Road Traffic Forecasting: Recent Advances and New Challenges

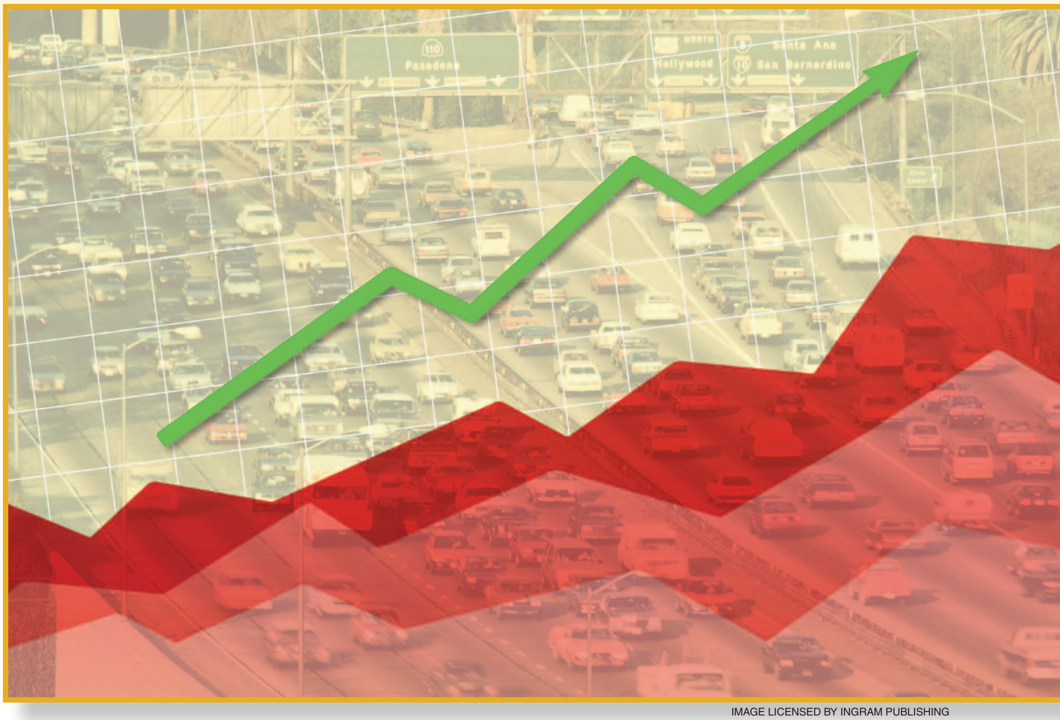

IMAGE LICENSED BY INGRAM PUBLISHING

Abstract—Due to its paramount relevance in transport planning and logistics, road traffic forecasting has been a subject of active research within the engineering community for more than 40 years. In the beginning most approaches relied on autoregressive models and other analysis methods suited for time series data. More recently, the development of new technology, platforms and techniques for massive data processing under the Big Data umbrella, the availability of data from multiple sources fostered by the Open Data philosophy and an ever-growing need

Ibai Laña
TECNALIA, Derio, Bizkaia, Spain, (E-mails: ibai.lana@tecnalia.com)

Javier Del Ser
TECNALIA, Derio, Bizkaia, Spain, Department of Communications Engineering, University of the Basque Country, Bizkaia, Spain, Basque Center for Applied Mathematics (BCAM), Bizkaia, Spain (E-mails: javier.delser@tecnalia.com, javier.delser@ehu.eus)

Manuel Vélez
Department of Communications Engineering, University of the Basque Country, Bizkaia, Spain (E-mails: manuel.velez@ehu.eus)

Eleni I. Vlahogianni
Dept. of Transportation Planning and Engineering, National Technical University of Athens, GR-15773 Athens, Greece (E-mail: elenivl@central.ntua.gr)

Digital Object Identifier 10.1109/MITS.2018.2806634
Date of publication: 20 April 2018

of decision makers for accurate traffic predictions have shifted the spotlight to data-driven procedures. This paper aims to summarize the efforts made to date in previous related surveys towards extracting the main comparing criteria and challenges in this field. A review of the latest technical achievements in this field is also provided, along with an insightful update of the main technical challenges that remain unsolved. The ultimate goal of this work is to set an updated, thorough, rigorous compilation of prior literature around traffic prediction models so as to motivate and guide future research on this vibrant field.

I. Introduction

Issues related to road traffic conditions are common practice in every major city around the world. Traffic congestion leads to social, economic and environmental problems, rationale for which public and private organizations have attempted at addressing them for more than 50 years. Efforts devoted to mitigate the effects of traffic congestion have been conducted in three directions [1]: increasing infrastructures, promoting transport alternatives and managing traffic flows. While the first direction is limited by topographical, budgetary and social factors and the second is mainly a matter of public policies, the latter has been continuously improving in the last decades with the expansion of data provided by sensors in roads and vehicles, as well as with the technology required to exploit those data. This allows measuring, modeling and interpreting traffic features such as flow, occupancy or travel times, which are useful to develop advanced traffic management systems (ATMS) and advanced traveller information systems (ATIS).

Data-driven traffic forecasting has been a research topic since the late 1970s. The first attempts at predicting traffic flows consisted mainly of time-series approaches with different techniques [2]–[5], as well as early explorations of Kalman filtering methods [6]. In any of the above, or in posterior attempts to predict traffic features, short-term prediction horizons were used. Since then, data availability, analysis methods, and computational capacity have evolved and grown remarkably, along with the interest of the research community in this field. Nowadays, Intelligent Transportation Systems (ITS) conform a vivid area of research, policy making and technology development, for which one of its foundations is predicting traffic features [7], [8]. Data from road sensors are available with fine-grained resolution, not only posted in regularly updated public repositories, but also made accessible in the form of floating car data, pedestrian mobility traces, bicycle traffic counts, or traffic lights operation conditions. Forecasting methods have evolved at a similar pace; although a great share of the latest research contributions still relies on time-series analysis, there is also a wide focus on machine learning methods. Simulation tools have proliferated likewise, allowing for network-wide traffic flow forecasts among many other functionalities. However, specific forecasting aspects such as the prediction horizon have remained essentially the same for decades. Roughly all literature is oriented towards short-term prediction, despite the usefulness of long-term estimations for road administration purposes and/or for enhancing short-term models as an additional input feature [7], [9]. Besides, implementation challenges in Big Data architectures are vaguely addressed in this discipline of study.

This manuscript systematically examines the recent developments gravitating on data-driven traffic forecasting methods, with an emphasis on the technical advances made in this dynamic field, the degree of achievement of the different technical challenges posed in the past, and a critical diagnosis of the unexplored areas of research that should be targeted by the community in forthcoming years. To this end the review capitalizes on recent surveys and enriches them with an analysis of the research trends, detected issues and identified challenges springing from the newest works in this field. The survey ends up by tracing the research niches that deserve further attention and efforts, solidly founded on the arrival of methodologies and tools related to Big Data, as well as on the characteristics of this particular data source. The discussions held through this article are intended to provide a comprehensive report of the current state of the art of traffic forecasting for early researchers and engineers, as well as to stimulate and steer future technical contributions towards directions of lacking maturity and reasoned potential.

II. Short-Term Traffic Forecasting: A Historical Perspective

The research activity and contributions dedicated to the development of traffic forecasting methods during the last three decades are assorted and can be classified under very diverse criteria. Indeed, a selection of works in the last ten years have sorted and classified the existing literature on traffic forecasting by adopting very diverging perspectives. As such, Van Arem et al. in [1] explored applications of traffic forecasting to dynamic traffic management by analyzing the overall process from an economical demand and supply approach; the first can be represented by origin-destination flows and is considered a human behavior concern, and the latter stands for the ability of the road network to satisfy the demand. This human aspect of traffic processes is relevant for ATMS and the forecasting methods themselves, which need to have in account that larger demands imply lower supplies (congested roads), which in turn leads to lower demands due to informed drivers who rearrange their planned routes. This noted impact of traffic conditions on their near-future selves and the stochastic nature of traffic [1], [10], [11] contribute to the fact that short-term forecasting is the main research focus of this and practically every other work on this subject.

Aside from the prediction horizon, this early review is concentrated on forecasting methodologies, performance evaluation techniques and real-world application examples. Besides, the authors provide a series of challenges that include representation, model validation, incorporation of human behavior to the model and design of the optimal monitoring network, among other aspects of relevance for the topic. As the authors suggested, the traffic forecasting field was "in its infancy" (*ad pedem literae*); from then on, a sprawl of research started and allowed [7] to review the subject in a more deep and principled manner. In this review the authors considered three main aspects of traffic forecasting: scope of application, output specification, and modeling features. The first aspect refers to the kind of roads for which the prediction is made and the type of application the prediction will be used for (mainly ATMS and ATIS). On the other hand, the output specification involves the prediction horizon and step concepts, and elaborates on which traffic parameters should be considered for developing the predictive model. The authors also provide a profound analysis of the modeling aspects of traffic forecasting. A classification of prediction models is presented, comparing their characteristics and performance. The authors contribute a methodological workflow to select and tune the model parameters that has been extensively referred [10], [12]–[17].

This paper systematically examines the recent developments gravitating on data-driven traffic forecasting methods, with an emphasis on the technical advances made in this dynamic field, the degree of achievement of the different technical challenges posed in the past, and a critical diagnosis of the unexplored areas of research that should be targeted by the community in forthcoming years.

Although most of the reviewed literature is focused in road features forecasting, specially traffic flow or volume, travel time is an alternative variable to predict. It is more human-understandable than flow, occupancy or speed, where a given value may have different interpretations depending on the type of road under analysis. A survey on travel time prediction was published in [18], exposing the techniques and main drawbacks and difficulties to estimate this traffic feature. According to this study, in 2005 the main handicap for this specific traffic forecasting scenario was concluded to be the lack of data, which the authors proposed to overcome with simulation techniques. Nowadays, the widespread proliferation of GPS-enabled devices and connected vehicles that can supply floating car data allows researchers to model and predict travel-times without resorting to simulation methods [19]–[22]. A more recent review on this field [23], delves in the latest travel-time prediction from floating car data (FCD) sources, and highlights the relevance of forecasting this feature for ATIS and operational planning of mass transit.

In 2007, a concise work on prediction modeling [10] delved into understanding and examining forecasting models proposed by previous research works. The authors considered a naïve category (predictions based in historic values or an average of them) in addition to the parametric/non-parametric category described by [7], and classified traffic simulation as a parametric method that can be used for traffic forecasting. The main comparison features were prediction horizon, scope of application, computational speed and accuracy, most of them aligned with previous surveys. Interestingly, network-wide predictions and a comparison method between simulation and the rest of models were first suggested as open challenges in this work. More recently, Bolshinsky et al. in [13] have agreed on the latter, stating the difficulties to compare not only simulation to every other forecasting method, but all methods to each other.

In this survey the main focus is set on techniques, adding a few not contemplated in previous reviews. Besides the forecasting method and the prediction horizon, the authors highlight the relevance of sources of data, as studies with different input data can be hardly compared. This work also ponders the pertinence of input data other than traffic data. An assortment of non-traffic factors affect traffic conditions, such as weather conditions, holiday periods, events, incidents and road works, or seasonal factors. Including these elements in a forecasting model can aid to refine predictions, but as reflected in [13], only a few previous authors have them into account. In the same year, [11] updated the taxonomy of forecasting methods provided by the same authors 5 years earlier [10] with more recent works. They conclude, as suggested by other authors before, that there is no universal method that fits every situation better than the rest. As a consequence, they propose as a challenge the development of model ensembles that outperform the existing ones, but more interestingly, the creation of a method to choose a technique, given the attributes of the forecast to make. A different view of the subject is proposed in [24], which offers a mathematical optimization perspective of the problem. This review summarizes the efforts made in

Growing open-data initiatives facilitate access to a variety of inputs for future models. This also influences the data sources criterion, which is used as comparison element only in most recent reviews, implying a more data-centered corpus of literature. Prediction models are often built upon data from one source (mainly traffic loops, and road sensors), but it might be difficult to compare a model with ATR input data to another with traffic camera data.

this direction and states that all observability, estimation and prediction problems can be integrated and formulated as an optimization paradigm.

To the date of this manuscript the most updated review on traffic forecasting is the one by Vlahogianni et al. in [25]. They studied literature on traffic forecasting since 2004, comparing scope of application, prediction horizon, input sources and methodological approach. According to all previous research contributions, they concluded that little effort had been dedicated to network-wide predictions, urban arterial forecasting, or multivariate models; besides, in most previous works predictions were made for traffic volume. Leaning on the ongoing development and expansion of data driven techniques and technologies, they gathered a series of challenges for future work in this area. Part of them coincide with previous surveys and the aforementioned conclusions; the key aspects to develop identified in these works are arterial and network-level predictions, shifting from traffic volume to travel-time prediction, spatio-temporal forecasts, model selection techniques and model comparison methodologies [7], [10], [11], [13], [18]. In addition to those recurrently formulated questions, the authors proposed new challenges related to data fusion, data aggregation, the explanatory capacity of variables and the responsiveness of the predictive models. A travel-time forecasting survey recently contributed in [26] portrays the main methods and issues in this field. Conclusions highlight similar relevant aspects: network-wide prediction, exogenous factors, data fusion, and the relevance of congestion situations for this kind of forecasts. Oh et al. in [27] review the data-driven approach of travel-time prediction by focusing on methods and techniques that complement the previous references.

A. A Taxonomy of Research

Literature reviews have so far analyzed the state of the art around traffic forecasting under different criteria. Table 1 summarizes the key criteria considered by the aforementioned reviews. This table reveals that the most used criteria in literature reviews are those related to prediction methods, horizon, scale and output variables. Forecasting techniques are a very relevant part of the reviewing methodology, but comparing their performance is a more complex task, as usually each type of technique delivers different performance metrics. Criteria such as the *optimization type*, the *data resolution* or the *streaming of data for online learning* are addressed only in most recent surveys, which cover a wider spectrum of contributions concentrated on data-driven approaches. Prediction horizon and context are likewise significant criteria for most authors, and they are indeed widely used with comparing purposes. Nonetheless, prediction context allows for little comparative; as highlighted by [25], most models are built on freeway or highway contexts, while urban arterial traffic forecasting is less addressed and much more challenging [28]. Prediction horizon is kept under one hour in most works and is varying enough not to be considered as a fair comparison factor.

Accounting for which non-traffic inputs are taken into account to make predictions is an interesting criterion that was first proposed by [1], but it has been rarely used ever since. A recent review by [13] shows that very few works have addressed these factors, which might be the reason why other reviewers do not consider them as a comparative criterion. In addition, calendar-related aspects are implicit in time-series models, which constitute a great part of forecasting literature. However, these and the other factors listed in Table 1 are proven highly relevant in forecasting [28], [31], [32], with a considerable impact in real traffic conditions. Growing open-data initiatives facilitate access to a variety of inputs for future models. This also influences the *data sources* criterion, which is used as comparison element only in most recent reviews, implying a more data-centered corpus of literature. Prediction models are often built upon data from one source (mainly traffic loops, and road sensors), but it might be difficult to compare a model with ATR input data to another with traffic camera data. In this line, data fusion techniques allow to combine data from different sources in the same model, including traffic data from different sensors, or non-traffic inputs [33], and it is considered by [25], [34] as one of the main challenges in this field.

Computational effort aspect was only considered in [1], back in 1998; this aspect has become less relevant since then. Generalizability was deemed as a key feature of simulation models. Their need for parametrization renders them too specific; they improve if they are applicable to other contexts. This is seldom analyzed for non-simulation models, but if standard test-beds and test data were to be available for testing and comparing algorithms, as suggested in [25], their results might be more easily generalizable. As data availability increases, more attention is paid on the type of learning process, which can rely on online streams of data [23].

Thus, plenty of aspects can be studied when evaluating and testing traffic forecasting methods, with different levels of relevance considering their abundance of use. The following section will describe the main issues found by researchers in traffic forecasting, in order to determine if the relevance of aspects is related to their impact in solving the issues, or some other aspects could be considered.

B. Some Well-Established Considerations

Reviewing traffic-forecasting literature has lead preceding researchers to diverse conclusions, future directions and challenges, which are next reviewed and analyzed. A fair share of them is maintained through the years, despite the advances in some of the fields involved, e.g. computer processing capacity, machine learning algorithms, simulation tools and access to data.

1. Stochastic Nature of Traffic

This feature is frequently stated in all kinds of traffic forecasting literature, and its effects are described to be considered by transport modelers since the first works. Traffic predictions have two natural limits: one is related to the randomness of events that can affect traffic; the other one is related to the effect predictions themselves can have on drivers decisions and habits [1].

Table 1. Prediction aspects assessed in previous reviews.

Criteria	Types	Referenced in
Prediction Method	Naïve (instant, historic average)	[10], [11], [13], [26], [29]
	Parametric (simulation, time series, Markov chains, Kalman filters)	[1], [7], [10], [11], [13], [25], [26], [29]
	Non-parametric (Neural Networks, bayesian, fuzzy logic, ATHENA)	[1], [7], [10], [11], [13], [25], [26], [29]
Prediction Horizon	Prediction horizon	[1], [7], [10], [11], [13], [25], [29], [30]
	Prediction Step	[7]
Prediction Scale	Single Location	[1], [10], [11], [13], [25], [26], [29]
	Road Segment	[1], [10], [25], [26], [30]
	Whole or part of the network	[1], [7], [10], [11], [25], [26], [30]
Prediction Context	Urban (arterial)	[1], [30], [7], [10], [25]
	Rural	[30], [10]
	Freeway	[1], [30], [7], [10], [25]
Data sources	Traffic management bureaus	[1], [13], [29]
	Automatic Traffic Recorders	[1], [13], [26], [29]
	Sensors (other than loops)	[13], [26], [29]
	Cameras	[1], [13], [26], [29]
	GPS-FCD	[13], [26], [29]
	Cellphone data	[13]
	Public transport information	[13]
	Crowd sourcing	[29]
	Social media	[29]
Exogenous factors	Calendar (Week days, weekends, bank holidays)	[1], [11], [13]
	Periods of the day	[1], [11], [13]
	Holidays	[1], [11], [13]
	Seasonal differences	[1], [11], [13]
	Weather	[1], [13], [25]
	Special events (demonstrations, parades)	[1], [13]
	Periodical events (sports, other social events)	[1], [13]
	Road Works	[1], [13]
	Traffic incidents	[1], [13], [25]
	Traffic source areas (malls, parkings, adjacent roads)	[1], [13]
Predicted variable	Traffic Flow (vehicles/hour)	[1], [7], [10], [11], [13], [25], [30]
	Traffic Density (vehicles/km)	[1], [7], [10], [11], [25], [30]
	Average speed	[1], [7], [10], [11], [25], [30]
	Travel time	[1], [7], [10], [11], [25], [26]
Uni/Multivariate	Not applicable	[1], [7], [13], [25]
Prediction performance	Not applicable	[1], [10], [30]
Optimization type	Not applicable	[25]
Computational effort	Not applicable	[10]
Generalizability	Not applicable	[30]
Scope of application	ATIS, AMTS, Logistic	[1], [25], [30]
Data resolution	Not applicable	[25]
Stream mining	Online, offline	[23]

Current works frequently combine methods with different purposes, and in recent research is common to find several types of optimizations for tuning the models.

2. Network Application

A conclusion common to all reviews is the lack of network-wide prediction models, compared to single-point or road segment predictions. The latter are useful for ATMSs, but the first help building more efficient ATISs [11], and thus can reach to the general public. Since the earlier reviews, when this was considered a network sensor coverage design problem [1], the subject has arisen in every review. This issue is related to the urban traffic prediction problem, for network predictions are more useful in urban environments. Network-wide predictions have been explored mainly via simulation [10], [11], although other approaches are found in literature [35], [36]. There is an increasing number of this kind of predictions in recent works (as shown in Tables 2 and 3), but they are usually referred to compact areas. Network-wide prediction models still remains a challenge in this field.

3. Urban Traffic

Predicting urban traffic is defined in previous reviews as a very complex task. Signals, interactions with close links, sources of traffic, and a more complex origin-destination relation have made some researchers state that traffic flow in urban arterials -even single location traffic- cannot be predicted as accurately as in freeways [28]. Again, simulation models are the best suited to address this issue, parametrizing preceding inputs; notwithstanding, some researchers have concentrated on this matter by incorporating spatial information to time series [28], neural networks [37]–[39] or other non-parametric approaches [14], [40]. Arterial traffic prediction has grown in interest, yet it constitutes a slight portion of works on traffic forecasting [25], and embodies along with network prediction one of the main challenges of forecasting.

4. Applicability and Model Selection

Or the ability of a model to adapt to different contexts. This is a typical simulation problem [30]; the more parameters are set, the more difficult it is to apply the simulation model to other circumstances. Non-simulation methods, and specifically non-parametric methods adapt better to different contexts [13], but all previous reviews coincide in asserting there is no best method that suits all situations [7], [10], [11], [13], [25], which implies an applicability at a higher level, not of the model, but of the method to choose the most suitable model given the characteristics of the forecasting problem [11]. A recent work [41] proposes a meta-modeling technique to tackle the model selection and parameter tuning. A lot of research is yet to be made to take advantage of these techniques, and even extend them to greater decision making tools.

5. Metrics of Performance and Comparison

Abundance of methods for traffic forecasting makes establishing a unified comparing metric an intricate task [1], [7]. Although Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE) are usual metrics for model performance measuring [10], [11], [42], they do not provide comparable measures when the complexities of compared models are too divergent [8]—e.g. a neural network and an Autorregresive Integrated Moving Average (ARIMA) model—or when the input datasets are completely different. Also, for network-wide models, errors can propagate through time and space along the network, so a spatio-temporal correlation between successive predictions can help measuring the performance. Although contemporary studies keep using the Root Mean Square Error (RMSE) and the Mean Absolute Percentage Error (MAPE) to assess performance, the definition of benchmark datasets, environments and metrics is currently identified as a necessity in the area [25].

6. Hybridization of Methods

Combining prediction techniques is a tendency that was explored in the first place with combination of ARIMA models with other methods to improve accuracy, and more recently combining the prediction method with techniques to preprocess data like clustering [31], [32] and to optimize the model [25]. Current works frequently combine methods with different purposes, and in recent research is common to find several types of optimizations for tuning the models (see Tables 2 and 3). Tselentis et al. [43] showed that combining models with different degrees of spatio-temporal complexity and exogeneities is most likely to be the best choice in terms of accuracy. Moreover, the risk of combining forecasts is lower than the risk of choosing a single model with increased spatio-temporal complexity.

III. New and Revisited Challenges

Existing background of traffic forecasting literature and reviews present countless efforts devoted to tackle foregoing challenges. This section is aimed to broaden the scope of previous questions and present the most recent developments in them. The latest survey in [25] presented a vast literature review up to 2014, and compiled challenges in

Table 2. Literature of the 2014-2016 period (1/2).

Reference	Scope Context	Predicts	ST	Max. Horizon	Data Extension	Step	Data Source	Exogenous Factors	Prediction Model	Comparative Models	Ageing
[67]	U-Sgm	Speed	×	2 min.	1 month	2 min.	RTMS	–	LSTM-NN	SVM, ARIMA, NN, KF	Memory Block
[59]	F-Sgm	Speed, TT	×	1 hour	1 month	15 min.	Camera, FCD, Loop	–	NN	GAM, ARIMA	×
[32]	U-Pts	Flow	×	NH	1 year	15 min.	Loop	Calendar	Cluster+RF	HA	×
[70]	U-Sgm	Volume, TT	×	2 days	15 days	–	Loop	–	NN	LR, ARIMA	×
[44]	U-Ntw	Speed	✓	1 hour	1 month	5 min.	FCD	Calendar	KNN	HA, LSSVM, NN	×
[9]	F-Sgm	Flow	×	1 week	1 week	10 min.	RTMS	Calendar	FNR	SARIMA, NN, LSSVM	×
[71]	U-Pts	Flow	×	3 days	15 days	1 hour	Loops	–	Block Regression	LR, SARIMA	×
[48]	U-Ntw	Flow	✓	30 min.	–	5 min.	Loops, Simulation	–	LLDR	○	×
[49]	U-Ntw	Speed	✓	15 min.	5 years	1 min.	Loops	Incidents	Context Aware Ensemble	○	Reward-regret
[72]	U-Pts	Flow	×	5 min.	1 month	5 min.	Loops	–	MVLR	○	×
[73]	U-Sgm	Congestion	✓	–	–	–	Simulation	–	MCS	○	×
[74]	U-Pts	Flow	×	1 hour	3 months	1 hour	Loops	–	NN Ensemble	Basic, HA	×
[75]	U-Pts	Speed	✓	30 min.	–	30 min.	Simulation	–	NN	○	×
[76]	F-Pts	Flow	×	5 min.	5 days	5 min.	Manual	Calendar, Type of Vehicle, Speed	NN	○	×
[77]	U-Pt	Flow	×	1 day	3 days	10 min.	Camera	–	SARIMA	Basic, HA	×
[60]	U-Pts	Flow	×	NH	1 year	15 min.	CCTV, Laser, Loops	GIS, Weather	Cluster + NN	Basic, HA, SVR, MVLR, KNN, NLR	×
[78]	U-Ntw	Flow	✓	1 hour	3 months	4.5 min.	Camera	Weather, pollutants, noise, FCD	CRF	○	×
[79]	F-Pt	Flow	×	2 min.	1 day	2 min.	RTMS	–	AKNN-AVL	KNN, ARIMA	AVL Case Database
[80]	F-Pt	Flow	×	4 hours	7 days	5 min.	Loops	–	Volterra	RFBNN	×
[81]	F-Sgm	Flow	×	1 day	1 day	1 min.	RTMS	–	MLPNN	RBFNN, WavenetNN	RL-Difference
[82]	F-Pt	Flow	×	10 min.	1 month	10 min.	Radar	Weather	CCGA-GSVR	SVR, WASVR, PSO-BP, ARIMA	×
[45]	F-Sgm	Flow	✓	30 min.	2 months	–	Loops	–	KNN	NN	×
[83]	F-Ntw	Flow	✓	2 min.	1 day	2 min.	RTMS	–	Custom TSA	ARIMA, LR	×
[51]	U-Ntw	Flow	✓	1 hour	1 month	5 min.	Loops	–	STRE	ARMA, STARMA, NN	×

(*continued*)

Table 2. Literature of the 2014-2016 period (1/2) (*continued*).

Reference	Scope Context	Predicts	ST	Max. Horizon	Data Extension	Step	Data Source	Exogenous Factors	Prediction Model	Comparative Models	Ageing
[84]	U-Ntw	Flow	✓	15 min.	–	5 min.	Simulation	Speed	LCG-BN	ARIMA, NN, BN	×
[85]	U-Pts	Flow	×	15 min.	25 days	15 min.	Loops	–	GPDM	KNN, RW	MTLNN
[86]	U-Pts	Flow	×	1 hour	7 days	–	Loops+manual	–	Fuzzy Logic	NN, LR	×
[87]	F-Pt	Speed	×	10 min.	2 months	2 min.	Loops	Calendar	HPSO-NN	ARMA	Error compensation
[88]	U-Ntw	TT	×	3 days	15 months	–	Public transport info	Calendar	EnsemblePPR, SVM, RF	–	×
[55]	U-Ntw	TT	×	NH	–	–	FCD	Calendar, time	Markov chain		×

Context: U → Urban; F → Freeway or highway; Ntw → network; Pt → single point; Pts → diverse points; Sgm → points in a road segment.
Predicts: TT → Travel Time
ST → Spatial-temporal prediction
Horizon, Extension, Step: - → No data available; NH → no horizon
Source: FCD → Floating Car Data; RTMS → Remote traffic microwave sensor.; Lat-lon → Latitude and longitude; AVL → Automatic Vehicle Location
Model, comparative model: LSTM-NN → Long Short-Term Memory Neural Network; KF → Kalman Filtering; GAM → Generalized Additive Model;
RF → Random Forest; HA → Historic Average; LR → Linear Regression; LSSVM → Least Squares Support Vector Machine; FNR → Functional Nonparametric Regression; LLDR → Link-to-link dividing ratio; ○ → Model tested against itself; MVLR → Multi-variable linear regression; MCS → Mobile crowd sensing; CRF → Conditional random field; AKNN-AVL → K-nearest neighbor combined with balanced binary tree;RL → Reinfor cement learning; CCGA → Cloud Chaos Genetic Algorithm; GSVR → Gaussian support vector regression; PSO → Particle Swarm Optimization; STRE → Spatio-temporal random effects; LCG-BN → Linear conditional Gaussian Bayesian network; GPDM → Gaussian Process Dynamical Models; PPR → Projection Pursuit Regression; EFGM → Grey Verhulst model with Fourier error corrections; MKNN → Multivariate k-nearest neighbors; ieRSPOP → incremental rough set-based pseudo outer-product with ensemble learning; DTT → Dynamic Topology-Aware Temporal traffic model; GA → Genetic Algorithm Optimization; PHFRBS → Parallel Hierarchical Fuzzy Rule-Based System; FO → Firefly optimization; TSA → Time series analysis; FfO → Firefly optimization; STW-KNN → Spatio-temporal weighted k-nearest neighbor; FIMT-DD → Fast incremental model trees-drift detection; AR → Auto Regression; MARS → Multivariate Adaptive Regression Splines; HS → Harmony Search.

ten global categories. We delve into some of those challenge categories, and propose future lines of development in the field of traffic forecasting, having in mind the shift to data driven machine-learning techniques that is also mentioned in [25].

A. The Context of Forecasting

By context we refer to the forecasting setting and the traffic parameters used. As can be observed in Tables 2 and 3, scope-context column displays the context of application (urban or freeway) and the scope (point, scattered points, segment, network). A significant increase in network predictions is revealed since previous state of the art, which considered it a challenge. Around 1 in 20 of works examined in [25] had network-wide coverage. Since their review, the proportion has risen considerably. Network-wide predictions usually require the model to know the influence of surrounding road links. This spatio-temporal correlation is mirrored in corresponding column (ST), and it shows that which was considered a challenge in 2014 has started to develop in current works. Data driven approaches facilitate the inclusion of spatial and temporal correlations in the model, and k-Nearest Neighbor (kNN) models are widely used [44]–[47]. Origin destination matrices are estimated to make predictions in [48], and context-aware models that modify forecasts depending on surrounding link predictions are employed in [49]–[52].

Travel-time replacing flow as predicted value was also considered a relevant challenge, as the first is a more helpful metric of traffic. While predicting volume can rely only in traffic counts, estimating travel-time is more complex [22]; although many attempts have been made to estimate travel-time from traffic counts [53], [54], availability of other data such as Floating Car Data (FCD) [23], [55]–[57] or vehicle identification [50], allows to build trajectories and perform improved estimations [58]. The relative amount of works centered in travel-time prediction is maintained similar to that found in previous reviews. Data sources column shows precisely that even when input data fusion is considered a challenge, and data can be obtained in many ways, most of works are still concentrated on traffic loops. It is possible to observe, though, some studies that fuse inductive loop readings with camera information, FCD or automatic number plate recognition (ANPR) [59]–[61] and an upsurge of the use of FCD, more reachable, and more exploitable with data driven methods.

B. Long-Term Prediction Horizons

Increasing the prediction horizon is not usually regarded as a challenge. Many authors describe the

Table 3. Literature of the 2014-2016 period (2/2).

Reference	Scope context	Predicts	ST	Max. Horizon	Data extension	Step	Data Source	Exogenous Factors	Prediction Model	Comparative models	Ageing
[89]	F-Sgm	TT	×	NH	15 days	5 min.	Loops	Simulation, occupation, volume, speed	Evolving Fuzzy NN	MLR, Basic, LM, NN, CP	×
[56]	F-Pts	TT, Speed	×	1 min.	24 days	1 min.	Loops	FCD	Grey Model EFGM	TSA	×
[90]	F-Pts	Flow	×	4 hours	3 months	5 min.	Loops	–	ML-KNN	ARIMA, KNN, Basic	×
[52]	F-Sgm	TT	✓	25 min.	93 days	1 min.	Loops	–	KF	–	×
[91]	F-Ntw	TT	✓	1 hour	383 days	10 min.	Loops	–	NN, LR, RTree, RF	○	×
[92]	F-Sgm	Flow	×	1 day	15 months	0.5 min.	Loops	Workzones data	LLR	–	×
[93]	F-Pt	Flow	×	1 day	7 days	5 min.	Loops	Calendar	Wavelet NN	○	×
[94]	F-Sgm	Flow	×	5 min.	1 month	5 min.	Radar	–	GRNN	Vector Regression, HA	×
[95]	F-Pt	Flow	×	5 min.	7 days	6 min.	Loops	–	Cluster + NN	–	×
[96]	F-Pt	Flow	✓	5 min.	–	5 min.	RTMS	–	Fuzzy NN	–	Real Time correction
[46]	F-Sgm	Flow	✓	1 hour	6 weeks	15 min.	Loops	Calendar	MKNN	HA, SARIMA, NBR, KNN	×
[69]	F-Pts	Flow	×	NH	82 days	5 min.	Loops	Calendar	Similarity, RBLTF	ARIMA, GRNN	×
[97]	F-Pt	Flow	×	10 min.	6 days	15 min.	Loops	–	ElmanNN+Bayes	BPNN, Wavelet NN	×
[98]	F-Pt	Flow	×	1 hour	9 days	5 min.	Loops	–	ieRSPOP	–	ieRSPOP
[99]	F-Ntw	Flow	×	30 min.	2 months	5 min.	Loops	–	DTT1, DTTinc	ARIMA, SVR	DTT
[100]	F-Pts	Flow	×	2 days	1 month	1 hour	Loops	Speed, occupancy	TS-TVEC (TSA)	MLPNN	Error correction
[101]	F-Pts	Incidents	×	15 min.	3 weeks	5 min.	Loops	Flow, occupancy	Drift3Flow	ARIMA, ETS	Concept drift
[102]	F-Sgm	Flow	×	30 min.	1 month	5 min.	Loops	Density	PHFRBS + GA	○	×
[103]	U-Ntw	Speed	×	15 min.	6 days	5 min.	FCD	–	BN, NN, SARIMA+BN	HA	×
[104]	F-Pts	Speed	×	10 min.	2 days	–	Loops	Calendar	LSSVM+FO	LSSVM	×
[47]	F-Sgm	Flow	✓	5 min.	15 days	5 min.	FCD	–	STW-KNN	KNN, NN, NB, RF, C4.5	×
[105]	U-Pt	Flow	×	10 min.	4 days	10 min.	Loops	–	LSSVM+FfO	LSSVM, RBFNN, LSSVM+PSO	×
[106]	U-Ntw	Speed	×	5 min.	15 months	5 min.	Loops	–	Custom TSA	○	×

(*continued*)

Table 3. Literature of the 2014-2016 period (2/2) (*continued*).

Reference	Scope context	Predicts	ST	Max. Horizon	Data extension	Step	Data Source	Exogenous Factors	Prediction Model	Comparative models	Ageing
[50]	U-Ntw	TT	✓	30 min.	166 days	5 min.	ANPR	–	STNN	HA, ARIMA, STARIMA	×
[107]	F-Sgm	TT	✓	15 min.	1 year	5 min.	FCD	Calendar, Lat-lon	GradientBoost	ARIMA, RF	×
[108]	U-Pt	Flow	×	1 day	4 days	15 min.	Loops	–	WaveletNN+GA	WaveletNN	×
[61]	F-Ntw	Flow	×	15 min.	5 years	15 min.	Loops, ANPR, Camera, FCD	Speed, TT, Calendar, Lat-lon	FIMT-DD	○	Drift Detect
[109]	U-Pts	Flow	×	2 hours	1 year	15 min.	Loops	–	MLP + HS	○	×
[110]	U-Pts	Flow	×	5 min.	4 years	5min.	Loops	–	Extreme Learning	○	Forgetting mechanism
[111]	F-Pt	TT	×	1 min.	15 hours	–	Loops	–	Cluster+ARIMA	○	×
[112]	U-Ntw	TT	✓	30 Seconds	1 year	Online	AVL	Calendar	RF	○	✓
[57]	U-Ntw	TT	✓	30 min.	15 days/2 months	5 min.	FCD	–	Graph Based Lag STARIMA	Basic, HA, kNN, RF, SVR	×
[113]	U-Ntw	Flow	✓	50 min.	1 month	10 min.	Loops	–	SVR	ARIMA, MARS, Spatio-Temporal Bayesian MARS, AR	×

Context: U → Urban; F → Freeway or highway; Ntw → network; Pt → single point; Pts → diverse points; Sgm → points in a road segment.
Predicts: TT → Travel Time
ST → Spatial-temporal prediction
Horizon, Extension, Step: - → No data available; NH → no horizon
Source: FCD → Floating Car Data; RTMS → Remote traffic microwave sensor.; Lat-lon → Latitude and longitude; AVL → Automatic Vehicle Location
Model, comparative model: LSTM-NN → Long Short-Term Memory Neural Network; KF → Kalman Filtering; GAM → Generalized Additive Model;
RF → Random Forest; HA → Historic Average; LR → Linear Regression; LSSVM → Least Squares Support Vector Machine; FNR → Functional Nonparametric Regression; LLDR → Link-to-link dividing ratio; ○ → Model tested against itself; MVLR → Multi-variable linear regression; MCS → Mobile crowd sensing; CRF → Conditional random field; AKNN-AVL → K-nearest neighbor combined with balanced binary tree;RL → Reinfor cement learning; CCGA → Cloud Chaos Genetic Algorithm; GSVR → Gaussian support vector regression; PSO → Particle Swarm Optimization; STRE → Spatio-temporal random effects; LCG-BN → Linear conditional Gaussian Bayesian network; GPDM → Gaussian Process Dynamical Models; PPR → Projection Pursuit Regression; EFGM → Grey Verhulst model with Fourier error corrections; MKNN → Multivariate k-nearest neighbors; ieRSPOP → incremental rough set-based pseudo outer-product with ensemble learning; DTT → Dynamic Topology-Aware Temporal traffic model; GA → Genetic Algorithm Optimization; PHFRBS → Parallel Hierarchical Fuzzy Rule-Based System; FO → Firefly optimization; TSA → Time series analysis; FfO → Firefly optimization; STW-KNN → Spatio-temporal weighted k-nearest neighbor; FIMT-DD → Fast incremental model trees-drift detection; AR → Auto Regression; MARS → Multivariate Adaptive Regression Splines; HS → Harmony Search.

degradation of predictions when this horizon is extended [1], [7], [10], but long-term predictions can be useful from the ATMS perspective [62], [63], or for scheduling in logistics planning [18]. The relevance of long-term predictions is claimed also for macroscopic network planning [64] for infrastructure development [65]. Besides, although long-term predictions cannot provide accurate outputs, they have been proposed as another input to short-term prediction models [9], [66], [67]. Notwithstanding, this characteristic is kept in almost every work before 2014 below 60 minutes [25] in the future. Greater accurate forecasting horizons might assist the progress of traffic management systems, and their achievement is connected to better prediction models, more and more linked data sources and spatial context aware predictions.

A shift to data-driven approaches is observable in most recent literature about traffic variables forecasting [25]. These models use large databases with plentiful records from which they learn and produce predictions.

In terms of forecasting horizon, Tables 2 and 3 show that most of recent works also predict traffic features under a 60 minute extent. Nonetheless, an increment in larger horizons or non-horizon models is noticeable, paired with a decline in ARIMA (and variants) models. Data driven models based on large data bases of readings are broadening in recent years, allowing researchers to make heuristic predictions at any point in the future [31], [32], [49], [68], [69], and data extension reaches to 5 years of data [61]. Recent literature shows that forecasts further than 60 minutes are possible, in the data-driven context, and can perform as good as short-term.

C. Exogenous Factors in Multi-Input Models

The main obstacles for accurate long-term and even short-term prediction are factors that affect traffic but are not part of its seasonal behavior and confer traffic its stochastic nature [1], [10], [11]: road works, incidents, events, weather, proximity to traffic affecting facilities (parking lots, shopping areas, work/study centers), and calendar matters (bank holidays, weekends). Although incidents or weather changes can happen suddenly and they can be difficult to predict, other traffic affecting factors like road works or events are usually foreseeable. Feeding these kind of inputs to data-driven prediction models can enhance their performance, enriching the provided forecasts [13]. Anticipated by [10] in 2007, mobility data sources and availability have been increasing since then. This entails a chance and a challenge regarding data fusion and integration [1], [25], [34]. Integrating exogenous factors from different data sources is a direction that should be considered in future works.

Exogenous factors continue to be barely addressed in recent research, and calendar information is the most considered one [9], [32], [44], [46], [55], [61], [69], [76], [87], [88], [93], [104], [107]. Weather is only used in three works [60], [78], [82], with less predictive relevance in the model than expected. Weather, pollutants [114] and noise [78] and incidents [49], are variables that need to be predicted too. A model can learn how traffic behaves when an incident happens, but it will need a forecast of future incidents to elaborate the output. Incident forecasting is addressed in [115] via Bayesian Neural Networks and detection through other traffic features. This work arises the many facets involved in incident prediction. A long as it is possible to build models of these inputs, it is also possible to encompass them in traffic forecasting models. Information about sport events, parades or in general any human-organized happenings is scarcely used [92], [116]. These events presumably have similar effects in traffic, if they are repeated over time, and combined with calendar information, they might provide a valuable input to prediction models. Road works are predictable, but they are not usually repeated over time in the same place, and each location is affected in a different way. However, it is possible to model the impact of roadworks depending on their location, affected lanes, and other factors, and include this in traffic conditions prediction model [117]–[119].

D. Data Ageing and Concept Drift in Data-Driven Models

A shift to data-driven approaches is observable in most recent literature about traffic variables forecasting [25]. These models use large databases with plentiful records from which they learn and produce predictions. As data extension increases, it is more probable that knowledge extracted from those data is less factual; road networks are constantly changing, specially in large urban areas. The usage of those roads is also variable within long periods of time, increasing in thriving metropolis and declining in economic crises affected areas. If a prediction model learns from a 5 year database, like [61], a change in flow direction, a closed or new lane, or a change in usage patterns during those years are possible, and depending on the urban area modeled, they can be highly probable.

Responsive models that adapt to unexpected short-term factors, like accidents, congestion situations or weather conditions, were proposed in [25]. An adaptation to long-

> The most relevant development in the field of traffic prediction in recent years is related to a shift in the prediction modelling paradigm. Advances in data oriented techniques and technologies, explosion of Big Data and machine learning, along with growing availability of traffic related data from plentiful sources have contributed to leave behind time-series analysis methods and given a boost to data driven models.

term factors is feasible through data ageing. Weighing down old measurements can be a naïve approach to provide a forgetting mechanism that gives a greater relevance to current input values of the model [110]. Concept drift techniques [120] allow an adaptive learning strategy that has recently started being applied to traffic prediction problem. The application of these techniques spans from adapting a model to detecting anomalies and recalling the inferred traffic patterns prior to the occurrence of the anomaly, which can be also exploited under incident or atypical road congestion [121].

Prediction models have experienced a considerable changeover. ARIMA models are in clear decline, and only three of the analyzed studies elaborate predictions with this kind of methods. They still are, though, usual models to compare to.

The shift to Machine learning (ML) techniques and Artificial Intelligence (AI) has roots to the nature of traffic datasets. Traffic data are usually messy, extremely irregular. The non-stationary and nonlinear nature of traffic has been frequently observed in forecasting literature [122]–[125]. ML and AI approaches may overcome the parametric nature of statistical models and the relevant modeling constraints. They are capable of mining information from messy and multi-dimensional traffic datasets. To this end, neural networks still maintain a hefty presence, with several variations, such [74], who uses an ensemble of neural networks that cooperate and aggregate predictions obtaining better results. [60] and [95] cluster data prior to perform a neural network regression. [89] and [96] use fuzzy rules with clustering purposes. In [110] an On-line Sequential Extreme Learning Machine (OSELM), a type of neural network, is used with a forgetting mechanism that allows the author perform adaptive learning. Particle swarm optimization is used in [87] to optimize the model parameters and in [108] a genetic algorithm is used to optimize a wavelet neural network. Aside from neural networks, fruit fly optimization [105] and firefly optimization [104] solvers are also used to optimize hyper parameters of support vector machine (SVM) algorithms. KNN models are also widely used [45]–[47], [79], [90], specially related to spatio-temporal forecasts.

Few works include any sort of concept drift techniques by now, but the proliferation of Big Data technologies and long-term prediction methods should lead to a more generalized usage of data ageing mechanisms. Adaptive learning techniques are applied in some of the works. In [98], an incremental learning algorithm (coined as *ieRSPOP*) is tested on three datasets, one of them with 9 days of traffic flow. Incremental learning implies any instance of data can only be used once for training [126] so older instances become less relevant when the algorithm evolves. A more traffic-specific study with incremental learning is made by [99], obtaining smaller MAPE values than the same method without incremental learning. Incident detection is achieved by [101] using *Drift3Flow*, an online incremental learning method that detects changes in traffic flow and occupancy and infers when they are caused by incidents. Same authors develop a prediction model to mitigate the effect of bus bunching by using information gathered by the vehicles, and requires the model to be constantly updated with online data [112]. Drift detection is also used in [61] to improve the prediction efficiency with 5 years of data.

E. Big Data and Architecture Implementation

The era of connected vehicles and sensed roads is near, and profuse data are becoming more available to exploit. Cloud and parallel computing big data paradigms can provide the means for macrosimulation, computationally inexpensive entire network predictions, traffic deep learning and more explanatory power of models [127]. Implementation of traffic forecasting tools in Big Data architectures also allows for real time predictions based on several types of input, taken from different open and not open sources in an effective way. Nevertheless, a significant work is required in this field; learning methods ought to be parallelized to be capable of mining chunks of intercorrelated data without jeopardizing the quality of the predicted variable. Likewise, the interaction of traffic information with non-structured datasets is also challenging due to the essential differences that may eventual characterize them in terms of ageing, drift, graph-like nature and spatial connectedness. The importance of developing Big Data approaches well suited for dealing with traffic forecasting will imminently become paramount and gain momentum in the research community.

To this end, the efficient representation for geospatial big data will play a decisive role. Most traffic information—mainly due to the crowdsourcing initiatives—have a clear location dimension (e.g. GPS data), which may hinder significant information on the recurrent and non-recurrent traffic patterns. The above along with the need to visualize and quantitatively process geospatial big data for decision support (e.g. real-time traffic management) with methods that conceptually defy the classical statistical modeling constrains of optimal design or model based sampling opens new opportunities to modeling and forecasting.

Nevertheless, traffic in both kind of networks needs to be managed: optimal routes, congestion situations, demand administration, priority control and alterations in the level of service, among others. In computer networks packages are identified and pinpointed at all times.

It should be noted that leveraging big data to improve short-term traffic predictions is strongly related to the availability of relevant data that are altogether linked. Traffic and mobility data openness is of extreme importance not only to the accuracy and adaptability of traffic forecasts, but also to the ability to provide users and planners with timely traffic information. Nevertheless, open data initiatives come with certain challenges, such as the economic cost of openness, issues of data reliability, access and control, security and legislative framework and so on. In any case, the concepts of opened and linked data are crucial and will have significant socio-economic perspectives. In the future, those that possess big data and can analyze them will most likely have a significant competitive advantage.

F. Computer Traffic Versus Road Traffic

Computer networks share some relevant features with road networks. Researchers have established analogies specially at microscopic levels, being data packages the counterpart to individual vehicles [128], [129]. At macroscopic—defining the origin and destination of a route— and mesoscopic—controlling the route and making decisions—levels differences are larger, as computer traffic is entirely controlled by the infrastructure [129]. Nevertheless, traffic in both kind of networks needs to be managed: optimal routes, congestion situations, demand administration, priority control and alterations in the level of service, among others. In computer networks packages are identified and pinpointed at all times. Nodes, links, their status, features and in general the complete map of the network are available, while in road networks drivers are ultimately in control of routing. Despite these essential differences the diversity of ITS technologies developed during the last decade (from adaptive cruise control to incipient autonomous driving and cooperative intelligent vehicles [130]) are converging at a microscopic level to computer networks, in an steady evolution towards fully managed road networks.

In the field of forecasting, computer network management usually requires predictions of future demand at different parts of the network, which are frequently based on the same principles than those used for road traffic predictions. In computer networks predictive information is used to rearrange their topology or route packages through the most cost-effective links (with *cost* often defined in familiar terms, e.g. end-to-end delay), while in road network traffic forecasts are used to inform drivers – who have the last decision – and to inform road managers for their traffic management duties. A more effectively and controlled infrastructure would make forecasts more relevant, hence the way traffic predictions are obtained and used in computer networks should motivate new forecasting models and applications in road networks. For instance, ARIMA and Artificial Neural Networks (ANN) are also mainly used methods in network traffic prediction [131]–[133]. Prediction methods and techniques have evolved in similar ways, and it can be possible to take some the more advanced models in network traffic prediction and apply them to road traffic prediction, subsequently triggering data-based traffic management strategies such as traffic rerouting [134].

IV. Conclusions

Traffic variables have been object of analysis and predictions for more than 40 years. Several surveys have examined the subject with different contrasting criteria and field challenge assessments. The most relevant development in the field of traffic prediction in recent years is related to a shift in the prediction modeling paradigm. Advances in data oriented techniques and technologies, explosion of Big Data and machine learning, along with growing availability of traffic related data from plentiful sources have contributed to leave behind time-series analysis methods and given a boost to data driven models. This shift has compelled most recent researchers to lay out new horizons and challenges for the field. This work has intended to compile and recapitulate previous work, to propose a comparing framework and to review most recent literature with respect to the updated criteria.

A set of common issues and challenges is found in all previous reviews, regarding prediction scope and context, most suitable model selection, metrics for different models comparison, and hybridization of models to improve performance. These aspects remain pertinent, 30 years after the first survey that exposed them, and regardless the shift to data-driven modeling. Besides, new issues and challenges arise in the light of this now prevailing way of modeling. Collecting data from new sources involves fusing different types, and managing data aggregation and resolution are some of the most recent challenges proposed by literature reviews. Additional challenges are introduced in this work: increasing the prediction horizon, incorporating exogenous factors to models, and adding data ageing mechanisms that allow models to adapt their learning to changing circumstances. The latter two are linked to data-driven modeling, and help achieving the first. Further prediction horizons are useful especially for traffic management, and having days or weeks forecast can change traffic managing measures from being reactive to being proactive. Their performance has always been considered poor, compared to short-term predictions, but these are slightly useful if there is no time for reaction. This performance can be improved when the models are data driven, by incorporating new sources of information that complete the traffic configuration scheme, and with adaptive learning methods. Another issue rises here that was anticipated since the first reviews on the field: the impact that road users having the information of future traffic status may imprint on the short-term traffic patterns themselves.

Traffic variables prediction literature has been studied thoroughly, so this review has focused in recent works. Updated reference criteria have been used to study new literature. In our inspection, the aforementioned shift to data driven models is clear: the use of ARIMA and other time-series analysis methods is lessening, and first works with prediction horizons longer than 60 minutes appear, laying the foundations for future work in this line. Models with concept drift or adaptive learning are also found in a small share of the reviewed works, but the progressive incorporation of models based in large databases which initial knowledge changes in time and requires adaptation, might precipitate this technique to be dominant. Exogenous factors are yet scantly introduced in traffic forecasting models. Their convenience has been specially proven when using calendar and time of day information as model parameters, but many other data are more available every day, which could boost traffic prediction performance of future models.

Acknowledgments
This work has been funded in part by the Basque Government under the ELKARTEK program (BID3ABI project), as well as by the European Commission under the Horizon 2020 program (REPLICATE project, grant ref. 691735).

About the Authors

Ibai Laña received his B.Sc. degree in Computer Engineering from Deusto University, Spain, in 2006, and the M. Sc. degree in Advanced Artificial Intelligence from UNED, Spain, in 2014. He is currently a senior researcher at TECNALIA (Spain), and a student pursuing his PhD at the Department of Communications Engineering of the University of the Basque Country (UPV/EHU). His research interests fall within the intersection of intelligent transportation systems, machine learning, traffic data analysis and data science.

Javier Del Ser (M'07-SM'12) received his first PhD in Telecommunication Engineering (Cum Laude) from the University of Navarra, Spain, in 2006, and a second PhD in Computational Intelligence (Summa Cum Laude) from the University of Alcala, Spain, in 2013. He is currently a principal researcher in data analytics and optimization at TECNALIA (Spain), a visiting fellow at the Basque Center for Applied Mathematics (Spain) and an adjunct professor at the University of the Basque Country UPV/EHU. His research activity gravitates on the use of descriptive, prescriptive and predictive algorithms for data mining and optimization in a diverse range of application fields such as Energy, Transport, Telecommunications and Security, among others. In these fields he has published more than 180 articles and conferences, co-supervised 6 Ph.D. theses, edited 4 books and coinvented 6 patents.

Manuel Vélez (M'03) Manuel Velez received the M.Sc. and Ph.D. degrees in telecommunication engineering from the University of the Basque Country in 1993 and 2008, respectively. In 1995 he joined the University of the Basque Country, where he is currently an Assistant Professor with the Department of Communications Engineering. In 2008, he was a visiting post-doctoral researcher at the Department of Information Technology of University of Turku, Finland. He has been involved in several research projects concerning broadcasting and wireless systems for 20 years and has supervised several PhD dissertations in this field. He is coauthor of more than 90 papers in international journals and conferences. He currently serves as an Associate Editor for the SET International Journal of Broadcast Engineering. His research interests are related to new broadcasting standards and cooperative wireless communications, in the areas of software defined radio and data analytics for spectrum and mobility management.

Eleni I. Vlahogianni (M'07) received the Diploma degree in civil engineering and the Ph.D. degree from National Technical University of Athens, Athens, Greece, specializing in traffic operations. She has been a Visiting Scholar with the Institute of Transportation Studies, University of California, Berkeley, CA, USA, working on the development of an arterial performance measure system. She is currently an Assistant Professor with the Department of Transportation Planning and Engineering, National Technical University of Athens. Her professional and research experience includes projects and consultancies, in national and European levels, focusing on urban traffic flow management, public transport, and traffic safety. Moreover, she has authored over 80 publications in journals and conference proceedings. Her primary research field is traffic flow analysis and forecasting. Her other research fields are nonlinear dynamics, statistical modeling, and advanced data mining techniques.

References

[1] B. Van Arem, H. R. Kirby, M. J. M. Van Der Vlist, and J. C. Whittaker, "Recent advances and applications in the field of short-term traffic forecasting," *Int. J. Forecasting*, vol. 13, no. 1, pp. 1–12, 1997.

[2] M. S. Ahmed and A. R. Cook, "Analysis of freeway traffic time-series data by using Box–Jenkins techniques," no. 722, 1979.

[3] M. Levin and Y. Tsao, "On forecasting freeway occupancies and volumes," *Transport. Res. Rec.*, no. 773, pp. 47–49, 1980.

[4] C. K. Moorthy and B. G. Ratcliffe, "Short term traffic forecasting using time series methods," *Transport. Plan. Technol.*, vol. 12, no. 1, pp. 45–56, 1988.

[5] N. L. Nihan and K. O. Holmesland, "Use of the Box and Jenkins time series technique in traffic forecasting," *Transportation*, vol. 9, no. 2, pp. 125–143, 1980.

[6] I. Okutani and Y. J. Stephanedes, "Dynamic prediction of traffic volume through Kalman filtering theory," *Transport. Res. Part B Methodological*, vol. 18, no. 1, pp. 1–11, 1984.

[7] E. I. Vlahogianni, J. C. Golias, and M. G. Karlaftis, "Short-term traffic forecasting: Overview of objectives and methods," *Transport Rev.*, vol. 24, no. 5, pp. 533–557, 2004.

[8] M. G. Karlaftis and E. I. Vlahogianni, "Statistical methods versus neural networks in transportation research: Differences, similarities and some insights," *Transport. Res. Part C Emerging Technol.*, vol. 19, no. 3, pp. 387–399, 2011.

[9] F. Su, H. Dong, L. Jia, Y. Qin, and Z. Tian, "Long-term forecasting oriented to urban expressway traffic situation," *Adv. Mech. Eng.*, vol. 8, no. 1, pp. 1–16, 2016.

[10] C. van Hinsbergen, J. van Lint, and F. Sanders, "Short term traffic prediction models," in *Proc. 14th World Congress Intelligent Transportation Systems*, Beijing, 2007.

[11] H. J. van Lint and C. P. van Hinsbergen, "Short-term traffic and travel time prediction models," *Artif. Intell. Applicat. Crit. Transport. Issues*, vol. 22, pp. 22–41, 2012.

[12] A. Bhaskar, E. Chung, and A. G. Dumont, "Fusing loop detector and probe vehicle data to estimate travel time statistics on signalized urban networks," *Computer-Aided Civil Infrastructure Eng.*, vol. 26, no. 6, pp. 433–450, 2011.

[13] E. Bolshinsky and R. Freidman, "Traffic flow forecast survey," Technion—Israel Institute of Technology, Tech. Rep., 2012.

[14] L. Dimitriou, T. Tsekeris, and A. Stathopoulos, "Adaptive hybrid fuzzy rule-based system approach for modeling and predicting urban traffic flow," *Transport. Res. Part C Emerging Technol.*, vol. 16, no. 5, pp. 554–573, 2008.

[15] M. G. Karlaftis and E. I. Vlahogianni, "Statistical methods versus neural networks in transportation research: Differences, similarities and some insights," *Transport. Res. Part C Emerging Technol.*, vol. 19, no. 3, pp. 387–399, 2011.

[16] G. Marfia and M. Roccetti, "Vehicular congestion detection and short-term forecasting: A new model with results," *IEEE Trans. Veh. Technol.*, vol. 60, no. 7, pp. 2936–2948, 2011.

[17] E. I. Vlahogianni, M. G. Karlaftis, and J. C. Golias, "Optimized and meta-optimized neural networks for short-term traffic flow prediction: A genetic approach," *Transport. Res. Part C Emerging Technol.*, vol. 13, no. 3, pp. 211–234, 2005.

[18] H. Lin, R. Zito, and M. Taylor, "A review of travel-time prediction in transport and logistics," in *Proc. Eastern Asia Society for Transportation Studies*, Bangkok, 2005, vol. 5, pp. 1433–1448.

[19] S. Cohen and Z. Christoforou, "Travel time estimation between loop detectors and FCD: A compatibility study on the Lille Network, France," *Transport. Res. Proc.*, vol. 10, pp. 245–255, 2015.

[20] X. Fei, C. C. Lu, and K. Liu, "A Bayesian dynamic linear model approach for real-time short-term freeway travel time prediction," *Transport. Res. Part C Emerging Technol.*, vol. 19, no. 6, pp. 1306–1318, 2011.

[21] C. J. Lu, "An adaptive system for predicting freeway travel times," *Int. J. Inform. Technol. Decision Making*, vol. 11, no. 04, pp. 727–747, 2012.

[22] F. Zheng and H. Van Zuylen, "Urban link travel time estimation based on sparse probe vehicle data," *Transport. Res. Part C Emerging Technol.*, vol. 31, pp. 145–157, 2013.

[23] L. Moreira-Matias, J. Mendes-Moreira, J. de Sousa, and J. Gama, "Improving mass transit operations by using AVL-based systems: A survey," *IEEE Trans. Intell. Transport. Syst.*, vol. 16, no. 4, pp. 1636–1653, 2015.

[24] E. Castillo, Z. Grande, A. Calviño, W. Y. Szeto, and H. K. Lo, "A state-of-the-art review of the sensor location, flow observability, estimation, and prediction problems in traffic networks," *J. Sensors*, vol. 2015, 2015.

[25] E. I. Vlahogianni, M. G. Karlaftis, and J. C. Golias, "Short-term traffic forecasting: Where we are and where we're going," *Transport. Res. Part C Emerging Technol.*, vol. 43, pp. 3–19, 2014.

[26] U. Mori, A. Mendiburu, M. Álvarez, and J. A. Lozano, "A review of travel time estimation and forecasting for advanced traveller information systems," *Transportmetrica A Transport Sci.*, vol. 11, no. 2, pp. 119–157, 2015.

[27] S. Oh, Y. Byon, K. Jang, and H. Yeo, "Short-term travel-time prediction on highway: A review of the data-driven approach," *Transport Rev.*, vol. 35, no. 1, pp. 4–32, 2015.

[28] A. Stathopoulos and M. G. Karlaftis, "A multivariate state space approach for urban traffic flow modeling and prediction," *Transport. Res. Part C Emerging Technol.*, vol. 11, no. 2, pp. 121–135, 2003.

[29] Y. Lv, Y. Duan, W. Kang, Z. Li, and F. Y. Wang, "Traffic flow prediction with big data: A deep learning approach," *IEEE Trans. Intell. Transport. Syst.*, vol. 16, no. 2, pp. 865–873, 2015.

[30] S. Hoogendoorn and P. Bovy, "State-of-the-art of vehicular traffic flow modelling," *Proc. Inst. Mech. Eng. Part I J. Syst. Control Eng.*, vol. 215, no. 4, pp. 283–303, 2001.

[31] R. Chrobok, O. Kaumann, J. Wahle, and M. Schreckenberg, "Different methods of traffic forecast based on real data," *Eur. J. Oper. Res.*, vol. 155, no. 3, pp. 558–568, 2004.

[32] I. Laña, J. Del Ser, and I. Olabarrieta, "Understanding daily mobility patterns in urban road networks using traffic flow analytics," in *Proc. IEEE Network Operations and Management Symp.*, Istanbul, Turkey, 2016.

[33] N. E. Faouzi, H. Leung, and A. Kurian, "Data fusion in intelligent transportation systems: Progress and challenges – A survey," *Inform. Fusion*, vol. 12, no. 1, pp. 4–10, 2011.

[34] N. E. Faouzi and L. A. Klein, "Data fusion for ITS: Techniques and research needs," *Transport. Res. Proc.*, vol. 15, pp. 495–512, 2016.

[35] Y. Kamarianakis and P. Prastacos, "Forecasting traffic flow conditions in an urban network: Comparison of multivariate and univariate approaches," *Transport. Res. Rec.*, vol. 1857, no. 1, pp. 74–84, 2003.

[36] C. Zhang, S. Sun, and G. Yu, "A Bayesian network approach to time series forecasting of short-term traffic flows," in *Proc. 7th Int. IEEE Conf. Intelligent Transportation Systems*, Washington, DC, 2004, pp. 216–221.

[37] H. Yin, S. Wong, J. Xu, and C. Wong, "Urban traffic flow prediction using a fuzzy-neural approach," *Transport. Res. Part C Emerging Technol.*, vol. 10, no. 2, pp. 85–98, 2002.

[38] M. Annunziato and S. Pizzuti, "A smart-adaptive-system based on evolutionary computation and neural networks for the on-line short-term urban traffic prediction," in *Proc. European Symp. Intelligent Technologies Hybrid Systems and Their Implementation on Smart Adaptive Systems*, Aachen, 2004.

[39] H. Dia, "An object-oriented neural network approach to short-term traffic forecasting," *Eur. J. Oper. Res.*, vol. 131, no. 2, pp. 253–261, 2001.

[40] W. C. Hong, Y. Dong, F. Zheng, and C. Y. Lai, "Forecasting urban traffic flow by SVR with continuous ACO," *Appl. Math. Model.*, vol. 35, no. 3, pp. 1282–1291, 2011.

[41] E. I. Vlahogianni, "Optimization of traffic forecasting: Intelligent surrogate modeling," *Transport. Res. Part C Emerging Technol.*, vol. 55, pp. 14–23, 2015.

[42] B. L. Smith, B. M. Williams, and R. K. Oswald, "Comparison of parametric and nonparametric models for traffic flow forecasting," *Transport. Res. Part C Emerging Technol.*, vol. 10, pp. 303–321, 2002.

[43] D. I. Tselentis, E. I. Vlahogianni, and M. G. Karlaftis, "Improving short-term traffic forecasts: To combine models or not to combine?" *IET Intell. Transport. Syst.*, vol. 9, no. 2, pp. 193–201, 2014.

[44] P. Cai, Y. Wang, G. Lu, P. Chen, C. Ding, and J. Sun, "A spatiotemporal correlative K-nearest neighbor model for short-term traffic multistep forecasting," *Transport. Res. Part C Emerging Technol.*, vol. 62, pp. 21–34, 2016.

[45] J. Zhong and S. Ling, "Key factors of K-nearest neighbors nonparametric regression in short-time traffic flow forecasting," in *Proc. 21st Int. Conf. Industrial Engineering and Engineering Management*, Bali, 2015, pp. 9–12.

[46] P. Dell'acqua, F. Bellotti, R. Berta, and A. De Gloria, "Time-aware multivariate nearest neighbor regression methods for traffic flow prediction," *IEEE Trans. Intell. Transport. Syst.*, vol. 16, no. 6, pp. 3393–3402, 2015.

[47] D. Xia, B. Wang, H. Li, Y. Li, and Z. Zhang, "A distributed spatial-temporal weighted model on MapReduce for short-term traffic flow forecasting," *Neurocomputing*, vol. 179, pp. 246–226, 2016.

[48] A. Abadi, T. Rajabioun, and P. A. Ioannou, "Networks with limited traffic data," *IEEE Trans. Intell. Transport. Syst.*, vol. 16, no. 2, pp. 653–662, 2015.

[49] J. Xu, D. Deng, U. Demiryurek, C. Shahabi, and M. Van Der Schaar, "Mining the situation: Spatiotemporal traffic prediction with big data," *IEEE J. Select. Topics Signal Processing*, vol. 9, no. 4, pp. 702–715, 2015.

[50] J. Wang, I. Tsapakis, and C. Zhong, "A space-time delay neural network model for travel time prediction," *Eng. Applicat. Artif. Intell.*, vol. 52, pp. 145–160, 2016.

[51] Y. Wu, F. Chen, C. Lu, and S. Yang, "Urban traffic flow prediction using a spatio-temporal random effects model," *J. Intell. Transport. Syst. Technol. Plan. Oper.*, vol. 2450, pp. 1–12, 2015.

[52] A. Ladino, A. Kibangou, H. Fourati, and C. C. de Wit, "Travel time forecasting from clustered time series via optimal fusion strategy," in *Proc. European Control Conf.*, Aalborg, 2016.

[53] D. J. Dailey, "Travel-time estimation using cross-correlation techniques," *Transport. Res. Part B Methodological*, vol. 27, no. 2, pp. 97–107, 1993.

[54] C. J. Lan and G. A. Davis, "Real-time estimation of turning movement proportions from partial counts on urban networks," *Transport. Res. Part C Emerging Technol.*, vol. 7, no. 5, pp. 305–327, 1999.

[55] B. S. Westgate, D. B. Woodard, D. S. Matteson, and S. G. Henderson, "Large-network travel time distribution estimation for ambulances," *Eur. J. Oper. Res.*, vol. 252, no. 1, pp. 322–333, 2016.

[56] A. Bezuglov and G. Comert, "Short-term freeway traffic parameter prediction: Application of grey system theory models," *Expert Syst. Applicat.*, vol. 62, pp. 284–292, 2016.

[57] A. Salamanis, D. D. Kehagias, C. K. Filelis-Papadopoulos, D. Tzovaras, and G. A. Gravvanis, "Managing spatial graph dependencies in large volumes of traffic data for travel-time prediction," *IEEE Trans. Intell. Transport. Syst.*, vol. 17, no. 6, pp. 1678–1687, 2016.

[58] B. S. Yoo, S. P. Kang, and C. H. Park, "Travel time estimation using mobile data," in *Proc. Eastern Asia Society for Transportation Studies*, Bangkok, 2005, vol. 5, pp. 1533–1547.

[59] H. Yang, T. Dillon, and Y. P. Chen, "Evaluation of recent computational approaches in short-term traffic forecasting," in *Proc. 4th IFIP TC 12th Int. Conf. Artificial Intelligence*, Daejeon, 2015, vol. 465, pp. 108–116.

[60] S. Oh, Y. Kim, and J. Hong, "Urban traffic flow prediction system using a multifactor pattern recognition model," *IEEE Trans. Intell. Transport. Syst.*, vol. 16, no. 5, pp. 2744–2755, 2015.

[61] A. Wibisono, W. Jatmiko, H. A. Wisesa, B. Hardjono, and P. Mursanto, "Traffic big data prediction and visualization using fast incremental model trees-drift detection (FIMT-DD)," *Knowledge-Based Syst.*, vol. 93, pp. 33–46, 2016.

[62] C. Lamboley, J. C. Santucci, and M. Danech-Pajouh, "24 or 48 hour advance traffic forecast in urban and periurban environment: The example of Paris," in *Proc. 4th World Congress Intelligent Transport Systems, Mobility for Everyone*, Berlin, 1997.

[63] F. Jin and S. Sun, "Neural network multitask learning for traffic flow forecasting," in *Proc. IEEE Int. Joint Conf. Neural Networks*, Hong Kong, 2008, pp. 1897–1901.

[64] P. Næss and A. Strand, "Traffic forecasting at 'strategic', 'tactical' and 'operational' level," *J. Crit. Realism*, vol. 51, no. 2, pp. 41–48, 2015.

[65] K. Jha, N. Sinha, S. Arkatkar, and A. K. Sarkar, "A comparative study on application of time series analysis for traffic forecasting in India: Prospects and limitations," *Current Sci.*, vol. 110, no. 3, pp. 373–385, 2016.

[66] C. Kai, K. Zhang, and Y. Hamamatsu, "Traffic Information real-time monitoring based on a short-long term algorithm," in *Proc. IEEE Int. Conf. Systems Man and Cybernetics*, Taipei, 2006, pp. 651–656.

[67] X. Ma, Z. Tao, Y. Wang, H. Yu, and Y. Wang, "Long short-term memory neural network for traffic speed prediction using remote microwave sensor data," *Transport. Res. Part C Emerging Technol.*, vol. 54, pp. 187–197, 2015.

[68] L. Li, X. Su, and Y. Zhang, "Trend modeling for traffic time series analysis: An integrated study," *IEEE Trans. Intell. Transport. Syst.*, vol. 16, no. 6, pp. 1–10, 2015.

[69] Z. Hou and X. Li, "Repeatability and similarity of freeway traffic flow and long-term prediction under big data," *IEEE Trans. Intell. Transport. Syst.*, vol. 17, no. 6, pp. 1786–1796, 2016.

[70] X. He, X. Gao, Y. Zhang, Z. H. Zhou, Z. Y. Liu, B. Fu, F. Hu, and Z. Zhang, "Research of traffic flow forecasting based on the information fusion of BP network sequence," in *Proc. Int. Conf. Intelligent Science and Big Data Engineering*, Suzhou, 2015, vol. 9243, pp. 548–558.

[71] H. Pan, J. Liu, S. Zhou, and Z. Niu, "A block regression model for short-term mobile traffic forecasting," in *Proc. IEEE/CIC Int. Conf. Communications China Symp. Next Generation Networking*, Shenzhen, 2015, pp. 1–5.

[72] L. Li, X. Su, Y. Wang, Y. Lin, Z. Li, and Y. Li, "Robust causal dependence mining in big data network and its application to traffic flow predictions," *Transport. Res. Part C Emerging Technol.*, vol. 58, pp. 292–307, 2015.

[73] J. Wan, J. Liu, Z. Shao, A. V. Vasilakos, M. Imran, and K. Zhou, "Mobile crowd sensing for traffic prediction in internet of vehicles," *Sensors*, vol. 16, no. 88, pp. 1–15, 2016.

[74] F. Moretti, S. Pizzuti, S. Panzieri, and M. Annunziato, "Urban traffic flow forecasting through statistical and neural network bagging ensemble hybrid modeling," *Neurocomputing*, vol. 167, pp. 3–7, 2015.

[75] A. Csikós, Z. J. Viharos, K. B. Kis, T. Tettamanti, and I. Varga, "Traffic speed prediction method for urban networks: An ANN approach," in *Proc. Models and Technologies for Intelligent Transportation Systems*, Budapest, 2015, pp. 102–108.

[76] K. Kumar, M. Parida, and V. K. Katiyar, "Short term traffic flow prediction in heterogeneous condition using artificial neural network," *Transport*, vol. 30, no. 4, pp. 397–405, 2014.

[77] S. V. Kumar and L. Vanajakshi, "Short-term traffic flow prediction using seasonal ARIMA model with limited input data," *European Transport. Res. Rev.*, vol. 7, no. 21, pp. 1–9, 2015.

[78] X. Niu, Y. Zhu, Q. Cao, X. Zhang, W. Xie, and K. Zheng, "An online-traffic-prediction based route finding mechanism for smart city," *Int. J. Distributed Sensor Netw.*, vol. 2015, pp. 1–16, 2015.

[79] M. Meng, S. Chun-fu, W. Yiik-diew, W. Bo-bin, and L. I. Hui-xuan, "A two-stage short-term traffic flow prediction method based on AVL and AKNN techniques," *J. Central South Univ.*, vol. 22, pp. 779–786, 2015.

[80] M. Hui, L. Bai, Y. Li, and Q. Wu, "Highway traffic flow nonlinear character analysis and prediction," *Math. Probl. Eng.*, vol. 2015, pp. 1–8, 2015.

[81] J. Abdi and B. Moshiri, "Application of temporal difference learning rules in short-term traffic flow prediction," *Expert Syst.*, vol. 32, no. 1, pp. 49–64, 2015.

[82] M. Huang, "Intersection traffic flow forecasting based on ν-GSVR with a new hybrid evolutionary algorithm," *Neurocomputing*, vol. 147, pp. 343–349, 2015.

[83] C. Dong, Z. Xiong, C. Shao, and H. Zhang, "A spatial–temporal-based state space approach for freeway network traffic flow modelling and prediction," *Transportmetrica A Transport. Sci.*, vol. 9935, pp. 547–560, 2015.

[84] Z. Zhu, B. Peng, C. Xiong, and L. Zhang, "Short-term traffic flow prediction with linear conditional Gaussian Bayesian network," *J. Adv. Transport.*, pp. 1–13, 2016.

[85] J. Zhao and S. Sun, "High-order Gaussian process dynamical models for traffic flow prediction," *IEEE Trans. Intell. Transport. Syst.*, vol. 17, no. 7, pp. 2014–2019, 2016.

[86] K. Balsys and D. Eidukas, "Traffic flow detection and forecasting," *Electron. Elect. Eng.*, vol. 5, no. 5, pp. 5–8, 2010.

[87] L. Si-yan, L. De-wei, X. Yu-geng, and T. Qi-feng, "A short-term traffic flow forecasting method and its applications," *J. Shanghai Jiaotong Univ.*, vol. 20, no. 2, pp. 156–163, 2015.

[88] J. Mendes-moreira, A. M. Jorge, J. Freire de Sousa, and C. Soares, "Improving the accuracy of long-term travel time prediction using heterogeneous ensembles," *Neurocomputing*, vol. 150, pp. 428–439, 2015.

[89] J. Tang, Y. Zou, J. Ash, S. Zhang, F. Liu, and Y. Wang, "Travel time estimation using freeway point detector data based on evolving fuzzy neural inference system," *PloS ONE*, vol. 11, no. 2, pp. 1–24, 2016.

[90] Y. Wu, H. Tan, P. Jin, B. Shen, and B. Ran, "Short-term traffic flow prediction based on multilinear analysis and k-nearest neighbor regression," in *Proc. 15th Int. Conf. Transportation Professionals*, Beijing, 2015, pp. 557–569.

[91] J. Rupnik, J. Davies, B. Fortuna, A. Duke, and S. S. Clarke, "Travel time prediction on highways," in *Proc. IEEE Int. Conf. Computer and Information Technology*, Liverpool, 2015, pp. 1435–1442.

[92] Y. Hou, P. Edara, and C. Sun, "Traffic flow forecasting for urban work zones," *IEEE Trans. Intell. Transport. Syst.*, vol. 16, no. 4, pp. 1761–1770, 2015.

[93] Y. Chun-xia, F. Rui, and F. Yi-qin, "Prediction of short-term traffic flow based on similarity," *J. Highway Transport. Res. Dev.*, vol. 10, no. 1, pp. 92–97, 2016.

[94] Y. Zhang and Y. Zhang, "A comparative study of three multivariate short-term freeway traffic flow forecasting methods with missing data," *J. Intell. Transport. Syst.*, vol. 20, no. 3, pp. 205–218, 2016.

[95] X. Li and W. Gao, "Prediction of traffic flow combination model based on data mining," *Int. J. Database Theory Applicat.*, vol. 8, no. 6, pp. 303–312, 2015.

[96] H. Lu, Z. Sun, W. Qu, and L. Wang, "Real-time corrected traffic correlation model for traffic flow forecasting," *Math. Probl. Eng.*, vol. 2015, pp. 1–11, 2015.

[97] W. Ma and R. Wang, "Traffic flow forecasting research based on Bayesian normalized Elman neural network," in *Proc. IEEE Signal Processing and Signal Processing Education Workshop*, Salt Lake City, 2015, pp. 426–430.

[98] R. T. Das, K. K. Ang, and C. Quek, "ieRSPOP: A novel incremental rough set-based pseudo outer-product with ensemble learning," *Appl. Soft Comput.*, vol. 46, pp. 170–186, 2016.

[99] J. Melorose, R. Perroy, and S. Careas, "Analysis and prediction of spatiotemporal impact of traffic incidents for better mobility and safety in transportation systems," Tech. Rep., 2015.

[100] T. Ma, Z. Zhou, and B. Abdulhai, "Nonlinear multivariate time-space threshold vector error correction model for short term traffic state prediction," *Transport. Res. Part B Methodological*, vol. 76, pp. 27–47, 2015.

[101] L. Moreira-Matias and F. Alesiani, "Drift3Flow: Freeway incident prediction using real-time learning," in *Proc. IEEE 18th Int. Conf. Intelligent Transportation Systems*, Las Palmas de Gran Canaria, 2015, pp. 566–571.

[102] P. Lopez-Garcia, E. Onieva, E. Osaba, A. D. Masegosa, and A. Perallos, "A hybrid method for short-term traffic congestion forecasting using genetic algorithms and cross entropy," *IEEE Trans. Intell. Transport. Syst.*, vol. 17, no. 2, pp. 557–569, 2016.

[103] G. Fusco, C. Colombaroni, L. Comelli, and N. Isaenko, "Short-term traffic predictions on large urban traffic networks: Applications of network-based machine learning models and dynamic traffic assignment models," in *Proc. Models and Technologies for Intelligent Transportation Systems*, Budapest, 2015, pp. 3–5.

[104] N. N. Yusof, A. Mohamed, and S. Abdul-Rahman, "Soft computing in data science," *Commun. Comput. Inform. Sci.*, vol. 545, pp. 43–53, 2015.

[105] Y. Cong, J. Wang, and X. Li, "Traffic flow forecasting by a least squares support vector machine with a fruit fly optimization algorithm," *Proc. Eng.*, vol. 137, pp. 59–68, 2016.

[106] S. Jeon and B. Hong, "Monte Carlo simulation-based traffic speed forecasting using historical big data," *Future Gener. Comput. Syst.*, 2015.

[107] Y. Zhang and A. Haghani, "A gradient boosting method to improve travel time prediction," *Transport. Res. Part C Emerging Technol.*, vol. 58, pp. 308–324, 2014.

[108] H. Yang and X. Hu, "Wavelet neural network with improved genetic algorithm for traffic flow time series prediction," *Optik*, vol. 127, pp. 8103–8110, 2016.

[109] I. Laña, J. Del Ser, M. Vélez, and I. Oregi, "Joint feature selection and parameter tuning for short-term traffic flow forecasting based on heuristically optimized multi-layer neural networks," in *Proc. Int. Conf. Harmony Search Algorithm*, 2017, pp. 91–100.

[110] Z. Ma, G. Luo, and D. Huang, "Short term traffic flow prediction based on on-line sequential extreme learning machine," in *Proc. 8th Int. Conf. Advanced Computational Intelligence*, Chiang Mai, 2016, pp. 143–149.

[111] G. Zhu, K. Song, and P. Zhang, "A travel time prediction method for urban road traffic sensors data," in *Proc. Int. Conf. Identification Information and Knowledge in the Internet of Things*, Beijing, 2015, pp. 29–32.

[112] L. Moreira-Matias, O. Cats, J. Gama, J. Mendes-Moreira, and J. F. de Sousa, "An online learning approach to eliminate bus bunching in real-time," *Appl. Soft Comput.*, vol. 47, pp. 460–482, 2016.

[113] Y. Xu, H. Chen, Q. Kong, X. Zhai, and Y. Liu, "Urban traffic flow prediction: A spatio-temporal variable selection-based approach," *J. Adv. Transport.*, vol. 50, pp. 489–506, 2016.

[114] I. Laña, J. Del Ser, A. Padró, M. Vélez, and C. Casanova-Mateo, "The role of local urban traffic and meteorological conditions in air pollution: A data-based case study in Madrid, Spain," *Atmos. Environ.*, vol. 145, pp. 424–438, 2016.

[115] H. Park and A. Haghani, "Real-time prediction of secondary incident occurrences using vehicle probe data," *Transport. Res. Part C Emerging Technol.*, 2014.

[116] F. C. Pereira, F. Rodrigues, and M. Ben-akiva, "Using data from the web to predict public transport arrivals under special events scenarios," *J. Intell. Transport. Syst.*, vol. 1, pp. 37–41, Apr. 2015.

[117] S. C. Calvert, J. W. C. Van Lint, and S. P. Hoogendoorn, "A hybrid travel time prediction framework for planned motorway roadworks," in *Proc. IEEE Conf. Intelligent Transportation Systems*, Madeira Island, 2010, pp. 1770–1776.

[118] S. Yousif, "Motorway roadworks: Effects on traffic operations," *Highways Transport.*, pp. 20–22, 2002.

[119] Z. Zhou, X. Yang, and P. Zheng, "A method for delay estimation using traffic monitoring data," in *Proc. Int. Conf. Automatic Control and Artificial Intelligence*, Xiamen, 2012, pp. 108–111.

[120] I. Žliobaitė, M. Pechenizkiy, and J. Gama, "An overview of concept drift applications," in *Big Data Analysis: New Algorithms for a New Society*. New York: Springer, 2016, pp. 91–114.

[121] G. Souto and T. Liebig, "On event detection from spatial time series for urban traffic applications," no. 9580, 2016.

[122] E. I. Vlahogianni, M. G. Karlaftis, and J. C. Golias, "Statistical methods for detecting nonlinearity and non-stationarity in univariate short-term time-series of traffic volume," *Transport. Res. Part C Emerging Technol.*, vol. 14, no. 5, pp. 351–367, 2006.

[123] Y. Kamarianakis, H. O. Gao, and P. Prastacos, "Characterizing regimes in daily cycles of urban traffic using smooth-transition regressions," *Transport. Res. Part C Emerging Technol.*, vol. 18, no. 5, pp. 821–840, 2010.

[124] E. I. Vlahogianni and M. G. Karlaftis, "Temporal aggregation in traffic data: Implications for statistical characteristics and model choice," *Transport. Lett.*, vol. 3, no. 1, pp. 37–49, 2011.

[125] J. Tang, Y. Wang, H. Wang, S. Zhang, and F. Liu, "Dynamic analysis of traffic time series at different temporal scales: A complex networks approach," *Phys. A Statist. Mech. Applicat.*, vol. 405, pp. 303–315, 2014.

[126] R. Elwell and R. Polikar, "Incremental learning of concept drift in nonstationary environments," *IEEE Trans. Neural Netw.*, vol. 22, no. 10, pp. 1517–1531, 2011.

[127] E. I. Vlahogianni, *Computational Intelligence and Optimization for Transportation Big Data: Challenges and Opportunities*. Springer International Publishing, 2015.

[128] S. Gábor and I. Csabai, "The analogies of highway and computer network traffic," *Phys. Statist. Mech. Applicat.*, vol. 307, no. 3, pp. 516–526, 2002.

[129] S. Clement, N. Vogiatzis, M. Daniel, S. Wilson, and R. Clegg, "Road networks of the future: Can data networks help?" in *Proc. 28th Australasian Transport Res. Forum*.

[130] J. Baber, J. Kolodko, T. Noel, M. Parent, and L. Vlacic, "Cooperative autonomous driving: Intelligent vehicles sharing city roads," *IEEE Robot. Automat. Mag.*, vol. 12, no. 1, pp. 44–49, 2005.

[131] H. Feng and Y. Shu, "Study on network traffic prediction techniques," in *Proc. Int. Conf. Wireless Communications Networking Mobile Computing*, Wuhan, 2005, vol. 2, pp. 995–998.

[132] A. Sang and S. Li, "A predictability analysis of network traffic," *Comput. Netw.*, vol. 39, no. 4, pp. 329–345, 2002.

[133] Y. Chen, B. Yang, and Q. Meng, "Small-time scale network traffic prediction based on flexible neural tree," *Appl. Soft Comput. J.*, vol. 12, no. 1, pp. 274–279, 2012.

[134] S. Salcedo-Sanz, D. Manjarres, Á. Pastor-Sánchez, J. Del Ser, J. A. Portilla-Figueras, and S. Gil-Lopez, "One-way urban traffic reconfiguration using a multi-objective harmony search approach," *Expert Syst. Applicat.*, vol. 40, no. 9, pp. 3341–3350, 2013.

ITS