https://doi.org/10.1111/mice.12688
86. Tang, J., Zeng, J., Wang, Y., Yuan, H., Liu, F., & Huang, H. (2020). Traffic flow prediction on urban road network based on License Plate Recognition data: combining attention-LSTM with Genetic Algorithm. *Transportmetrica A: Transport Science*, *17*(4), 1217–1243. https://doi.org/10.1080/23249935.2020.1845250
87. Tang, K., Cao, Y., Chen, C., Yao, J., Tan, C., & Sun, J. (2020). Dynamic origin-destination flow estimation using Automatic Vehicle Identification data: A 3D convolutional neural network approach. *Computer-Aided Civil and Infrastructure Engineering*, *36*(1), 30–46. https://doi.org/10.1111/mice.12559
88. Tang, K., Wu, H., Yao, J., Tan, C., & Ji, Y. (2021). Lane-based queue length estimation at signalized intersections using single-section License Plate Recognition data. *Transportmetrica B: Transport Dynamics*, *10*(1), 293–311. https://doi.org/10.1080/21680566.2021.1991504
89. Wan, L., Tang, J., Wang, L., & Schooling, J. (2021). Understanding non-commuting travel demand of car commuters – Insights from ANPR trip chain data in Cambridge. *Transport Policy*, *106*, 76–87. https://doi.org/10.1016/j.tranpol.2021.03.021
90. Wang, D., Guo, J., & Cai, Z. (2021). Real time estimation of saturated flow rate of mixed traffic flow based on Automatic License Plate Recognition data. *Journal of Transportation Systems Engineering and Information Technology*, *21*(02), 37–43. https://doi.org/10.16097/j.cnki.1009-6744.2021.02.006
91. Wang, P., Lai, J., Huang, Z., Tan, Q., & Lin, T. (2021). Estimating traffic flow in large road networks based on multi-source traffic data. *IEEE Transactions on Intelligent Transportation Systems*, *22*(9), 5672–5683. https://doi.org/10.1109/tits.2020.2988801
92. Wei, H., Nie, C., Shi, J., & Zhang, M. (2021). *Vehicular license plate detection data fusion and algorithms for dynamic traffic signal control*. https://doi.org/10.1061/9780784483534.036
93. Wu, H., Luo, L., Oguchi, T., Tang, K., & Zhu, H. (2024). Stochastic queue profile estimation using license plate recognition data. *Physica A: Statistical Mechanics and its Applications*, *643*. https://doi.org/10.1016/j.physa.2024.129790
94. Wu, H., Yao, J., Cao, Y., & Tang, K. (2024). Left-turn queue spillback identification based on single-section license plate recognition data. *Transportmetrica B: Transport Dynamics*, *12*(1). https://doi.org/10.1080/21680566.2024.2363273
95. Xie, Y., Zhang, Y., Akkinepally, A. P., & Ben-Akiva, M. (2020). Personalized choice model for managed lane travel behavior. *Transportation Research Record: Journal of the Transportation Research Board*, *2674*(7), 442–455. https://doi.org/10.1177/0361198120923355
96. Xing, J. (2022). Estimating traffic volumes in an urban network based on taxi GPS and limited LPR data using machine learning techniques. *2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC)*, *20*(25). https://doi.org/10.1109/itsc55140.2022.9922445
97. Xu, C., Zhang, Z., Fu, F., Yao, W., Su, H., Hu, Y., Rong, D., & Jin, S. (2023). Analysis of spatiotemporal factors affecting traffic safety based on multisource data fusion. *Journal of Transportation Engineering Part A: Systems*, *149*(10). https://doi.org/10.1061/jtepbs.Teeng-7990
98. Xu, L., Xu, J., Qu, X., & Jin, S. (2022). An origin-destination demands-based multipath-band approach to time-varying arterial coordination. *IEEE Transactions on Intelligent Transportation Systems*, *23*(10), 17784–17800. https://doi.org/10.1109/tits.2022.3150977
99. Yang, H., Yan, C., Chen, Z., & Wang, P. (2024). A K-shape clustering based transformer-decoder model for predicting multi-step potentials of urban mobility field. *IEEE Transactions on Intelligent Transportation Systems*, 1–15. https://doi.org/10.1109/tits.2024.3355211
100. Yang, J., & Sun, J. (2015). Vehicle path reconstruction using Automatic Vehicle Identification data: An integrated particle filter and path flow estimator. *Transportation Research Part C: Emerging Technologies*, *58*, 107–126. https://doi.org/10.1016/j.trc.2015.07.003
101. Yang, J., Sun, J., Cao, Y., Yang, H., & Huang, W. (2025). Path flow reconstruction under sparse automatic vehicle identification coverage and limited probe vehicle penetration. *Transportmetrica B: Transport Dynamics*, *13*(1). https://doi.org/10.1080/21680566.2025.2582570
102. Yao, J., Wu, H., & Tang, K. (2023). A bi-Level programming method for SPaT estimation at fixed-time controlled intersections using License Plate Recognition data. *Transportmetrica B: Transport Dynamics*, *11*(1), 1045–1070. https://doi.org/10.1080/21680566.2023.2165191
103. Yao, W., Chen, C., Su, H., Chen, N., Jin, S., Bai, C., & Fang, Z. (2022). Analysis of key commuting routes based on spatiotemporal trip chain. *Journal of Advanced Transportation*, *2022*, 1–15. https://doi.org/10.1155/2022/6044540
104. Yao, W., Zhang, M., Jin, S., & Ma, D. (2021). Understanding vehicles commuting pattern based on License Plate Recognition data. *Transportation Research Part C: Emerging Technologies*, *128*. https://doi.org/10.1016/j.trc.2021.103142
105. Ye, P., & Pan, E. (2025). Adaptive filtering algorithm for estimating signalized arterial travel times. *Transportation Research Record: Journal of the Transportation Research Board*, *2679*(9), 769–787. https://doi.org/10.1177/03611981251340379
106. Yu, H., Yang, S., Wu, Z., & Ma, X. (2018). Vehicle trajectory reconstruction from Automatic License Plate Reader data. *International Journal of Distributed Sensor Networks*, *14*(2). https://doi.org/10.1177/1550147718755637
107. Zeng, J., & Tang, J. (2023). Modeling dynamic traffic flow as visibility graphs: A network-scale prediction framework for lane-Level traffic flow based on LPR data. *IEEE Transactions on Intelligent Transportation Systems*, *24*(4), 4173–4188. https://doi.org/10.1109/tits.2022.3231959

108. Zhan, X., Li, R., & Ukkusuri, S. V. (2015). Lane-based real-time queue length estimation using License Plate Recognition data. *Transportation Research Part C: Emerging Technologies*, *57*, 85–102. https://doi.org/10.1016/j.trc.2015.06.001
109. Zhan, X., Li, R., & Ukkusuri, S. V. (2020). Link-based traffic state estimation and prediction for arterial networks using License-Plate Recognition data. *Transportation Research Part C: Emerging Technologies*, *117*. https://doi.org/10.1016/j.trc.2020.102660
110. Zhang, S., Jiang, Q., Li, H., Cao, B., & Fan, J. (2024). PURP: a scalable system for predicting short-term urban traffic flow based on license plate recognition data. *Big Data Mining and Analytics*, *7*(1), 171–187. https://doi.org/10.26599/bdma.2023.9020017
111. Zhao, Z., Lu, S., & Han, Y. (2017). Similar trajectory query method based on massive Vehicle License Plate Recognition data. *Journal of Tsinghua University*, *57*(02), 220–224. https://doi.org/10.16511/j.cnki.qhdxxb.2017.22.018
112. Zheng, F., Liu, J., van Zuylen, H., Wang, K., Liu, X., & Li, J. (2019). 27–30 Oct. 2019). Dynamic OD prediction for urban networks based on Automatic Number Plate Recognition data: Paramertic vs. non-parametric approaches. In *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, 4037–4042. https://doi.org/10.1109/ITSC.2019.8917229
113. Zheng, F., Liu, X., Zuylen, H., Li, J., & Lu, C. (2017). Travel time reliability for urban networks: Modelling and empirics. *Journal of Advanced Transportation*, *2017*, 1–13. https://doi.org/10.1155/2017/9147356
114. Zhu, M., Liu, C., Wang, X., & Han, Y. (2017). Approach to discover companion pattern based on ANPR data stream. *Journal of Software*, *28*(06), 1498–1515. https://doi.org/10.13328/j.cnki.jos.005220

## Publisher's note