Deliverable 4.1

Exploring prediction perspectives

Grant Agreement number:	688082
Project acronym:	SETA
Project title:	An open, sustainable, ubiquitous data and service ecosystem for efficient, effective, safe, resilient mobility in metropolitan areas
Funding Scheme:	H2020-ICT-2015

Authors
DEL, TSS, USF and AIZ

Internal Reviewer
Ben Heller B.Heller@SHU.AC.UK

State:	Final
Distribution:	Public

Deliverable History

Date	Author	Changes
02-03-2016	Oded Cats	Draft table of contents
06-04-2016	All partners	Synopsis of partners' contribution
23-05-2016	All partners	First round contribution
01-06-2016	TU Delft	Comments to partners
27-06-2016	TU Delft	Complete draft version for partners final review
18-07-2016	TU Delft	Final version following partners' input and internal review comments

Contents

Summary

This report presents an overview of the state-of-the-art of short-term traffic prediction techniques from two perspectives - data-driven and simulation (model)-based. Given that a great number of different predictive approaches have been developed and tested in the past three decades, the adopted taxonomy in this report is in line with other researchers' reviews on this topic, and moreover, matches well with SETA's objective of developing different use cases.

In the data-driven part, six specific groups of methods that are more inherently similar are presented, including linear regression and its variations, coordinate transformation methods/dimensionality reduction methods, neural networks, decision trees, K-nearest neighbour and scalable Bayesian inference. In each group, basic principles of the methods are described first, followed by an overview of existing literature based on such method. Some new and novel methods that are beyond the scope of these six groups are also briefed, given the upward trend of applying new predictive algorithms and approaches to short-term traffic forecasting.

In the simulation (model)-based part, the concept of simulation-based traffic prediction is highlighted followed by a short review of different levels of model representation. Overall, simulation (model)-based approaches adopt a very different path towards successful traffic prediction than data-driven methods, since it is highly dependent on the knowledge of traffic and travellers' behaviour in a sense that prediction is achieved by having traffic, involving all types of agents and parameters, evolve over time under the governance of a set of traffic models. These models, as significant components of traffic simulation, include traffic flow models, demand estimation and prediction models, and transport assignment models, are also described and discussed in the context of the latest research progress.

The state of the art reviewed in this report is instrumental in identifying the most promising avenues for developing short-term traffic predictions in the context of the SETA project. The report concludes with a discussion of the data availability of requirements for enabling the predictions envisioned in this project and sketching a plan for developing both data-driven and simulation-based predictors.

Glossary of Terms

ABC	*Approximate Bayesian Computation*
ANN	*Artificial Neural Networks*
AR	*Auto Regressive*
ARIMA	*Auto Regressive Integrated Moving Average*
BEM	*Basic Ensemble Method*
DBMS	*Database Management System*
DBN	*Deep Belief Networks*
DLM	*Dynamic Linear Models*
DM	*Data Mining*
DNN	*Deep Neural Networks*
DTA	*Dynamic Traffic Assignment*
DTC	*Dynamic Tensor Completion*
GARCH	*Generalized Auto Regressive Conditional Heteroscedasticity*
GEM	*Generalized Ensemble Method*
IMF	*Intrinsic Mode Function*
ITS	*Intelligent Transport Systems*
K-NN	*K-nearest neighbour*
MA	*Moving Average*
MCMC	*Markov Chain Monte Carlo*
MSA	*Method of Successive Averages*
OD	*Origin Destination*
PCA	*Principal Component Analysis*
PDF	*Posterior Density Function*
SDE	*Stochastic Differential Equation*
SVM	*Support Vector Machines*
SVR	*Support Vector Regression*
WP	*Work Package*

1 Introduction

1.1 Background

Traffic congestion has become one of most severe issues in many cities worldwide, while simply building more transport infrastructure does not turn out to be an ultimate solution. The function of traffic management systems is to efficiently manage existing transportation resources; the goals are to maximise safety and transport productivity, minimise congestion and damage through incidents, distribute information on traffic-conditions, road-conditions, weather, etc. The challenging tasks of traffic modelling and forecasting aim at developing reactive and even proactive systems that allow for dynamic traffic management and timely incident handling. This calls for the development of short-term mobility predictions.

Single-point traffic prediction systems can be trained to estimate the traffic flow at a given point (e.g. intersections, tollbooths) given the history of traffic levels and correlated data (e.g. weather conditions). Network-level traffic prediction systems consider the detection points as nodes in a graph and model the propagation of traffic flow within the whole network: those systems are more powerful but they require a model of the roadmap - it may also be necessary to simplify the network using graph reduction techniques, such as 'collapsing' a block or district to a single point in the network for better performance.

There are many applications of these systems. The first example is the aforementioned dynamic traffic management, that is, re-routing traffic through less congested paths, or regulating traffic lights' cycle in order to better accommodate the traffic flow. As a second example, context-aware navigation apps would be able to estimate and compare traffic conditions along alternative paths towards the selected destination. An important by-product of a traffic management system lies in the collected data, which in turn can be used for a multitude of purposes, ranging from transport (service) planning and design, policy evaluation and monitoring, and uses beyond the mobility domain itself, e.g. asset management, city planning, etc.

The novelty of SETA in this domain is manifold in that it spans from the involvement of citizens as active data-providers to the development of sensors, and from the fusion of heterogeneous data sources to the application of cutting-edge machine learning algorithms and traffic simulation models. Living labs, data acquisition and integration are the core contributions of WP1, WP2 and WP3.

1.2 Purpose, Scope and relation to other SETA Work Packages

Figure 1 depicts a schematic representation of the ideal traffic (or transport) control cycle, in which three tasks are closely intertwined. These tasks include (a) traffic state estimation, in which data from diverse traffic sensors (inductive loop detectors, cameras, probe vehicles, etc.) and traffic flow models are used to reconstruct a network-wide picture of the traffic state (e.g., in terms of traffic densities and/or speeds); (b) traffic state prediction, in which a projection of current traffic in the future is computed; and (c) the optimization of traffic control measures (e.g., algorithms such as route guidance, dynamic speed-limits, or ramp metering, etc.), the results of which are fed back into the traffic system using actuators (traffic lights, reversible lanes, information panels, etc). Roughly speaking, traffic estimation is addressed in SETA WP3; traffic prediction is the subject of SETA WP4; and optimization (e.g. developing routing engines based on predicted info) is jointly covered by WP4 and WP5.

The examples in the section above relate to car traffic, but it will be clear that a similar picture as in Figure 1 can be identified for public transport, or multi-modal transport systems in general. Even information provision (about e.g. route or mobility options in terms of travel

time and costs) can be understood as an actuator in a control system. In this case actuation depends upon usage of and sensitivity to the information (or compliance with provided alternatives). The bottom line is that accurate, timely and reliable traffic estimation and prediction are key to successful traffic and transport control and information systems.

As noted, traffic state estimation refers to reconstructing state variables (i.e. variables with which the dynamics of a traffic or transport system can be described) from the available sensor data and is the primary focus of SETA WP 3. It is complex due to the heterogeneity of sensor data in terms of semantics, availability, reliability and accuracy. In many cases those variables that are of most interest and relevance for management and control (e.g. densities, space mean speeds) are poorly (if at all) observable through sensor data and need to be derived from the variables that are measured. In some cases, estimating key state or input variables poses a highly underdetermined problem for which many assumptions need to be made. Clearly, the reliability and accuracy of measurement data can also have a major impact on traffic prediction, which is the central objective of this work package (SETA WP4). This impact is direct or indirect or both – this depends on the type of prediction approach that is taken. This we discuss below.

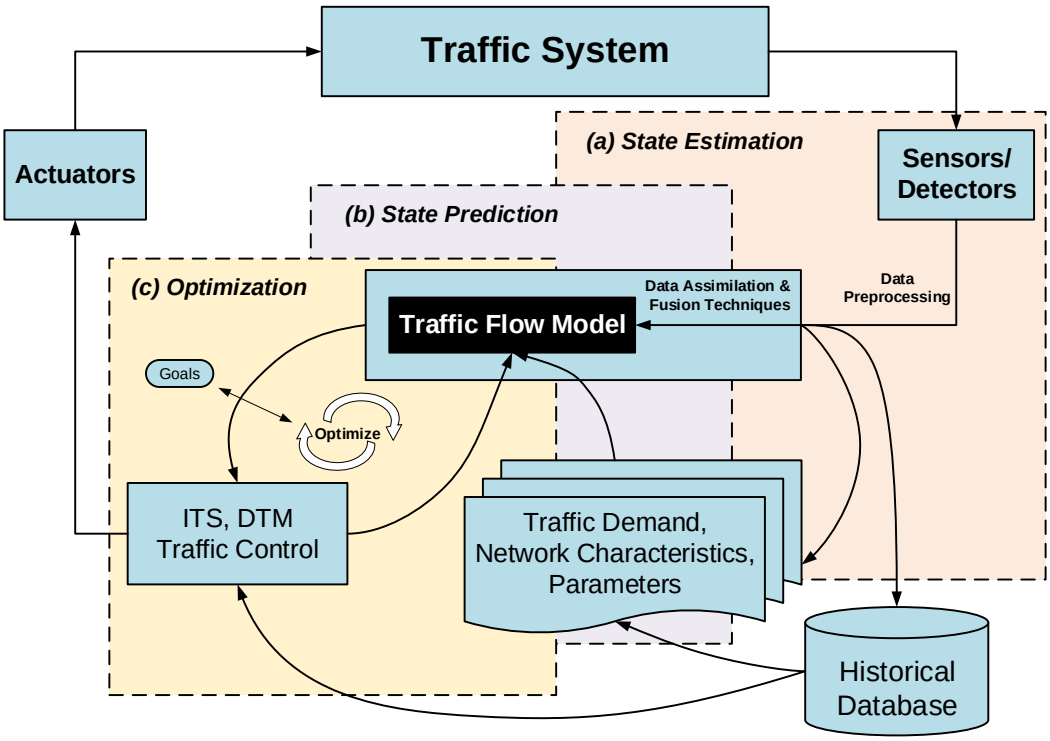


Figure 1 A schematic representation of an ideal model-based decision support system. Adapted from Yuan (2013)*. Roughly speaking, traffic estimation is addressed in SETA WP3; traffic prediction is the subject of SETA WP4; and optimization (e.g. developing routing engines based on predicted info) is jointly covered by WP4 and WP5.*

1.3 Motivation for a Taxonomy of Short-term Traffic Prediction Approaches that match SETA's needs

Short-term traffic prediction has been one of the main research topics since the early 1980s. A large number of different methods have been developed by researchers from all over the world as advanced techniques emerge and better data sources become available. There are many different reviews on these methods with different ways to categorize them. For example, Vlahogianni et al. (2004) reviewed short-term traffic forecasting studies up to 2003 from three perspectives: determination of scope, modelling and conceptual output specification. Van Lint and Van Hinsbergen (2012) review all methods based on a taxonomy that considers three branches: naïve, data-driven and model-based approaches.

Vlahogianni et al. (2014) review publications between 2004 and 2013 and identify existing challenges as well as future directions. Oh et al. (2015) present a dedicated review on data-driven short-term traffic prediction studies. In this paper a taxonomy of the data driven approaches for travel time prediction is provided.

In this report, we use a simple two-branch taxonomy that is in line with the overviews above and that matches well with SETA's objective of developing different user cases. These use cases differ in two dimensions. We differentiate between *professional* users (e.g. road operators) and *end-users* (travellers) and between *offline* ("what-if" questions and evaluation) and *real-time* applications.

From the literature we know that *theoretically,* (simulation) model-based approaches provide the ideal solution for traffic prediction for reasons that become apparent when looking at Figure 1. A network traffic / transport model provides a single multi-purpose network-wide solution. It can be used for both network-wide estimation and prediction purposes, and most importantly, it also serves as the computational engine for control-optimization and "what if" (scenario) type inference. This is particularly relevant for the SETA use cases that focus on transportation professionals (traffic managers, control operators, urban planners, etc). However, there is a downside. A full-fletched model-based solution for city networks like Santander and particularly Turin and Birmingham is a highly complex and labour-intensive solution (i.e. network coding, data gathering, model calibration). This is because model-based approaches require many inputs that need to be derived from data:

- Consistent graphs that represent the road infrastructure – deriving and keeping a large graph consistent with actual infrastructure is a huge challenge. A single mistake can, however, lead to huge estimation and prediction errors.
- Boundary conditions, most importantly an initial state (estimate). A model based prediction engine thus by definition requires a state estimation engine to reconstruct the initial state (or more generally the prevailing state based on the latest sensor information). This initial state usually encompasses variables that are difficult to observe directly (e.g. densities, queues), so a high quality state estimation algorithm to reconstruct these from observable data is crucial.
- Inputs, of which some must also be estimated from sensor data origin-destination flows, route choice patterns, etc) and some come from other sources (e.g. public transport schedules)
- Parameters of the mathematical (traffic flow) models for driving (walking) and traveling behaviour. Also these need to be properly tuned on the basis of historical data and/or physical considerations

Clearly, a model based approach is only as good as the quality of each of these inputs. Data driven approaches, in which the correlations between link "traffic states" are not determined by consistent (physical) propagation over a network graph but through statistical modelling (regression, clustering, etc) with generic mathematical models (e.g. neural networks); are much less labour intensive; less "meta-data hungry"; and more robust to missing and faulty input data. Moreover, they can be used to mine and explore the new SETA data. The downside here evidently, is that data driven models are only as good as the degree in which the training data is representative for the traffic dynamics in a network. It is not possible to do "what if" reasoning in case e.g. a bridge is removed or added from or to the network or in case the bus-frequencies are doubled.

Within SETA we opt for a dual approach that uses the best of both worlds (data driven and model-based), and that allows for growth during and after the project.

- For the professional use cases we develop a model-based solution, in which as many of the inputs as possible will be automatically derived from data. For scale

reasons, we focus first on small (corridor) networks to develop, test and verify a model-based solution.

- For the end-user use cases, we focus from the onset on data driven solutions that cover the entire network. This is crucial, because in these cases we require (predictive) information (multiple route options) across the entire network.

The final SETA solution for estimation and prediction then becomes a (prototype) hybrid system with both data-driven ("black-box") and model-based ("white-box") components (Figure 2). In this hybrid approach data-driven methods are used for predicting traffic states on parts of the network not used in the professional use cases, and for those quantities (e.g. demand and other input quantities) that are notoriously difficult to estimate from "classic" traffic data. This hybrid solution is modular and scalable.

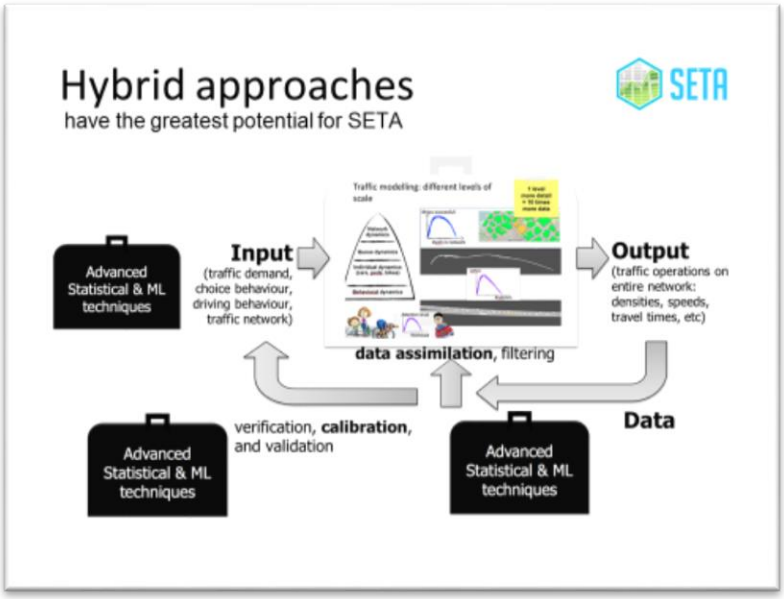


Figure 2: a hybrid approach to traffic estimation and prediction: simulation model-based core and data driven assimilation, estimation and calibration methods.

1.4 Report outline

The remainder of this report is organized as follows. Chapter 2 provides a review of data-driven prediction approaches which are relevant for predicting traffic and travel demand. The review focuses on approaches grouped into the following categories: linear regression, coordinated transformation, neural networks, decision trees, K-nearest neighbours and Bayesian inference. In particular, applications in the transport domain are described. Chapter 3 describes the simulation-based approach to traffic and travel demand predictions with an overview of traffic simulation models, demand estimation and prediction, transport assignment models and model calibration and validation. Special consideration is given to the requirements for short-term predictions and the implications of modelling choices. This report concludes with a discussion on data requirements and availability, and suggests directions for developing prediction schemes in the upcoming research and development activities in the context of the SETA project.

2 Data-driven Approaches to Mobility Predictions

A data-driven approach predicts traffic conditions by estimating the current or future traffic state from historical traffic patterns, without detailed descriptions of intrinsic network dynamics based on traffic flow mechanisms.

Fundamentally, the goal of prediction is to estimate some future values based on current and past data samples. Mathematically stated:

$$\hat{x}(n+\Delta_n) = f(x(n-a), x(n-b)x(n-c), \dots.),$$

where, $\hat{x}$ is the predicted value of a discrete series x sampled at n space and n is a series of N discrete samples: $n = \{0,1,2, \dots N-1\}$ and $\hat{x}(n+\Delta)$ is some predicted value in the future. For a prediction algorithm, a function $f\ (x)$ is defined based on the given data that will have an output equal to the predicted value for some prediction horizon. Based on the properties of the input data and the prediction parameters, the predictions can be broadly classified as linear and non-linear.

For both linear prediction models (for which the parameters are proportional to the output), e.g. $f(\boldsymbol{x}) = \sum_i w_i g_i(\boldsymbol{x}) + w_0$; and nonlinear models (in which case the parameters are nonlinearly related to the output), e.g. $f(\boldsymbol{x}) = h(\sum_i w_i g_i(\boldsymbol{x})) + w_0$; the prediction problem boils down to finding the optimal parameter set w, and defining the criteria for finding the optimal set of weights. There are methods to assign or tune these parameters and they can be broadly be classified as supervised, unsupervised and reinforcement learning (calibration) methods.

In supervised learning, an input is presented and an output computed based on the proposed method. This is compared with the output desired for these inputs (referred to as the ground truth), and a global error function computed. This is then used to update the weights or parameters in order to move the output towards the desired output. Over the course of many examples being presented to the method, it is hoped that the global error will gradually decline, as the network converges into a steady state. The learning is described as supervised, because an exact description of the behaviour required after each iteration is given.

Unsupervised learning operates purely on input data, with all the criteria for updating the weights being determined internally within the method. The main use of such learning is for classification problems where it is not certain beforehand what the definition of the classes should be.

Reinforcement learning is a less direct way of tuning the parameters by informing whether it performed well or badly for each iteration. If the prediction was accurate, the user assigns a positive score and if wrong, a negative score, essentially teaching the predictor to perform better.

Here, we review some of the short-term prediction methods that have been employed in transport and how the parameters of the predictors are estimated and tuned.

2.1 Linear Regression and its Variations

Predictive methods based on linear regression pertain to constructing a linear relationship between desired traffic parameters and a set of correlated variables. Because of their relatively simple structures, the high computational efficiency of such methods has been widely confirmed by researchers.

Kwon et al. (2000) applied linear regression to travel time prediction on freeway using a heterogeneous dataset. In their study, it was found that the prevailing traffic states were the most influencing factor on short-term predictions, while the historical data were more useful in predictions for longer prediction horizons. Later, time-varying coefficient linear regression models were developed and applied to freeway travel time predictions (Zhang and Rice, 2003; Rice and Van Zwet, 2004). A linear relationship between the current traffic and future travel time for discrete time intervals was revealed. In the study conducted by Zhang and Rice (2003), regression analysis with varying coefficients according to the departure time was performed and they also forecasted the future travel time using data from two real-life loop detectors. Rice and Van Zwet (2004) presented a similar method in which the speed values were estimated with flow and occupancy using the unique 'g-factor'. Sun et al. (2003) developed a travel time prediction method based on local linear regression models and the main difference from Rice's studies is that the linearity between the current information and future predictions is not assumed.

In many cases, linear regression models are employed to construct more sophisticated predictors. For example, Danech-Pajouh and Aron (1991) proposed so called ATHENA model by adopting a clustering technique to group the data and assigning each cluster to a linear regression model. Although the ATHENA model produces excellent numerical performances benefiting from its layered structures, its downside is also obvious that the practical application is difficult due to its inherently complicated computational processes associated with 192 different clusters. More recently, Van Hinsbergen and Lint (2008) developed a Bayesian framework that combines two linear regression models (linear model and locally weighted model) to increase practicality in the context of real-time application. Later, a Bayesian inference-based dynamic linear model (DLM) was also developed by Fei et al. (2011) to predict online short-term travel time on a freeway stretch. The Bayesian inference-based model considers travel time as a linear combination of three components, including the median historical travel times, the variations in travel times, and the model error. Their case study shows that the developed DLM outperforms the auto-regressive (AR) model particularly for the non-recurrent traffic.

A limited number of studies can be found in the existing literature in terms of the short-term prediction of bus passenger demand. For instance, a short-term passenger demand forecasting method algorithm based on time series prediction models was proposed and further integrated in a framework for real-time demand-responsive bus dispatching control (Sheu, 2005).

A great advantage of linear regression predictive models is that they provide a parsimonious, efficient and intuitively appealing way of solving this problem. However, its disadvantage is also very obvious that such models (like all data-driven models) can in many cases only provide site-specific solutions and cannot be transferred from one application to another.

2.1.1 The ARIMA Family of Models

Methods based on Auto Regressive Integrated Moving Average (ARIMA) have been extensively applied to perform short-term traffic prediction since its first application in forecasting freeway traffic volume and occupancy time series in 1979 (Ahmed and Cook, 1979). Autoregressive (AR) modes are modes in which the value of a variable in one period is related to its values in previous periods, while moving average (MA) models account for the possibility of a relationship between a variable and the residuals from previous periods. By modelling a variable y_t as a parameterized (weighted) linear function of past observations y_{t-n} of that variable (AR) and past error terms e_t (MA), one can then predict time series. In essence, it is a special case of regression models in a sense that the input $\boldsymbol{x}_t$ is composed of past observations of y. However, the successful application of the ARMA method requires a process to be stationary, which means it has mean and variance that do

not change over time and it does not have trends. Given the fact that traffic (as well as many other realistic) processes are usually not stationary, a common solution is to incorporate a term for structural trends in the data, which leads to the ARIMA (p, d, q) model with *p* autoregressive lags, *q* moving average lags and difference in the order of *d*. For instance, when a variable y_t is not stationary, a differenced variable: $\Delta y_t = y_t - y_{t-1}$ is used for first order differences. The variable y_t is integrated of order one, denoted as $I(1)$, if taking a first difference produces a stationary process.

A structured methodology for applying ARIMA model refers to the book by Box and Jenkins (1976) which elaborates the procedures of this methodology in the iterative steps of model identification, model estimation, and model diagnosis. ARIMA models soon became a novel and prevailing approach for transportation researchers to perform traffic forecast owing to its well defined theoretical foundation and effectiveness in prediction. An early overview was presented by Nihan and Holmesland (1980). Ever since then, rapid development, extension and application of this ARIMA family of models have been witnessed until the recent significant leap to Computational Intelligence (CI) – Data Mining (DM) approaches to analysing the data (Vlahogianni et al., 2014). A number of variations and additions on ARIMA can be explored in the existing literature, including the Seasonal ARIMA (SARIMA) in which periodic terms about trends in traffic data are added (Smith et al., 2002; Williams and Hoel, 2003), subset ARIMA models which partition the time series in subsets with corresponding terms and components (Lee and Fambro, 1999), Kohonen ARIMA which combines Kohonen maps with ARIMA models to forecast traffic flow (Van Der Voort et al., 1996a), ARIMAX models which are employed to conduct multivariate traffic flow (Williams, 2001), and vector autoregressive moving average (VARMA) and space–time ARIMA (STARIMA) models (Kamarianakis and Prastacos, 2003), etc.

It was pointed out that although time domain models, including ARIMA, provide promising ability to forecast the expected value of traffic flow, the inherent volatile nature of traffic data is still left unexplained (Zhang et al., 2014). An issue in ARIMA approaches was discussed by Van Lint and Van Hinsbergen (2012) that they do not apply to departure time prediction. As explained, with time periods *k* of fixed size Δt (of typically 1, 5, or 10 min in short-term prediction problems), there is no guarantee that a previous departure travel time TT_{k-n} is available for input, with *n* typically depicting 1 or a few discrete time periods. In fact TT_{k-n} is actually measured at time period $k^* = k - n + TT_{k-n}$, which in most nontrivial cases will be later than *k*. This problem becomes rapidly worse when one considers a reasonable sized (e.g., more than a few kilometres) and congested route with travel times typically (much) larger than Δt. It implies that time series approaches could only be solutions to offline (when the data are available) travel time prediction, yet not to real (online) traffic information or control applications. When predicting departure travel time on longer routes, measured (by definition arrival) travel times are of little use as inputs (Van Lint, 2008).

The stochastic SARIMA + GARCH structure is gradually emerging as one of the promising candidates for modelling traffic flow series. GARCH refers to the Generalized Auto Regressive Conditional Heteroscedasticity (GARCH) methodology, which was first proposed by Bollerslev (1986) and has been transferred to studies on short-term prediction of traffic variability (Guo, 2005; Kamarianakis et al., 2005; Tsekeris and Stathopoulos, 2006a; Karlaftis and Vlahogianni, 2009). Tsekeris and Stathopoulos (2006a) used the standard GARCH model and a long-memory extension of it. The GARCH model can be processed using Kalman filters (Guo and Williams, 2010) so that its performance can be improved through an adaptive mechanism.

In the stochastic SARIMA + GARCH structure, the SARIMA component captures the first conditional moment, i.e., the dynamics of traffic flow levels, and the GARCH component captures the second conditional moment, i.e., the dynamics of traffic flow variances. An

adaptive Kalman filter is proposed to update the process variance to a SARIMA + GARCH prediction model (Guo et al., 2014). A new hybrid model for multi-step traffic flow forecasting in a freeway system is proposed in (Zhang et al., 2014). Their methodology consists of an assumption that traffic flow can be comprised of three different components: an intra-day or periodic trend detected by using spectral analysis, a deterministic part modelled by the ARIMA model, and the volatility part estimated by the GJR-GARCH model.

2.2 Coordinate Transformation Methods/Dimensionality Reduction Methods

In these methods, the non-linear input space needs to be mapped into a linear feature space so that linear operations can be done for prediction and estimation purposes.

2.2.1 Support Vector Machine

Support vector machines (SVMs) were proposed by Cortes and Vapnik (1995) and are based on the structural risk minimization principle (minimizing an upper bound on the generalization error). SVM has the potential to overcome the drawbacks of neural networks, and can be effective in problems of nonlinearity, small samples, high dimension, local minima and over-fitting.

The basic idea of SVM is to map the training data from the input space into a higher dimensional feature space via a kernel function and then construct a separating hyperplane with maximum margin in the feature space. The use of kernels is the key in SVM/SVR applications. It provides the capability of mapping non-linear data into “feature” spaces that are essentially linear, where the optimization process can be duplicated as in the linear case. Some of the most commonly used kernel functions are Linear, Polynomial and Radial Basis Functions (RBF includes Gaussian).

The application of SVM to time-series forecasting, called support vector regression (SVR), has also shown many breakthroughs (Sapankevych and Sankar, 2009) However, there are few SVR results on time-series analysis for ITS. The first work that employed SVR in traffic domain was Ding et al. (2002) in which the traffic flow was predicted based on the past flow. Wu et al. (2004) setup a more elaborate experiment for examining the feasibility of applying SVR to travel-time prediction. They experimented with three kernel functions before finding a linear function that had the best performance and proposing a set of SVR parameters that can predict travel times very well. They also did a comparison studies with two other baseline predictors. Bin et al. (2006) used the travel time of preceding/current buses on links to estimate the traffic conditions of links and developed bus arrival time prediction models, which consist of the SVM, segment, current segment travel time, and latest next segment travel time. The research used the RBF kernel function.

A new online SVR method was proposed in the work of Castro-Neto et al. (2009) for short-term horizon predictions. In this work, two scenarios for prediction were considered for short-term traffic flow predictions – normal or non-incident conditions and atypical conditions such as holidays or incidents and compared the performance to other prediction algorithms. While other methods performed slightly better for normal conditions, online SVR outperformed them under atypical conditions which is more relevant for real-world applications. Jeong et al. (2013) proposed a novel approach which is a variation of the online SVR with weights for short-term traffic flow predictions. In the method, the most recent data are given the highest weight, which makes the model able to capture the unusual recent change relatively quicker and consider it on its next prediction making it perform better for atypical traffic conditions. Lippi et al. (2013) also found that seasonal SVR showed similar performance to SARIMA models for forecasts during the most congested periods.

Least squares support vector machine (LSSVM) is a new SVM which can be used to approximate nonlinear systems with higher accuracy. The LSSVM algorithm is an extension

of the SVM using linear least squares criterion for the loss function instead of inequality constraints. Cong et al. (2016) used the LSSVM with a fruit fly optimization algorithm to forecast traffic flow and showed that it performed better than the single LSSVM, RBF neural network and LSSVM with particle swarm optimization.

Besides the advantages of the SVM model, it also has some limitations. Perhaps the biggest one lies in choice of the appropriate kernel function for the practical problem. Wang and Shi (2013) constructed a new kernel function using a wavelet function to capture the non-stationary characteristics of the short-term traffic speed data. This new wavelet function proved to be effective especially at the traffic state transitional period when traffic changes from one state to another.

2.2.2 Principal Component Analysis

Principal Component Analysis (PCA) is a popular terminology in a wide range of fields and is aimed at dimensionality reduction (Jolliffe, 2002). PCA is a coordinate transformation method that maps or projects the input data onto a new set of axes. These axes are called the principal axes or components. Each principal component has the property that it points in the direction of maximum variation given the variation already accounted for in the preceding components. The main disadvantage of PCA is its sensitivity to outliers.

Tsekeris and Stathopoulos (2006b) first introduced the PCA method in the study of traffic matrices, and they found that traffic matrices can be well approximated by a few principal components that correspond to the largest singular values of the matrices. One important drawback is that if the traffic matrix is corrupted by large volume anomalies, the resulting principal components will be significantly skewed from those in the anomaly-free case. This prevents the traditional methods from accurately decomposing the total traffic into different sub-matrices, and decreases the efficiency of PCA-based applications such as anomaly detectors.

Recently, the robust PCA (Candès et al., 2011) (a special case of matrix completion) has attracted much attention in decomposing matrices into low rank and sparse components matrices. The low rank component can be viewed as a de-noised version of the data. The sparse component is useful in detecting the outliers. Jin et al. (2008) used the PCA outlier detection to isolate abnormal traffic flow patterns. Xing et al. (2015) use RPCA to accurately decompose the observed traffic flow matrix into sub-matrices that correspond to different classes of traffic flow which was then used for traffic flow prediction and anomaly detection. The results show that robust PCA succeeded in isolating most of the large volume anomalies in the residual traffic matrix while the traditional PCA failed. For traffic flow prediction, the robust PCA outperformed PCA and naïve average method.

PCA is also used together with other forecasting methods to improve their efficiency. Zhang and He (2007) used PCA to process the historical data of the forecasted traffic volume and interrelated volumes thus reducing the dimension without losing the main information. The output of PCA is fed into a combined neural network and this approach outperformed the typical back propagation network.

2.3 Neural Networks

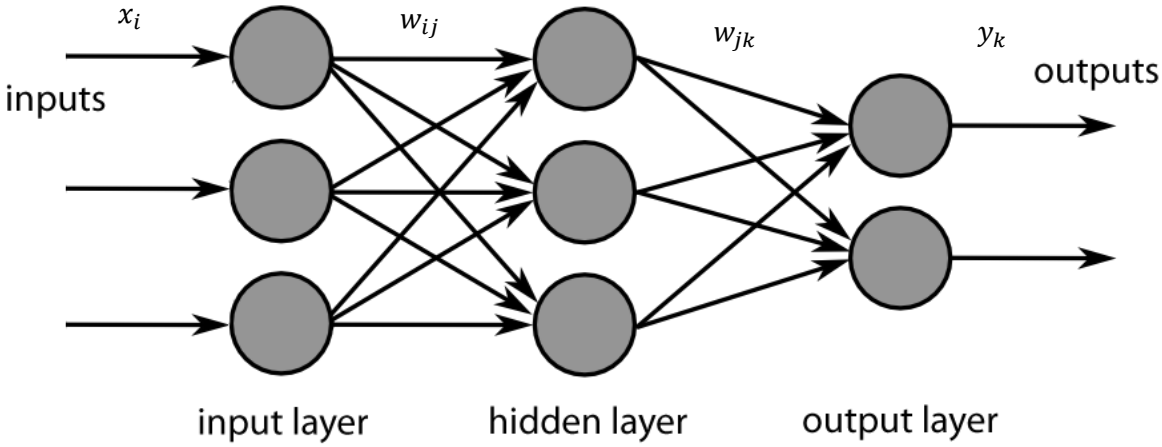
Neural networks is a broad term covering a great many different architectures, or paradigms. The operation of these paradigms can vary enormously. However, all neural networks share some basic common features. They are composed of a number of very simple processing elements, known as neurons or nodes. These elements take data in from a number of sources and compute an output dependent in some way on the values of the inputs, using an internal “transfer function”. The neurons are joined together by weighted connections;

data flows along these connections and is scaled during transmission according to the values of the weights (Dougherty, 1995). And finally, there is a topological structure relating them with each other. In general terms the relationship between the inputs $x_0, x_1 \ldots x_n$, of neuron j and its output y, is given by the following equations which is similar to the general objective function except with multiple weights. This function is typically a non-linear function.

$$I_j = \sum_{i=0}^{n} w_{ij}\, x_i$$
$$y_j = f(I_j)$$

The output of a particular neuron can be the input of another. The neural network's functionality is very much bound up in the values of the connection weights, which can be updated over time, causing the neural network to adapt and possibly "learn". Based on how these weights are updated or tuned, there are different categories of neural networks

Neural networks have been used for various applications in transport from parameter estimation to prediction. Here, we are mainly interested in prediction and most of the previous work use feed-forward networks with back propagation where supervised learning is used to tune the weights and hence we focus on this paradigm. A typical topological structure of a feedforward neural network is shown in the figure.



Neural networks has been exhaustively used for both strategic and short-term forecasts. There is much work that uses variations of backpropagation neural networks for short term predictions during 1990s which saw an explosion of neural network applications (Dougherty, 1995) However, most of them used simulated data until the work of Smith and Demetsky (1994) which predicted real time traffic volume and made a comparison study with ARIMA time series analysis. An important problem of designing a neural net is the large number of input parameters, also known as the "*curse of dimensionality"*. Dougherty and Cobbett (1997) first addressed this in their work where they predicted the traffic flow, speed and occupancy by considering only a subset of the inputs and this subset was found by reducing the network size stepwise by elasticity testing the large neural networks. Jeong and Rilett (2005) also used a neural network to predict bus arrival times and showed that they can be used for real time applications.

It is well known that the optimal number of degrees of freedom in the neural network model depends on the number of training samples, amount of noise in the samples and the complexity of the underlying function being estimated. A new Bayesian approach for neural networks was introduced to solve this with which prior knowledge is used to define the hyper

parameters that determine the model complexity (Lampinen and Vehtari, 2001) This approach was adopted in traffic prediction applications as well, mainly for short-term traffic flow predictions (Zheng et al., 2006; Wang et al., 2014)

Van Lint et al. (2002) introduced state-space neural networks (SSNN), a recurrent neural network for freeway travel time prediction and later improved the predictions for real-time applications by solving for missing or corrupt data (Van Lint et al., 2005). Zhou et al. (2007) developed a hybrid approach, which combined a back-propagation neural network and three traditional travel demand forecasting models (trip generation, trip distribution, and mode choice) to simulate the travellers' inter-city behaviours. In addition, Tsai et al. (2009) developed two neural network based approaches depending on distinctive railway data for short-term railway passenger demand forecasting, in which some temporal features were adopted to improve the forecasting performance. Extracting meaningful patterns or features embedded implicitly in data by pre-processing may enhance the capability of forecasting models, but relatively few studies have focused on developing data pre-processing techniques in neural network based forecasting models.

There is also a body of work that looks into comparative studies between time series predictions and neural networks, mainly focusing on reliable "goodness of fit" measures which is hard for such inherently different classes of methods (Kirby et al., 1997; Karlaftis and Vlahogianni, 2011) What becomes clear from these works is that despite the countless research papers in transportation that utilize NN, researchers often implement them blindly, ignoring some of their shortcomings such as limited inherent explanatory power or their inherent inability to produce a unique solution to a problem (leading many to refer to NN as "black-boxes").

2.3.1 Ensemble Neural Networks

Another extension of neural networks that has gained a lot of attention is committee or ensemble neural networks. Models ensemble is a technique where many prediction models cooperate on the same task. The aggregation of multiple prediction of the same variable may lead to better results and generalization than using a single model prediction. There are several ensembling methods such as Basic Ensemble Method (BEM) where arithmetic means of the individual neural net outputs are considered, Generalized Ensemble Method (GEM) where the outputs are combined in a weighted average and Bagging (Bootstrap Aggregating) where part of the training data is replaced with a random combination of the training data itself. Chen and Chen (2007) employed a bagging ensemble approach on RBF neural networks for short-term traffic flow prediction. This was compared to a single RBF neural network and the ensembling method showed better prediction performance. Hinsbergen et al. (2009) used Bayesian inference theory introduced in Van Hinsbergen and Lint (2008) to combine the trained neural network models in a committee for travel time prediction. The committee predictions led to improved accuracy over the single neural network predictions and also allowed for accurate confidence intervals estimation. Moretti et al. (2015) introduced a hybrid model that combined statistical models with neural ensembles for urban traffic flow forecasting. The neural ensembles used in this method were the BEM and bagging. The bagging ensemble combined with the statistical model outperformed the hybrid BEM.

2.3.2 Self-Organizing Maps

A self-organizing map (SOM) is a type of ANN that is trained using unsupervised learning to produce a low-dimensional, discretized representation of the input space of the training samples, called a map. The two important features of SOM, topological preservation and easy visualization has attracted a great deal of interest in a wide variety of fields ever since it

was introduced in 1981 (Kohonen, 1990). However, this ANN paradigm has so far been quite rare within the transport sector, especially for forecasting applications.

Self-organizing maps have just two layers; a linear input layer which is then wholly connected to the map itself via weighted connections. As a vector of input data is presented, each of the neurons in the map is stimulated. The neuron with the highest activity is awarded the status of 'winner' and its connection weights are increased. When a neuron is declared a 'winner', neurons within a certain radius of influence, the 'neighbourhood function', of that neuron also have their connection weights increased (though by less than the winner). The two popular 2-D map topology configurations are the rectangular and hexagonal lattice of neurons.

Van Der Voort et al. (1996b) proposed a hybrid method known as KARIMA method for short-term traffic flow forecasting. The technique uses a Kohonen hexagonal SOM as an initial classifier with each class having an individually tuned ARIMA model associated with it. The traffic flow was forecasted for half-hour and one hour horizons. The main advantage of this method was that it requires fewer classes, which eased the problem of retraining the model which can aid in tracking long-term changes in traffic flow, making it a reinforced learning process. Hu et al. (2003) introduced a phase reconstruction theory coupled with SOM for seeking the nearest neighbour for clustering to forecast short-term traffic flow. The study considered the chaotic property of traffic flow time series. Chen et al. (2008) investigated the application of SOMs in the representation and prediction of multi-dimensional traffic flow time series. SOMs were applied to cluster the time series and to project each multi-dimensional vector onto a two-dimensional SOM. The SOM visualization facilitated the analysis of the characteristics of the traffic flow evolution in multiple links simultaneously. The kNN algorithm was applied to the clustering results to perform short-term predictions of the traffic flow vectors.

2.3.3 Deep Learning

In 2006, the so-called Deep Learning or Representation learning emerged as a new area of machine learning research to solve the limited modelling and representational power of shallow learning models such as SVM, Multilayer Perceptron(MLP), etc. Deep learning exploits multiple layers of information-processing in a hierarchical architecture for pattern classification and representation learning. It has been applied with success in classification tasks, natural language processing, dimensionality reduction, object detection, motion modelling, and so on. However, deep learning gained popularity in the traffic domain only recently.

Deep neural networks (DNNs) are multilayer networks with many hidden layers, whose weights are fully connected and often initialized or pre-trained using deep belief networks (DBMs). DBM is a pre-training unsupervised step that utilizes large amount of unlabelled training data for extracting structures and regularities in input features, and also provides good initialization weights for DNN.

As a traffic flow process is complicated in nature, deep learning algorithms can represent traffic features without prior knowledge, leading to good performance for traffic flow prediction. Lv et al. (2015) used a deep architecture model with stacked autoencoders as building blocks to learn generic traffic flow features to predict short term traffic flow from loop detector data. The deep network was pretrained using a greedy layerwise unsupervised learning algorithm. The results show that the proposed method was superior to the competing methods. Polson and Sokolov (2016) used deep learning for forecasting traffic flow during special events where the sharp traffic flow regime can occur very suddenly and hardly predictable from historical patterns. Instead of using historical values, they used recent observations of traffic conditions (i.e. within last 40 minutes) for the prediction. Their

hypothesis was that the future traffic conditions were more similar to current ones as compared to those from previous days. The results showed that the vector autoregressive model had good performance on the data for typical days, special event days and poor weather days. However, a deep learning model produced better predictions for non-recurrent events.

2.4 Decision Trees

Tree methods have many varieties and have been successfully used for prediction. However, there is not much work to date in the traffic domain. There is some work on tree based predictors mainly in classification and regression trees (CART) and random forests (RF). The basic idea of the tree method is that the model is fitted using recursive partitioning, whereby the data are successively split along coordinate axes of the predictor variables. The split is done so that, at any node, the response variable is maximally distinguished in the left and right branches. The splitting continues until data are too sparse for each node. Then the tree is "pruned" using cross-validation. Terminal nodes are called "leaves," whereas the initial node is called the "root". CART is used for analysing classification issues for either categorical or continuous dependent variables. When the dependent variable is categorical, CART produces a classification tree and continuous variables produce regression trees.

Nikovski et al. (2005) used a simple regression tree to predict travel times and found that although regression trees can predict travel time at longer prediction horizons (more than one hour) equally well as linear regression, their accuracy at short horizons is quite poor, and much worse than that of linear regression. Xu et al. (2013) applied CART to predict the short-term traffic volume at single locations. The traffic volume series were shaped into state vectors to form the data space which was then clustered into different subsets, on which linear regression was applied. The future traffic state could be predicted through assigning the current traffic state vector to an optimal subset and then calculating it by using the corresponding regression model. The method performed much better than k-NN and Kalman Filters at predicting the traffic volume.

Most of the existing methods work poorly when the data used for forecasting is incomplete, i.e., partially missing or substantially contaminated by noise, while this situation often occurs in practice. Sun and Zhang (2007) used random subspace predictors to use the incomplete data to predict the traffic flow and showed robust performance for both complete and incomplete data. RF also works under the same principle of random space (Ho, 1998). RF is an ensemble method based on feature bagging by which a random subset of the samples of the data are chosen and then the tree regressions are fitted to these samples. The individual regressions trees are averaged after training to provide a more stable prediction. Jomaa et al. (2016) employed CART, RF and other methods to predict traffic speeds which was then used to derive the optimal trigger speed for vehicle activated signs. The results showed that RF was successful at predicting the speed in terms of both performance and efficiency.

2.5 K-Nearest Neighbour

The K-nearest neighbour (K-NN) method has its origin in statistical pattern recognition. The idea behind this method is to locate a set of observations termed as "nearest neighbour" from a well maintained historical database and then the forecast is produced based on this nearest neighbour set. The nearest neighbour set actually reflects the historical traffic data which are similar in a sense to the traffic condition in question. The similarity is generally defined as Euclidean distance. The forecast method applied to the nearest neighbour set can be straight average, weighted average, average adjusted by flow rate, etc. In K-NN method, the size of the nearest neighbour set (span) is important.

Davis and Nihan (1991) applied the K-NN method in forecasting traffic conditions for southbound Interstate-5 at the boundary between Kin and Snohomish counties, Washington

State. A simple first order univariate autoregressive model based on the Box-Jenkins approach was developed as the benchmark. The results showed that the forecasting ability of K-NN model was comparable to, but not definitely better than, that of the autoregressive model. Most of the works are based on traditional KNN mainly focusing on improving the performance and efficiency of traffic flow forecasting (Xiaoyu et al., 2013; Zhang et al., 2013). However, these traditional KNN algorithms are single-step, depending on the limited data regarding a single road, which has two main disadvantages: (i) generating overlapping nearest neighbours when the method is extended to multiple-step forecasting; (ii) sensitive to noisy neighbours. To remedy these limitations, various enhanced KNN algorithm were developed recently.

Zheng and Su (2014) proposed a two-step approach to enhance KNN's performance in forecasting short-term traffic volume using loop detector data. A time constraint window was introduced for selecting the k-nearest neighbours, and then local minima of the distances between the state vectors were ranked to avoid overlapping among candidates. Principal component analysis was used to control extreme values' undesirable impact. Cai et al. (2016) tried to solve the multiple-step forecasting by determining the spatiotemporal correlative road segments from floating car data. The time vectors of the road segments constitute the spatiotemporal state matrices and these vectors are used to describe the traffic states instead of the time series as in ordinary KNNs. The nearest neighbours are selected according to the Gaussian weighted Euclidean distance instead of the use of correlation coefficients as proposed in the work of Zheng and Su (2014). The results showed that the improved KNN performed better for multiple-step forecasting compared to ANN, SVM and other non-parametric methods. However, when the state noise and measurement noise are approximately Gaussian and the complexity is reasonable, the Kalman method is still preferred.

2.6 Scalable Bayesian Inference

In Bayesian based inference methods for traffic prediction, the aim is to evaluate the posterior density function (PDF) of the state vector (vector containing the traffic information). If a recursive framework is applied this PDF is given by (Ristic, Arulampalam and Gordon 2004)

$$p\left(x_k \middle| Z^{k-1}\right) = \int_{\mathbb{R}^{n_x}} p(x_k|x_{k-1}) p(x_{k-1}|Z^{k-1})\, dx_{k-1}$$

where $x_k^T = [x_{1,k}^T, x_{2,k}^T, \dots, x_{n,k}^T]^T$ is the state vector at time t_k and $Z = \{z_1, z_2, \dots, z_k\}$ gives the set of available measurements. Note, the first term in the integral gives the current state conditionally on the previous states, whereas the second gives the previous traffic states conditional on the previous measurements. From Bayes rule

$$p\left(x_k \middle| Z^k\right) = \frac{p(z_k|x_k) p\left(x_k \middle| Z^{k-1}\right)}{p\left(z_k \middle| Z^{k-1}\right)}$$

where, $p\left(z_k \middle| Z^{k-1}\right)$ is a normalising constant. As a result the following approximation can be used

$$p\left(x_k \middle| Z^k\right) \approx p(z_k|x_k) p\left(x_k \middle| Z^{k-1}\right).$$

Unfortunately this process is intractable and as a result approximations are instead used to form the basis of scalable Bayesian inference based methods for traffic prediction. Kalman filters, particle filters (PFs) and Markov Chain Monte Carlo (MCMC) methods are examples of such methods. Each will now be considered in turn.

2.6.1 Kalman Filters and other incremental learning / estimation methods

In his seminal paper (Kalman, 1960), Kalman proposed a filter based on the state space representation of a linear dynamic system, which provides a minimum variance estimator that guarantees an optimal solution to all linear problems that are cast in state space form. State space systems constitute a very general class of mathematical models with applications in many different fields, including control, automation, process technology, and transportation, etc. They consist of the following equations:

$$\mathbf{x}_{k+1} = \mathbf{F}_k \mathbf{x}_k + \mathbf{w}_k \text{ (process equation)}$$

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{v}_k \text{ (observation equation)}$$

The Kalman filter is a very powerful tool in many estimation and prediction problems because all regression and time series problems can be cast in state space form with some creativity. One of its appealing properties is that it requires no memory. For instance, it predicts the future state of a system using the last available state and whatever new information is available. This makes it very suitable to apply in online estimation or calibration settings. Applications in traffic prediction based on both linear Kalman filters (Okutani and Stephanedes, 1984; Ben-Akiva, 1992; Chen and Grant-Muller, 2001; Chien et al., 2003; Stathopoulos and Karlaftis, 2003; Yang, 2005) and nonlinear (Extended) Kalman filters (Chien and Kuchipudi, 2003; Wang and Papageorgiou, 2005; Wang et al., 2007, 2008) can be found in the literature, but in the nonlinear case, an optimal estimator cannot be guaranteed.

Another early contribution was delivered by researchers from Canada who aimed to develop a bus travel time model capable of providing real-time information to travellers and to transit controllers for proactive control strategies (Shalaby and Farhan, 2004). The whole system consisted of two separate algorithms which both relied on the Kalman filter and took advantage of Automatic Vehicle Location(AVL) and Automatic Passenger Counting(APC) data. The demand prediction was achieved through a passenger arrival rate prediction algorithm. The dwell time at a particular stop was predicted by multiplying the predicted arrival rate by the predicted headway and by the passenger boarding time.

Recently, the ensemble Kalman filter (EnKF), which is an extension of the KF to non-linear situations using a Monte Carlo sampling approach, has also been employed to conduct traffic state estimation and prediction (Yuan et al., 2015). The main difference between the EKF and the EnKF is that the EKF is a one-shot procedure that uses only one state vector as an estimation of the true state. The EnKF uses an ensemble of (up to hundreds of) state vectors: the average of the ensemble represents the estimated true state. This distinction has quite some impact on the algorithm: the EKF has to maintain both the average state and covariance matrix separately, whereas the EnKF estimates both quantities via the ensemble.

2.6.2 Particle filters

The Kalman filter discussed above is a special case of Bayesian filtering for linear models and Gaussian noise that gives a closed form solution. For the case of non-linear models EKF and unscented Kalman filters (UKF) can be used. In the first a linearization around an estimate of the mean is used, while the desired PDF is still assumed to be Gaussian (which can lead to divergence). However, this can introduce large errors if the model is highly non-linear. Instead the UKF uses deterministic sampling and sigma points to calculate the mean and covariance. Unfortunately, both the EKF and UKF are unable to track multi modal distributions.

Instead PFs work by representing the PDF of interest as a random set of samples/particle. Then when the number of particles is large enough an exact representation can be achieved (Gordon, Salmond and Smith 1993). This can then be recursively updated when measurements are available (Doucet, Freitas and Gordon 2001) (Gordon, Salmond and Smith 1993). Methods based on PFs have proven to be successful in the areas of traffic state estimation and prediction. This is due to their ability to deal with both uncertainties and non-linearities of any kind (Hoogendoorn et al., 2006; Mihaylova et al., 2007). They have been used with both microscopic, hybrid and macroscopic traffic models for traffic state estimation and prediction in real time scenarios. Gaussian PFs (GPFs), which approximate the state filtering and prediction proababilities with a Gaussian (Kotecha and Djuric, 2003), or Gaussian sum PFs (Kotecha and Djuric, 2001, 2003), a bank of GPFs being used together, can also be applied to traffic related problems (Mihaylova et al., 2012).

Ren, et al. (2010) made short term predictions relating to traffic speed and densities using a PF with a second order traffic model. In (Zhang, et al. 2009) an autoregressive conditional model was used with a PF. The PF was used to estimate/predict the traffic flow speeds and densities using measured vehicle speeds in the update step. The autoregressive model can be used to predict the duration of time intervals when there are no vehicles present. The authors of (Sutarto, Joelianto and Sugian 2013) consider a model based on interconnected junctions with traffic lights on the junctions. A PF was used to estimate the traffic state and model parameter values as explained in section 5. These were then used to predict traffic flow and provide an input to traffic control measures (altering the timing of the traffic lights).

2.6.3 Markov Chain Monte Carlo Methods

In high dimensional problems PFs can be inefficient (Septier, et al. 2009) (Septier and Peters 2015). As a result, MCMC based methods become a desirable alternative. An added advantage of such methods is that they allow sampling from a large, varied group of distributions (Bishop 2006). Two popular MCMC methods, Metropolis-Hastings (MH) algorithm (MH) (Hastings 1970) and Gibbs sampling (Geman and Geman 1984) will be briefly reviewed here. Both are used when it is not possible to sample directly from a desired/target probability distribution.

MH works by taking a sample from a proposal distribution $\boldsymbol{x}_k^* \sim q(\boldsymbol{x}_k|\boldsymbol{x}_k^m)$, where $\boldsymbol{x}_k^*$ is the proposed sample and $\boldsymbol{x}_k^m$ is the current state, resulting in a Markov chain (MC) being formed. By an appropriate selection of $q(\boldsymbol{x}_k|\boldsymbol{x}_k^m)$ this can be guaranteed to be possible. This proposed sample is then accepted or rejected, based on an acceptance probability given by

$$\rho = min\left(1, \frac{p(\boldsymbol{x}_k^*|\boldsymbol{Z}^k)q(\boldsymbol{x}_k^m|\boldsymbol{x}_k^*)}{p(\boldsymbol{x}_k^m|\boldsymbol{Z}^k)q(\boldsymbol{x}_k^*|\boldsymbol{x}_k^m)}\right),$$

where the target distribution is given by $p(\boldsymbol{x}_k^*|\boldsymbol{Z}^k)$ for the proposed sample and $p(\boldsymbol{x}_k^m|\boldsymbol{Z}^k)$ for the current sample. As a result, the new sample will either be given by the proposed sample (i.e. the proposed sample is accepted) or it will be the same as the previous sample (if the proposed sample is rejected). The Metropolis algorithm, (Metropolis, et al. 1953), can be considered as a special case of the MH algorithm, where the proposal distribution is symmetric.

Gibbs sampling is a special case of MH, where the acceptance probability is always 1 (Bishop 2006).When sampling the traffic state from distribution $p(\boldsymbol{x}_k^m|\boldsymbol{Z}^k) = p(\boldsymbol{x}_{1,k}^m, \boldsymbol{x}_{2,k}^m, \dots, \boldsymbol{x}_{n,k}^m|\boldsymbol{Z}^k)$, Gibbs sampling involves drawing each $\boldsymbol{x}_{i,k}^m$ individually, using either the previously found new samples, or the previous samples for the remaining individual states. After a given number of iterations of this process (burn-in period) either the mean of

the states in the remaining iterations, or a maximum likelihood estimate, can be used to give a final estimate of the traffic states.

Short term to midterm predictions were considered in (Mai, Ghosh and Wilson 2011). An ARMA-based model for traffic, which can also account for noise cross-correlations, was considered. As the aim was use the model to predict traffic flow, accurate estimates of the model parameters were required. These were obtained by using both MH algorithm and Gibbs sampling. The proposed methods was successfully applied to multiple junctions in Dublin city centre.

Alternatively in (Xu, et al. 2014) a spatiotemporal Bayesian multivariate adaptive splines model was used for the purposes of short term prediction/forecasts. For such a model the traffic prediction for a given road segment is based on both historical measurements but also the measurements from the upstream and downstream road segments. This means the spatial relationship in traffic data is adequately addressed. In this case a reversible jump MCMC, (Green 1995), is used and in particular the framework of (Denison, Mallick and Smith 1998) is followed. Reversible jump MCMC methods are required when there is a change in the dimensionality of the state space. Here this change refers to the number of basis function required by the Bayesian multivariate adaptive splines model to accurately represent the traffic flow. Loop detector data from Portland was used to verify the effectiveness of predictions made based on this model.

In (Beheshti and Sukthankar 2015) long term traffic flow and parking usage prediction were considered. A MH algorithm was used alongside an agent-based model in order to achieve this. The agent based models were used to obtain the proposal distribution for the MH algorithm. This was done by generating simulated data to initialise the proposal distribution. The MH algorithm was then used to obtain the posterior distribution of the model parameters and the corresponding estimates were used to make long term predictions. Using the combination of the two methods allowed for an improvement in performance compared to using each method separately. This is for the following reasons:

1. The agent based models can be used to accurately represent complex phenomena such as traffic flow in an accurate manner. However, they suffer from poor replicability.
2. On the other hand MCMC methods are relatively simple in comparison. However, they are sensitive to the proposal distribution selected. Therefore, if a more accurate distribution can be obtained via the agent based model methods than the MCMC methods should perform better.

It is also worth noting that unlike using stochastic approximation methods to obtain the proposal distribution, agent based methods do not use the same applicability limiting assumptions.

2.7 Other Methods

New predictive methods developed through either combinations of different existing common approaches or application of some novel approaches from other disciplines are still emerging. A brief overview is then presented in this section.

A hybrid EMD–BPN forecasting approach which combines empirical mode decomposition (EMD) and back-propagation neural networks is developed to predict the short-term passenger flow in metro systems (Wei and Chen, 2012). This approach consists of three stages, with the first one (EMD Stage) decomposing the short-term passenger flow series data into a number of intrinsic mode function (IMF) components. The second stage (Component Identification Stage) identifies the meaningful IMFs as inputs for BPN which serves to conduct forecasting for the final stage.

Based on Support Vector Machine (SVM) regression theory, a hybrid model called Chaos–Wavelet Analysis-Support Vector Machine model (C-WSVM) is proposed to perform short-term traffic speed forecasting (Wang and Shi, 2013). The authors construct a new kernel function using a wavelet function to capture the non-stationary characteristics of the short-term traffic speed data. In addition, Phase Space Reconstruction theory is also used to identify the input space dimension.

Tahmasbi and Hashemi (2014) developed a methodology for the short-term prediction of traffic flow using the Stochastic Differential Equation (SDE). The Hull-White model, a popular choice in modelling financial time processes, was employed in their study. Their proposed method was tested with Tehran's highways and their results show that with this method, a better fit (as compared to GARCH/ARMA methods) to the traffic volume was achieved.

A novel short-term traffic flow prediction approach was proposed by Tan et al. (2016) based on dynamic tensor completion (DTC), in which the traffic data are represented as a dynamic tensor pattern, which was able to capture more information of traffic flow than traditional methods, namely, temporal variabilities, spatial characteristics, and multimode periodicity. A DTC algorithm is designed to use the multimode information to forecast traffic flow with a low-rank constraint.

Nicholson and Swann (1974) developed a tunnel traffic forecasting method based on spectral analysis for the sake of the operation of ventilation system. With a stable daily traffic pattern recognized, the daily traffic volume was decomposed into mutually orthogonal components plus an independent expansion error using the random process discrete expansion. The expansion coefficients were estimated based on the historical daily traffic volume and a partial set of the coefficients were updated on receiving new traffic volume data for the corresponding intervals of the day. Predictions for ensuing time intervals were constructed based on the updated coefficients. The stable daily traffic pattern assumption is fundamental and will limit its ability of adapting to the changing traffic environment. Ever since Nicholson and Swann proposed a sequential least-squares prediction based on spectral analysis for traffic flow prediction in 1970s, this method has not been further developed for traffic prediction purpose. It was concluded by Tchrakian et al. (2012) that what sets spectral-analysis based forecasting apart from most other forecasting methods is its representational power, i.e., how well it represents real traffic flows, albeit at single points in the roadway. The modes that are obtained in the spectral forecasting technique represent trends in traffic and thus go some way in capturing the underlying traffic dynamics. In this sense, spectral-analysis-based forecasting considerably differs from forecasts based on statistical time-series models such as ARIMA or those based on ANNs.

In comparison with the profusion of short-term road traffic prediction methods, the development of bus passenger demand lags far behind, despite its importance in minimizing operation cost and improving bus service quality by properly allocating the constrained resources (Ma et al., 2014). More recently, an Interactive Multiple Model-based Pattern Hybrid (IMMPH) approach was proposed by Ma et al. (2014) to fill this specific research gap. They constructed different demand relevant pattern time series and then developed a weekly, a daily and an hourly model to forecast passenger demand. The sub-model predictions were combined by the occurrence probabilities of pattern models using an amended IMM algorithm. The amended IMM algorithm imported a control input vector to control the structure of different pattern models and used time-dependent TPMs instead of constant ones. The core idea of the IMMPH approach was to dynamically estimate the priori optimized combinations of pattern models using real-time observations and TPMs. In addition, Samaras et al. (2015) presented the passenger demand prediction model of a novel information system called BusGrid which was developed for the improvement of productivity and customer service in public transport bus services. They employed the supervised learning of a regression model to predict people demand for any given bus stop

and route.

2.8 Future Challenges

Due to the big data that will be associated with the SETA project, a key future challenges will be the development of computationally efficient Bayesian algorithms. Such algorithms can take various forms.

For PFs (and the Gaussian and Gaussian sum PF), it has been argued that the parallel form of the filter is an efficient way of solving this problem (Wang et al., 2016). This has been considered for both PFs and Gaussian PFs being applied to traffic problems (Mihaylova et al., 2007, 2012). There are two ways of implementing the PF in this way, both based on splitting the road network into subnetworks. The first method shares particles between all the subnetworks, while the second has a separate set of particles for each subnetwork. An efficient version of a multiple model PF with incident detection has recently been proposed in (Wang, Fan, & Work, 2016).

For the MCMC based methods this computational efficiency can be achieved by using either subsampling or divide and conquer methods illustrated in Figure 3 (De Freitas et al., 2015). Here the subsampling based method relies on uniformly selecting a portion of the measurements to be used when evaluating the likelihood (Bardenet, Doucet, & Holmes, 2015). This method scales well with increases in data size. In fact the more data that there is, the greater the potential computational savings are. Alternatively the divide and conquer method processes all of the data but in a distributed way (Xu, Teh, Zhu, & Zhang, 2014)(Gelman, et al. 2014). For sequential MCMC methods this is implemented with expectation propagation, which is a variational message passing scheme.

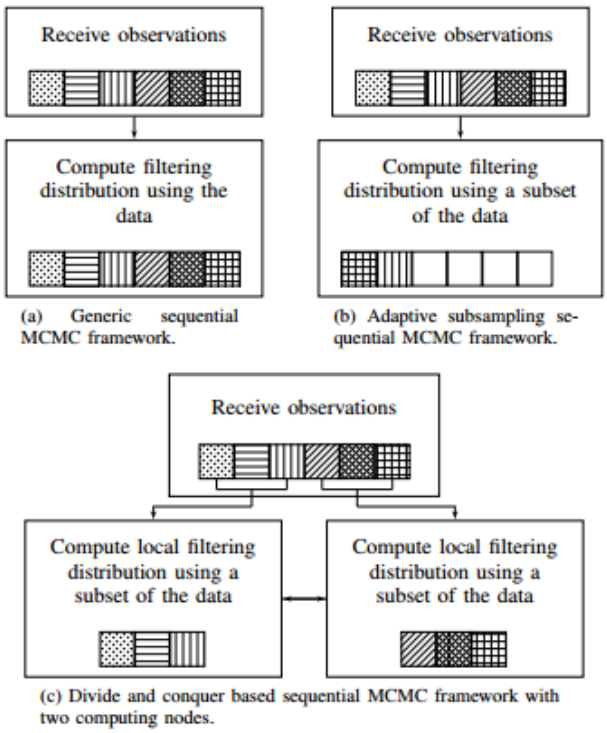


Figure 3 Subsampling and divide and conquer MCMC frameworks, from (De Freitas, Septier and Mihaylova 2015)

The use of the MCMC methods also relies upon the accuracy of the model used. As a result, model errors can lead to less accurate predictions. This means that it may be

beneficial to consider model free methods, which are known as approximate Bayesian computation (ABC) methods (Csilléry et al., 2010) (Fearnhead and Prangle, 2012)(Wilkinson, 2013). Such methods rely on model parameters being drawn from a prior distribution being used to generate data points which can be compared to the observations made. Similar ideas can also be applied to more traditional MCMC methods (Marjoram et al., 2003; Fearnhead and Prangle, 2012). Although these methods have not been directly applied to traffic prediction related problems, there are some obvious similarities between this concept and the methods involving models and MCMC-based methods that have been previously discussed, e.g. (Tossavainen and Work, 2013) (Mai, Ghosh and Wilson 2011) (Xu, Kong, Klette, & Liu, 2014)(Beheshti and Sukthankar 2015), which suggests a possibility for these methods to be exploited.

3 Simulation-based Approaches to Mobility Predictions

The emergence of traffic simulation parallels the rapid development of digital computers, with the first developments in both technologies occurred about 50 years ago in the United States and in Europe and later developments originated in Asia. May (1990) defines simulation as a numerical technique from conducting experiments on a digital computer, which may include stochastic characteristics, be microscopic or macroscopic in nature, and involve mathematical models that describe the behaviour of the system over extended periods of real time. Simulation thus can be seen as an alternative to analytical models constituting a technique that imitates on a computer the operation of a real-world system as it evolves over time.

Of all different types of simulation models that have been used in various subjects, traffic simulation models have been developed by transportation researchers to mimic complex dynamics of traffic systems and the implications of ITS applications. Simulation models for mobility prediction represent the mathematical modelling of transportation system through application of computer software to better support planning, design, management and control of transportation systems. Simulation in transportation is vital because it can study models too complicated for analytical or numerical treatment, can be used for experimental studies, can study detailed relations that might be difficult or impossible to consider analytically and can produce attractive visual demonstrations of present and future scenarios. In the last forty years, many simulation models have been developed where each have a specific purpose and approach towards modelling mobility.

Typically, traffic simulations are classified into four categories based on their level of detail and aggregation, including macroscopic, mesoscopic, microscopic and hybrid. In the context of macroscopic models, traffic is described as a continuum flow based on flow-density functions and explicit modelling of detailed components, such as lanes and vehicles, is not incorporated. Microscopic traffic simulation models, on the contrary, attempt to mimic the real traffic dynamics in a very detailed manner. Individual vehicles are modelled and represented in the simulation with their interactions with other vehicles and geometry. Models concerning driver's behaviour, such as car following, lane changing and gap acceptance behaviours, play a critical role in the performance of simulation results. Mesoscopic models are another available option to many researchers and practitioners lying between microscopic and macroscopic models. Although individual vehicles are represented in mesoscopic simulations, detailed modelling of their second-by-second movement is avoided. Due to computational constraints, the level of detail is always inversely proportional to the network size and complexity. Recently, hybrid models have been developed to enable the simultaneous performing of microscopic and mesoscopic simulation in a way that modelling of the large areas can be realized by zooming out without a coarser level of detailed presentation. Combining an event-based mesoscopic model with a more detailed time-sliced microscopic simulation offers a best-of-both-worlds scenario, blending superior computational efficiency with precise representation of traffic dynamics. In conclusion, different simulation models need to be selected in accordance to the faced problems.

In this chapter, the first section introduces what simulation-based traffic predictions entail and is followed by brief reviews of different levels of simulation models. It is believed that to be able to understand which simulation models are available for mobility prediction and which applications are relevant, a brief overview of existing simulation models along with their characteristics is necessary. A section pertaining to demand estimation and prediction methods (§3.3) is then presented, followed by two other sections about transport assignment models (§3.4) and calibration and validation of simulation models (§3.5). An additional part (§3.4) is included to review some existing parameter estimation techniques. Furthermore, it

is apparent from the emerging mobility data sources and rapid industrial evolution that a shift in modelling approaches is necessary in which modelling packages and suites are considered important to empower mobility prediction. These future perspectives and challenges that the SETA project aims to answer are discussed in the final section.

3.1 Simulation-based Traffic Prediction

The essence of a simulation (model) - based traffic prediction is to feed the model with all necessary components and let every single subject evolve over time under the governance of the model so that the traffic condition that is closest to reality can be reflected. In this sense, the most critical task for simulation-based estimation and prediction boils down to finding the most appropriate components that are going to feed the simulator. These components include demand inputs, boundary conditions, the initial state of the network and model parameters.

Demand
Concerning the definition of traffic demand as an input to the simulator, most of the existing simulators can operate in two alternative modes: aggregate travel demand can be fed into simulation either in the form of origin-destination (OD) matrix or inflows. Route choice models involving dynamic traffic assignment (DTA) and/or dynamic user equilibrium (DUE) became an essential part of the simulation model. In cases where no route choices are used, inflows and turn fractions at intersections need to be specified as well. Details of demand prediction models are presented in §3.3.

Boundary conditions
Boundary conditions also need to be specified and involve, for example, determining reduced capacities at destinations.

Initial states
The initial states of the network concerning a number of traffic parameters, such as density, flow, speed, queues, vehicle mix must be loaded at the beginning of a simulation run.

Model parameters
Carefully tuned model parameters are important for the performance of simulation models in a way that all behaviours relevant to the simulation are essentially determined by the parameters of behavioural models, such as driving and lane changing behaviours in microscopic models and route choice behaviour models. The calibration and validation of simulation models are discussed in 3.5.

In practice, traffic simulation models are implemented as synchronous simulators that – at each simulation step – explore all entities in the model and update the model's state by updating the entities' states. The generic computational framework for the traffic simulation models that uses OD matrices as inputs and deal with route choice models can be described as follows:

1. Initialization:
   1.1. Define a cycle time to update paths in the network
   1.2. Calculate initial shortest paths for each OD pair on the basis of some definition of initial link costs, e.g., free-flow link travel times
2. Repeat until all the demand has been loaded onto the network:
   2.1. Calculate path flow rates according to a route choice model and the proportion of the demand for each OD pair for the selected cycle length.
   2.2. Dynamic network loading: propagate the flows along the paths in accordance with the microscopic flow dynamics:
   At every simulation step, update the position of every vehicle in the model:
   2.2.1. Determine the next move at the current simulation step and the associated

model

2.2.2. Apply the corresponding model: lane-changing, car-following, etc.

2.2.3. Calculate the new position at the end of the simulation step

3. Collect statistics according to a predefined data collection plan
4. Update link costs
5. Update shortest paths with the updated link costs, for use in the next cycle.

Due to the complexity of traffic simulation, in this section we simplify the problem by distinguishing only between (a) predicting supply, i.e. the resulting traffic patterns that occur when assigning / distributing the demand to the network, which implies that under "supply prediction" we also include route choice (assignment) and (b) predicting demand (so either OD patterns or inflows).

3.2 Traffic Flow Modelling

Modelling the dynamics of traffic flows to simulate their temporal propagation through traffic networks is a nice illustration of Marvin Lee Minsky's, an American cognitive scientist in the field of artificial intelligence, statement that a system can be modelled in different ways according to various approaches depending on the modeller's purposes. As shown in Figure 4, traffic flows can be modelled macroscopically from an aggregated point of view based on a hydrodynamic analogy by regarding traffic flows as a particular fluid process whose state is characterized by aggregate macroscopic variables: density, volume, and speed. Alternatively, they can be modelled microscopically, that is, from a fully disaggregated point of view aimed at describing the fluid process from the dynamics of the individual particles (the vehicles) that compose it. Mesoscopic models represent a third intermediate modelling alternative based on a simplification of vehicular dynamics.

Finally, we discuss how the model equations can be discretized for application in computer simulations. The discretization and simulation methods will be applied in the following section discussing applications.

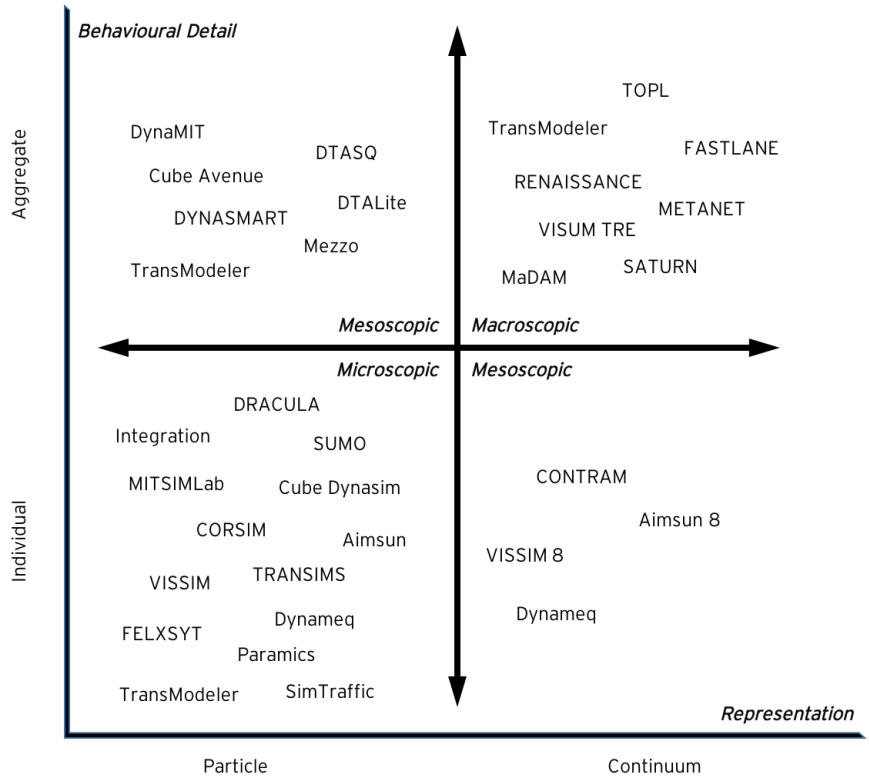


Figure 4 A taxonomy of traffic flow models with corresponding typical examples

3.2.1 Macroscopic Traffic Flow Models

The macroscopic modelling of traffic flows is usually based on the continuum traffic flow theory of which objective is the description of the time-space evolution of the aggregated variables that are capable of characterizing the macroscopic flows, including (average) flow, (average) speed, and (average) density at all network locations. Two major branches, including kinematic wave models and higher-order models, can be found in the family of macroscopic modelling approaches (Kessels, 2013; Wageningen-Kessels et al., 2015). In order to include differences between types of vehicles (e.g., passenger cars and trucks), multiclass versions of both types of macroscopic models are developed as well.

Kinematic wave models

The development of dynamic macroscopic models can be traced back to half a century ago. The kinematic wave model, also known as the Lighthill–Whitham–Richards (LWR) model, was introduced by Lighthill and Whitham (1955) and complemented by Richards (1956) with the introduction of "shock-waves on the highway". Normally the simplest (first order) approach employs three equations: a dynamic equation for the evolution of average density ρ (conservation of vehicles), an equilibrium relation to compute the flow q (or speed u) from density (the fundamental diagram) and the continuity equation $q = \rho$. On a network scale, route choice patterns increase the complexity and number of degrees of freedom substantially. Further model extensions were led by overcoming two main drawbacks of the LWR model; i.e. the transition from free-flow to congestion regime (breakdown) always happens at the same density without capacity drop, and the assumption that vehicles attain new speeds immediately after a change in the density leads to an infinite acceleration and deceleration.

Higher-order models

The second issue has mainly been dealt with higher-order macroscopic models introduced by Payne (1971). Payne's model provided good results under certain traffic conditions but was found to lack accuracy under dense traffic conditions near on-ramp and/or lane drops. A number of extensions (Papageorgiou et al. 1990) contributed to an accuracy improvement of Payne's model. Most of these extensions include a relaxation term that represents the traffic flow tendency to adjust speeds prescribed by the fundamental relation due to changes in free-flow speeds along the roadway. However, also this type of model received much criticism since vehicles do not only react on their leader but also to their follower which results in vehicles driving backward in certain situations. This criticism motivated other researchers to develop anisotropic models (Zhang, 1999; Aw and Rascle, 2000; Klar et al., 2003), including a multiclass higher-order model (Bagnerini and Rascle, 2003) and a generalized higher-order model (Lebacque et al., 2007).

Application of macroscopic models

Since individual vehicles are not modelled in a macroscopic model, the resulting predictions of speeds correspond to an average value of the traffic flow on a specific road section in a network. Therefore, these models cannot predict the lane choice and speed choice of individual vehicles at a random cross-section in the network and can therefore not be used for evaluating measures focusing on lane, speed and headway choice in great detail. In many cases predicting an 'average' speed for a group of drivers on a specific section of road suffices for the purpose for which a macroscopic model is used. Such purposes focus mainly on network and demand changes for future years rather than short-predictions, without considering major changes to the composition of traffic. Examples are infrastructural changes, traffic management measures and road pricing.

Traditionally macroscopic models have been applied most frequently due to their relative availability and lower computation times compared to microscopic models. Initially they were applied as static models, meaning that time dynamics were not considered. Some examples of static macroscopic software packages include Aimsun, Cube Voyager, OmniTRANS, PTV

VISUM, EMME, TransModeler, Synchro. First dynamic macroscopic simulation models were developed in academia, such as METANET (Messner and Papageorgiou, 1990; Papageorgiou et al., 2010). Extensions from this tool have also been available ever since then, including RENAISSANCE (Wang et al., 2006) and AMOC (Kotsialos and Papageorgiou, 2004). More recently, Fastlane was developed based on a first-order multi-class traffic flow model that is faster than second-order models (such as METANET). TOPL was developed based on a cell transmission model. Among the leading commercial traffic modelling packages that include dynamic macroscopic simulation are PTV VISUM TRE, TransModeler, SATURN.

While developments are still ongoing in macroscopic modelling, a shift in focus can be observed from improving macroscopic models to combining the advantages of different modelling approaches by deploying hybrid models. This is further discussed in §3.2.4.

3.2.2 Microscopic Traffic Flow Models

A microscopic traffic flow model simulates the motion of each individual vehicle, such that the dynamic variables of the model represent microscopic properties like the position and speed of the vehicle. One can predict individual speeds, acceleration, deceleration, lane-changing and car-following distances at any time and any place in the network. Microscopic traffic flow models represent the oldest family of the traffic flow simulation models that can be classified in three broad groups, car-following models (safe-distance, stimulus-response and psycho-physical distance models), cellular-automata models and lane-change models (lane selection and lane change models). All these models are in continuous state of improvement due to their significant role in describing complex drivers' behaviour (for example, in traffic breakdowns, crash-prone situations, and adverse weather conditions) to improve traffic conditions and to better understand widely-reported puzzling traffic flow phenomena, such as capacity drop, stop-and-go oscillations, and traffic hysteresis.

Car-following models

The car-following theory describes the longitudinal traffic behaviour, i.e. how one vehicle follows a leading vehicle in an uninterrupted flow, using two microscopic variables: distance-headway and distance gap. Various models were formulated to represent how a driver reacts to the changes in the relative positions of the leading vehicle, and can be classified based on key modelling assumptions as safe-distance, stimulus-response and psycho-physical distance models. A comprehensive review on the state-of-art in car-following models is available in (Brackstone and McDonald, 1999) and Saifuzzaman and Zheng (2014).

The pioneering car-following model was a safe-distance model introduced by Pipes (Pipes 1953). Pipes' model theory follows the assumption that the follower driver will adjust the vehicle speed according to a minimum safe distance headway to their leader, that increases linearly with speed. Taking into account that other behavioural factors, for example the target or desired speed of a driver, should also been taken into account, (Gipps 1981) refined Pipes' model by introducing two components: acceleration and deceleration. In the last decade, researchers put forward safe-distance models by introducing driver's reaction delay (Laval and Leclercq 2010; Newell 2002) and hybrid combination with the LWR model (Bourrel and Lesort 2003; Leclercq, Laval, and Chevallier 2007; Leclercq 2007). Comprehensive field experiments led to the formulation of the car-following models in the form of stimulus-response model, where the response is the reaction of a driver to the motion of the vehicle immediately preceding him in the traffic stream (Gerlough and Huber 1975). The response is always to accelerate or decelerate in proportion to the magnitude of the three stimuli: vehicle's current speed, distance to the leader, and relative speed with respect to the leader. The first and simplest model, now famous GHR-model, proposed by Gazis (Gazis, Herman, and Potts 1959) corresponds to the case where the response is

directly proportional to the stimuli, variation in the relative speed with respect to the leader. More recent developments in response-stimuli models include introduction of the driver's optimal velocity (Bando et al. 1995, 1998), intelligent driver (Treiber, Hennecke, and Helbing 2000; Treiber, Kesting, and Helbing 2010), and generalizations (Klenov 2006; Wilson 2008). Psycho-physical distance models form a yet another group of the car-following models that have been introduced by Wiedemann (1974) and The concept of these models is derived from two assumptions: at large headways, the driver of the following vehicle is not influenced by the size of a speed difference; and at small headways, there are combinations of the relative speeds and distance headways for which driver of the following vehicle only reacts if the change is large enough for him to be perceived. This concept is in contrast to the other car-following models, since driving behaviour is influenced by perceptual threshold of the driver to perceive the changes in relative speed or headways. This model was further developed by (Fellendorf 1994; Brackstone & McDonald 1999) while (Leutzbach 1988) shows that the psycho-physical distance model corresponds to a particular case of the general car-following models.

Cellular-automata models

Cellular-automata models are usually categorized in the microscopic model family, as a class separated from the car-following models. In contrast to car-following models, space and sometimes time are discretized in the cellular-automata models. The first use of stochastic cellular automata for traffic flow modelling was introduced by (Nagel and Schreckenberg 1992), referred as the NaSch's model. The road is discretized into cells (usually 7,5m long), and in each time step every single vehicle is advanced zero, one, or more cells, according to a certain algorithm. This model became the basis for several developments and improvements, such as anticipation models that have been proposed in order to take into account the drivers' movement in the next instant of time and the concept adopted in those methods is that every driver considers that its leading vehicle, the one immediately ahead of it, will move with the same velocity as in the previous time frame (Lárraga and Alvarez-Icaza 2010; Lárraga, del Río, and Schadschneider 2004). More recent developments include brake-light models (Knospe et al. 2000, 2002; Tian et al. 2014) or combination with the car-following models (Helbing and Schreckenberg 1998) and (Kerner, Klenov, and Wolf 2002).

Lane-changing models

The lane-changing theory describes the lateral traffic behaviour. One of the important effects of lane changing manoeuvres on traffic flow characteristics could be speed and traffic flow oscillations. During heavy traffic conditions, the oscillation appears as a result of lane changing rather than car following (Laval and Daganzo 2006; Mauch and Cassidy 2002). A comprehensive review on the state-of-art in lane changing models is available in (Toledo 2007) and (Moridpour, Sarvi, and Rose 2010). Lane changing models can be classified according to assumed two-step decision processes, lane selection and lane change execution.

The lane selection process is influenced by the underlying nature of motivation for lane change. Among the earliest lane changing models, the most popular one was that developed by (Gipps 1986) and followed by (Halati, Lieu, and Walker 1997; Hidas and Behbahanizadeh 1999). The basic structure of these models treated lane selection as a rule-based process, mandatory or discretionary, where the set of lanes in the driver's choice set were repetitively evaluated based on different considerations or rules prioritized through a deterministic sequence. (Yang and Koutsopoulos 1996) introduced a random utility framework for lane selection to overcome the limitations of the deterministic framework of the previous models. Further model improvements were focused on capturing simultaneously mandatory and discretionary situations and over multiple lanes (Ahmed 1999; Toledo et al. 2003; Toledo, Choudhury, and Ben-Akiva 2005).

The lane change execution process is distinguished from the lane selection process, and modelled using gap acceptance models. Gap acceptance models are typically modelled as a binary choice process, such that the driver compares a representative measure of the available adjacent gap with a critical threshold, also known as critical gaps. (Drew et al. 1967; Herman and Weiss 1961; Miller 1972) were among the first to investigate the application of gap acceptance models by modelling critical gap only with the mean and the variance of distributions. Variation of the critical gap variables has been introduced by (Daganzo 1981) and (Mahmassani and Sheffi 1981). Further gap acceptance model improvements were focused on the improvement of modelling drivers' behaviour at non-mandatory lane-changing situations (Ahmed et al. 1996; Kita 1993; Toledo et al. 2003, 2005), and at merge locations (Ahmed 1999; Choudhury et al. 2007; Hidas 2005).

Application of microscopic models

All microscopic models described so far depend on a number of parameters aimed at mimicking as closely as possible the way in which drivers of follower vehicles adjust their driving to that of leader vehicles, while the increasing number of model parameters could in theory replicate better the dynamics of this process with behavioural components. On the other hand, this makes it harder to find the right values of these parameters. Compared to the meso- and macroscopic approach it has the highest level of detail but on the other hand also the longest calculation time. Due to the relatively large computation times, microscopic models are limited to a reduced network size and forecast horizon. As microscopic models consider individual vehicles, they are dynamic by definition. It must be noted that obtaining an equilibrium state is not always a goal in microscopic modelling studies, which means finding a balance in route choice which can be reproduced time and time again. For instances in which reproducible results are required, it may mean making use of multiple simulations runs, of which the results are averaged out to obtain statistically sound and robust results.

The development of the microscopic traffic simulation models has attracted a lot of attention as presented through the literature review resulting in the existence of around 60 microscopic simulation software packages. These software are primarily research products. However, around 20% of them are commercialized and are continuously in development. Some examples of the commercial packages include Aimsun, Paramics, VISSIM, Dracula, Integration, MITSIMLab, FLEXSYT, CORSIM, HUTSIM, THOREAU and TRAF-NETSIM.

3.2.3 Mesoscopic Traffic Flow Models

Mesoscopic modelling of traffic flow dynamics lies in between microscopic models and macroscopic models by representing the behaviour of the individual vehicles (microscopic) and describing the traffic dynamics as a continuum flow (macroscopic). Mesoscopic models describe vehicle movements in aggregate terms such as probability distributions. A summary of existing mesoscopic simulation models which can be used for short-term traffic prediction is available in Burghout (2005). More recently, Di Gangi et al. (2015) presented a brief state of the art in their paper which aims to discuss the classification of mesoscopic models in terms of flow representation. Mesoscopic models include headway distribution models (Branston 1976; Buckley 1968) and cluster models (Kühne et al. 2002). However, the oldest and most extended and popular branch within this family consists of gas–kinetic traffic flow models.

Gas–kinetic traffic flow models were first introduced in the early 1960s (Prigogine and Andrews 1960). They describe traffic flow in a way similar to how gas is modelled in gas–kinetic models. The movements of vehicles (or molecules in a gas) are not modelled individually. Instead, distributions of density and speeds are used to calculate and lead to expected densities and speeds. A first revival of the branch took place in the mid- and late 1970s with an improved model (Paveri-Fontana 1975) and a continuum gas–kinetic model

(Phillips 1979). A second revival of gas–kinetic traffic flow models started in the mid-1990s. The older models were extended and generalized (Helbing 1997; Hoogendoorn and Bovy 2001) and more continuum models were derived (Helbing et al. 2001; Hoogendoorn and Bovy 2000; Tampère, van Arem, and Hoogendoorn 2003; Treiber, Hennecke, and Helbing 1999).

Application of mesoscopic models

In the academic field, DynaMIT (Ben-Akiva et al., 2001) and DYNASMART (Jayakrishnan et al., 1994; Mahmassani et al., 2005) are the two most developed and tested mesoscopic simulation models with predictive capabilities, both of which were benchmarked against a number of online and offline test cases by the Federal Highway Administration (FHWA). Recently the development of DynaMIT 2.0 has been carried out by the collaboration research alliance of MIT and Singapore. The extension of the original project aims to take advantage of data-driven methods to improve the efficiency of the model-based approach (Milkovits et al., 2010; Lu et al., 2015). For commercial applications of mesoscopic modelling tools, Aimsun and Paramics are examples that have been widely employed in practice.

3.2.4 Hybrid or Combined Traffic Simulation Models

Besides models that consist purely of one specific modelling type, a new type of model has emerged which makes use of different modelling types. These combined models make use of different traffic models under certain conditions or for a certain part of the network, such as macroscopic simulation for high level representation of traffic operation on motorways and major roads, coupled with micro-simulation at (controlled) intersections. While some of these are marketed as mesoscopic models, as they represent something between micro- and macroscopic models, they are actually not, but rather a combined or hybrid model. There are several possible combinations between various main characteristics of traffic models. Probably the most common combination is the use of a macroscopic part to model motorway or main traffic flows together with a microscopic model being applied to model intersections or road sections with increased traffic interaction. Other hybrid combinations also exist which make use of mesoscopic modelling, or even combine continuous flow models with discrete models. This last type applies generalized traffic flow characteristics over long homogenous road sections and time and space segmented simulations where more detail is required. The main advantage of approaches with multiple models is that each model is only used when the conditions apply (fit-for purpose). This allows hybrid models to take advantage of the strengths of various model parts and avoid their disadvantages.

Most hybrid models only exist in an experimental state and are not yet available in the market. Within the package AIMSUN 8 (*Aimsun Dynamic Simulators Users Manual v8*, 2015) travel and traffic demand modelling can be fused with mesoscopic, microscopic and hybrid simulation – all within a single software application. It also covers traffic demand modelling within the same interface.

3.3 Demand Estimation and Prediction

Demand estimation and prediction models are used to predict the number of the vehicle trips or travellers in the network with a specific mode of transport. Concerning the definition of traffic demand as an input to traffic simulator, most of the existing simulators can operate in two alternative modes: traffic demand input is defined in terms of inflows and turning fractions or in terms of Origin-Destination (OD) matrices. It is worth noting that in the literature, the term demand estimation is often referred to as a demand calibration or adjustment problem.

3.3.1 Traffic demand as input flows and proportions

Traffic demand input is defined in terms of inflows at origin links and turning fractions at bifurcation nodes. Bifurcation nodes are nodes with one incoming and multiple outgoing links, where each outgoing link has assigned turn fraction. Vehicles travel stochastically in

the network, leaving the network occasionally according to turning fractions. This is the usual mode in most practical applications of traffic simulation models in the networks without route choice options. Definition of traffic demand as inflows and turn fractions is less memory cost for traffic simulations, since it requires storing of information at the node level but at the cost of losing route information. The number of turn fractions is not dependent on the number of routes, link lengths, but just on the size of the network, which is both computationally as well as memory friendly (Raadsen et al., 2009)

This reduced complexity comes with additional cost that turn fractions must be determined beforehand in some way. The majority of traffic simulation models that use turn fractions as an input data are applied to model traffic state on the corridor networks that are covered with sufficient number of traffic detectors. Traffic counts from these traffic detectors can be used to directly derive turn fractions and inflows. Alternatively, turn fractions at the corridor networks can be directly derived from OD matrices. OD matrix estimation approaches that may be used to estimate inflows and turn fractions for linear corridors include Camus et al. (1997), Cremer and Keller (1987) and Van Der Zijpp and De Romph (1997) and extensions for intersections include Bell (1991) and Zuylen and Branston (1982). All these methods assume that all entries and exits in the network are continuously monitored. However, more OD estimation methods that relax this assumption are presented in the section 3.3.2. Furthermore, turn fractions can be determined by the minimum demand supply principle in macroscopic simulation models, where the available supply is distributed according to the effective capacity of the outgoing links (Van Lint et al., 2008).

It is worth noting that destination specific split fractions are the special case of turn fractions on the networks with route choice options. In this case, the route choice behaviour inside the network is described by the use of splitting fractions which express the portion of drivers deciding at a bifurcation node to use a certain alternative output link toward their destination (Papageorgiou et al. 2010). Splitting fractions can be looked upon as turning fractions by destination (i.e., the ration of the traffic volumes in each output link of a node). Depending on the chosen dynamic traffic assignment (DTA) model the number of splitting fractions which must be kept in memory varies. There is no question that this approach is extremely memory intensive and the larger the network gets and the longer each link gets the more fractions need to be kept in memory. Also, the more routes exist the more fractions need to be stored.

3.3.2 Traffic demand as OD matrices

Traffic demand input is defined in terms of origin-destination (OD) matrices, where each cell represents the number of vehicles travelling from the origin node to the destination node along the available paths between these nodes. Route choice models and therefore dynamic traffic assignment and/or dynamic user equilibrium become an essential part of the traffic simulation models.

Disaggregated demand models

The disaggregate demand modelling approach views travellers' behaviour and activities as the unit of analysis, using the observed link counts to calibrate the parameters of the behavioural model to estimate OD trips. The behavioural models can range from a simple trip-based model to a finer activity-based model.

Trip-based demand models

Trip-based demand models were originally developed in the late 1950s, which used individual trips as the unit of analysis and consisted of four sequential steps: trip generation, trip distribution, mode choice (these three steps are also usually referred as 'travel demand') and assignment (also referred as 'travel supply'). Time-of-day of trips is either not modelled or is modelled in only a limited way in the trip-based approach. Most commonly, time is introduced by applying time-of-day factors to 24-hour hour travel volumes at the end of the traffic assignment step or at the end of the trip generation step. The behavioural inadequacy of the trip-based approach, and the consequent limitations of the approach in evaluating

demand management policies, has led to the emergence of the activity-based approach to demand analysis.

Activity-based demand models

The activity-based demand models view travel as a derived demand from the need to pursue activity distributed in space (Axhausen and Gärling, 1992). The seminal works by Chapin (1971), Cullen and Godson (1975) and Hägerstrand (1970) formed the basis for much of the research on activity analysis. Activity-based travel research has received much attention and seen considerable progress since these early studies (Kitamura, 1988; Pas, 1997; Jovicic, 2001; Bhat, 2003). The activity-based models require time-use survey data for estimation and prediction. A time-use survey entails the collection of data regarding all activities (in- home and out-of-home) pursued by individuals over the course of a day (or multiple days). The activity-based approach does require more careful and extensive preparation of data to construct entire "sequences" of activities and travel. On the other hand, such intensive scrutiny of data helps identifies data inconsistencies which might go unchecked in the trip-based approach.

In general, activity-based models can be categorized in discrete choice models, hazard duration models, structural equation models, and computational process models. Here we briefly discuss the discrete choice models as the latter methods have emerged more recently because of the need to model travel as part of a larger activity-travel pattern and involve relatively non-traditional methodologies such as duration analysis and limited-dependent variable models. The multinomial logit model proposed by McFadden (1973) has been the most widely used structure for modelling discrete choices in travel behaviour analysis. This model has been generalized and extended to relax some assumptions of the multinomial logit model (Chu, 1989; Bunch and Kitamura, 1990; Lam and Mahmassani, 1991; McFadden and Train, 2000; Flötteröd and Rohde, 2011). Castiglione et al. (2015) provided an extensive review of the literature on the activity-based demand models.

The advantage of this type of models is that it is possible to include demand management measures and even combine them with traffic management. The disadvantage is the amount of input data needed to feed these models and to calibrate them. Well known simulation software examples of this type of model are the American TRANSIMS (Barrett et al., 2001), the Dutch ALABATROSS model (Arentze and Timmermans, 2000; Arentze, 2005), TASHA (Roorda et al., 2007). More information can be found in a report written by Castiglione et al. (2015).

Aggregated demand models

Aggregated demand approach views trips between every origin to every destination as the unit of analysis, that is, the observed link counts per time interval are used to directly estimate/calibrate the OD trips over time intervals. Dynamic OD demand estimation models have been proposed since the 1980s, and since then a great evolution of these models has taken place. Dynamic OD flow estimation started with the seminal contributions by Okutani and Stephanedes (1984) and Cremer and Keller (1987), which were translated into a practical generic optimization problem by Cascetta et al. (1993). Since then, dynamic OD demand estimation has received a lot of attention in the literature, resulting in a number of generalizations to solve the following challenges: selection of measurements, non-linearity, observability, and online formulations. A review of early contributions in this field is reported by Balakrishna et al. (2005), whilst Toledo et al. (2015) surveys more recent research attempts.

Selection of measurements

Selection of traffic measurement as inputs for dynamic OD demand estimation has attracted a large amount of research attention. In addition to link counts, other traffic variables, including speed, density, and travel time have also been extensively used for such purpose (Ashok and Ben-Akiva, 2000; Tavana and Mahmassani, 2001; Balakrishna and Koutsopoulos, 2008). New technologies for probe vehicle re-identification and tracking

represent additional data source of growing importance. The examples include: AVI data (Dixon and Rilett, 2002; Antoniou et al., 2006; Zhou and Mahmassani, 2007), GSM based trajectories (Sohn and Kim, 2008), Bluetooth data (Barceló et al., 2010). Cipriani et al. (2013) and Djukic et al. (2015) presented an analysis on the impact of different kinds of information to the estimation of dynamic OD flows underlining the importance of using a sample of path travel time measurements, together with measures of speeds and flows on link sections.

Non-linearity

Research efforts have emphasized the importance of accurate modelling of the dependence of link-flow proportions on OD flows in congested networks, and of the consistent non-linear relationship between traffic measurements and OD flows, tackled with different solution frameworks. Chang and Wu (1994) were the first to attempt to formulate a non-linear relationship between OD distributions and link traffic counts under congested conditions on highway networks, using Extended Kalman Filter (EKF) estimators. This work was subsequently extended by Chang and Tao (1996) to more general networks. The influence of non-linearity of the link-route proportion matrix on OD estimation has been studied further by several authors (Flotterod and Bierlaire, 2009; Frederix et al., 2011; Frederix R. and Tampère, 2011). Normally, in terms of formulation, the dynamic OD flows estimation problem for congested networks has typically been specified either as a bi-level optimization problem (Tavana and Mahmassani, 2001; Lindveld, 2003; Zhou et al., 2003; Zhou and Mahmassani, 2006) or bypassing the complex relationship between the OD flows and traffic flow measurements by treating the assignment model as a black box (Balakrishna et al., 2011; Cipriani et al., 2011) or through proper linearization (Nie et al., 2005; Flötteröd and Bierlaire, 2009; Larsson et al., 2010).

Observability

Drawing upon Marzano et al. (2009) and Yang et al. (1991), an important issue in dynamic OD flows estimation refers to the high indeterminateness of the problem. In this respect, some authors tackled this issue using various approaches, for instance, Flötteröd and Bierlaire (2009), Djukic, Flötteröd, et al. (2012) and Djukic, Van Lint, et al. (2012) explored methods to reduce the high dimensionality of OD estimation problem using principal component analysis (PCA). Cascetta et al. (2013) proposed a quasi-dynamic approach, providing a more effective unknowns/equations balance, shown to outperform traditional dynamic OD flow estimators. Cipriani et al. (2014) presented an application of a quasi-dynamic traffic assignment model that approximates the dynamic traffic model by steady-state intervals and applies approximate performance functions to reduce the computation burden of the estimation. Marzano et al. (2015) extended the quasi-dynamic framework to online applications based on the work by Ashok et al. (1993).

Online formulations

A specific research task in dynamic OD estimation research refers to online applications, wherein OD flow estimates for recent time slices together with predictions for future time slices should be provided, recursively and promptly. Online estimation was first studied by Okutani and Stephanedes (1984) and subsequently generalized by many researchers who acknowledged the importance of structural information in OD flows (Cremer and Keller, 1987; Nihan and Davis, 1987, 1989; Ashok et al., 1993; Ashok, 1996; Ashok and Ben-Akiva, 2000). Computational issues in online within-day OD estimation in large networks represent also a main issue, firstly addressed by Bierlaire and Crittin (2004) and more recently by Antoniou et al. (2009).

Commercial software packages, such as OmniTRANS, VISUM and AIMSUN are able to construct demand models and perform related calculations. In the United States the packages CUBE from Citilabs and EMME from INRO are often used.

3.4 Transport Assignment Models

To complete our model building process according to the proposed systems approach, we must be able to formalize the relationships between the capacity and the demand. To model this interaction, the main underlying hypothesis is that travellers travel from origin to destinations in the network along the available routes connecting them, which involves modelling how travellers chose their routes through the network. The modelling hypothesis that supports the main transportation models is based on the concept of user equilibrium, which assumes that travellers try to minimize their individual travel times, that is, travellers chose the routes that they perceive as the shortest under the prevailing traffic conditions. This modelling hypothesis is formulated in terms of Wardrop's first principle (Wardrop, 1952): The journey times on all the routes actually used are equal, and less than those which would be experienced by a single vehicle on any unused route.

Traffic assignment is the process of determining how traffic demand, usually defined in terms of an origin–destination matrix, is loaded onto the network, and it provides the means for computing traffic flows on the network links. Traffic assignment models based on Wardrop's principle are known as user equilibrium models (Sheffi, 1985; Florian and Hearn, 1995). This modelling hypothesis, implemented for traffic demands and average flows not depending on the time of day, has supported the traditional transport-planning models used in practice for strategic planning analysis.

3.4.1 Dynamic Traffic Assignment

The dynamic traffic assignment (DTA) problem can be considered as an extension of the traffic assignment problem described above, capable of describing how traffic flow patterns evolve in time and space on the network (Mahmassani, 2001). Research into DTA started in the late 1970s, and since then received a lot of attention in practice due to its potential to accommodate changes in travel demand and network supply over very short time intervals such as 5 to 15 minutes, and ability to model the spatial and temporal results of their interactions. The two approaches proposed to solve the DTA problem can be classified into two broad classes: mathematical formulations looking for analytical solutions (Wu, 1991; Wu et al., 1998; Xu et al., 1998, 1999) and simulation-based approaches looking for approximate heuristic solutions (Tong and Wong, 2000; Lo and Szeto, 2002; Varia and Dhingra, 2004; Liu et al., 2005). Florian et al. (2001) and Mahut et al. (2002) proposed a systematic framework to solve the DTA problem consisting of two main components:

- Route choice approach determining how the OD flows are assigned onto the available routes at each time step; and
- Dynamic network loading approach to determine how vehicles propagate along the route.

How vehicles propagate along the route is modelled with microscopic, mesoscopic, macroscopic and hybrid models as discussed in Section 3.2. Here we would discuss various route choice approaches implemented in DTA, since the DTA equilibrium framework is independent of the network loading. Route choice approaches in DTA have two components: the generation of a route choice set; and the route choice selection among the alternatives in the choice set.

The generation of a route choice set

Since the number of paths connecting an origin and destination nodes grows exponentially with network size, the choice set of alternative routes cannot include all the feasible routes for most real life networks. In such networks, the route choice set can thus be seen as a subset of the full choice set. Path generation algorithms, deterministic, e.g., Azevedo et al. (1993), Ben-Akiva et al. (1984), Barra et al. (1993), Friedrich et al. (2001), Hoogendoorn-Lanser et al. (2005) and Prato and Bekhor (2006) or stochastic, e.g., Bovy and Fiorenzo-Catalano (2007), Frejinger et al. (2009) and Ramming (2002) can be used to define choice set.

The route choice selection among the alternatives in the choice set

The travellers route choice decisions within DTA are built on the premise that a dynamic user equilibrium (DUE) exists in the network. The definition of DUE is an extension of Wardrop's 1st principle along the temporal dimension formulated by Ran and Boyce (1996). Route choice algorithms can be further grouped into two classes: *preventive* (Papageorgiou 1990), which implicitly assumes that traffic conditions in the network are predictable and travellers are aware of these conditions (e.g., by previous experience), and *reactive*, which assumes that traffic conditions in the network are not predictable (e.g., due to incidents, variability of demand, stochasticity of the traffic system). Friesz et al. (1993) proved that DUE solutions are reached through the implementations of the preventive route choice mechanism, combining experienced travel times with conjectures to predict temporal variations in flow and travel costs. A variety of solution algorithms have been proposed for solving a DTA problem to provide DUE solutions with preventive route choice decision making: from projection algorithms, or methods of alternating directions to various versions of the method of successive averages (MSA) and gradient-based methods.

Other DTA models that consider modelling of reactive route choice decision making are those that model the process from the point of view of probabilistic theory and discrete choice modelling. Discrete choice models consider that the set of available routes for each traveller is a finite choice set of alternatives, each one with a perceived utility by the traveller (Ben-Akiva and Lerman, 1985). Examples of this could be perceived travel time or travel costs. In general, the utility for each alternative path can be considered a random variable consisting of deterministic component, the measured utility and an additive random error (i.e., perception error due to the lack of perfect information). The most used probabilistic models are Logit, modified Logit models (C-Logit and Path-Size Logit), and Probit models (Cross-Nested Logit, Probit, Logit Kernel), and they depend on behavioural parameters that have to be calibrated. These models are used regularly in the recent literature , see e.g., for Logit (Yang, 1999; Yang et al., 2001; Meng et al., 2004) and for Probit (Connors et al., 2007; Uchida et al., 2007; Meng et al., 2012). Probit models can account for partially overlapping routes and routes with significantly different lengths, while simple Logit models can not properly handle overlap, and assume that all route costs are subject to the same level of stochasticity. However, Logit models have closed form solutions for choice probabilities, while Probit equilibria can only be determined using sampling techniques or numerical integration, and are therefore highly computationally intensive. In conclusion, it would be very useful to know how important the differences between these models are likely to be in real-world situations.

In the context of urban traffic networks, Flötteröd (2010) and Gentile and Meschini (2011) developed a tool that calibrates in real-time dynamic assignment models and allows reconstruction of current traffic conditions and prediction of future conditions to support real-time decision making. Peeta and Ziliaskopoulos (2001) provide a conceptual review of various DTA approaches.

3.4.2 Dynamic Transit Assignment

The developments in the field of traffic assignment models point to the potential role that simulation models can play in the context of public transport assignment models. Simulation models provide an appropriate platform to enhance the realization of public transport system modelling. These includes the capabilities to reproduce time-dependent trip generation; the dynamic evolution of network conditions; the interaction between supply and demand; the variation among travellers and their adaptive behaviour; representing operational management strategies, and; emulating the generation and dissemination of passenger information services. Liu et al. (2010) concluded in their review that the developments in public transport route choice and assignment lag behind the counterpart developments in car traffic networks. While the two problems have important similarities and modelling issues, there are also important differences which limit the transferability of developments in

the car traffic network sphere to the public transport network domain. Most important among these differences are the additional service layer which implies limited temporal and spatial availability of the public transport network and results in a more complex definition of path alternatives and consequently the sequence of travel decisions.

Public transport assignment models are conventionally classified into frequency-based and schedule-based models - differing in their network supply representation and its implications on passenger loading procedure. Passengers are assigned to common line corridors in frequency-based models while schedule-based models assign passengers to specific vehicle trips. Since service reliability issues and the respective operations management strategies often apply to individual vehicles, the dynamic public transport model should represent the movement and assign passengers to individual vehicle runs. While schedule-based models represent individual vehicle runs, the evolution of service reliability cannot be fully captured and en-route choices of individual passengers are prohibitive due to demand representation.

Following developments in the sphere of traffic assignment models, there are indeed a few recent efforts in the public transport domain. Although still in its early stages, agent-based simulation models emerged recently as an alternative approach to public transport assignment models. A review of the simulation-based approach to public transport assignment models and description of the features of the main models developed in this domain in recent years is available in Cats et al. (2016). The so-called agent-based approach used in a range of sciences is aimed at modelling complex systems by representing the strategies of individual agents and the dynamics between agents and the environment as well as interactions between agents. Agent-based models represent complex systems using a bottom-up modelling approach where each individual entity is represented as an agent. Simulation models can facilitate the dynamic loading of passengers over a dynamic representation of the public transport system. Wahba and Shalaby (2006) discuss the potential advantages of a multi-agent simulation framework for modelling the public transport assignment problem, in particular in the context of ITS.

Frequency-based assignment models offer a robust modelling framework for long-term strategic planning. However, they are not capable of modelling system dynamics and support real-time applications. The schedule-based approach facilitates the modelling of congestion effects at the individual vehicle trip rather than on a common corridor. However, similarly to frequency-based models, it has a limited capability to capture the dynamics of service reliability and its evolution along the line. Furthermore, the static and aggregate representation of passenger demand prevents the consideration of en-route travel decisions at the individual level. Agent-based simulation models facilitate the dynamic representation of individual passengers and the emergence of dynamic network effects based on numerous inter-dependent local decisions.

Conventional offline public transport assignment models obtain passenger flows by solving equilibrium conditions for a public transport network graph. However, in the context of real-time applications there is no reason to assume that equilibrium conditions will prevail due to the short time frame. In reality, service perturbations are expected to result in non-equilibrium assignment conditions. Hence, the passenger flow module is solely concerned with within-day dynamics rather than day-to-day network evolutions. Notwithstanding, in the absence of complete real-time information concerning passengers' locations and travel plans, a day-to-day dynamic network loading is an essential component of network initialization in order to estimate passenger departure times as well as waiting and on-board flows (e.g. impact of expected reliability and congestion levels on route choice which their assessment requires an iterative day-to-day assignment).

3.5 Calibration and Validation of Simulation Models

In order to use the model as an experimental substitute for the actual system, the reliability of this decision making depends on the ability to produce a simulation model that represents the system's behaviour with a sufficient level of realism (Barceló and Casas, 2004; Dowling, Skabardonis, Halkias, et al., 2004; Jeannotte et al., 2004). Validation is the iterative process to determine whether the simulation model and its parameters are an accurate representation of the system under study. The calibration process has the objective of finding and fine-tuning the simulation model parameters that will produce a valid model. This implies that calibration and validation require obtaining adequate traffic field data in a particular area under consideration, filtering out the unreliable measurements, and completing the missing data. In addition, this process may require the availability of data that are not directly observable, such as transport demand expressed in terms of dynamic OD matrices as an input to the simulation models. For example, macroscopic model calibration will focus on correct macroscopic traffic flow parameters such as fundamental diagram parameters (e.g. density at capacity, jam density). Macro-, meso- and microscopic models should use correctly estimated traffic demand (number of cars measured at on-street counting sites should be as close as possible to the counts in the traffic assignment), otherwise the origin-destination matrix should be updated or route choice model parameters adapted. Microscopic models should include well calibrated driver behaviour parameters (desired headway to predecessor, lane change behavioural parameters, etc.).

The calibration and validation of simulation models is still a major challenge in the use of simulation for practical purposes. This is particularly true in the case of microscopic traffic simulation models that combine high level of uncertainty of the modelled system with a large number of parameters. Consequently, calibration and validation have attracted the attention of many researchers (Bayarri et al., 2002, 2012; Kuwahara et al., 2002; Dowling, Skabardonis, Halkias, et al., 2004; Hollander and Liu, 2008; Ciuffo and Lima Azevedo, 2014) and governmental agencies (Dowling, Skabardonis, and Alexiadis, 2004) in recent years. In accordance with research insights, Dowling, Skabardonis, and Alexiadis (2004) provided guidelines on calibration and validation structured in four stages: error checking, capacity calibration, route choice calibration, and performance validation. Error checking today is a straight forward process as most of the available commercial traffic simulation software provide advanced user-friendly graphic user interfaces to assist analysts in the model building process and reducing the number of errors. Increasing effort has been put into capacity calibration to produce empirically more realistic simulation outcomes, resulting in a variety of approaches that use for example well known trajectory data, NGSIM (Rakha et al., 1996; Wilson, 2001; Choudhury et al., 2007; Punzo and Tripodi, 2007). Route choice calibration has been studied by many researchers (Janson, 1991; Mahut et al., 2004; Liu et al., 2005; Barceló and Casas, 2006; Fox, 2008), as improperly calibrated route choice models yield inappropriate flow distributions on the road network. Other calibration and validation methodologies in the field of traffic flow simulation can be found, amongst others (Ossen et al., 2006; Zhou and Mahmassani, 2006; Ossen and Hoogendoorn, 2008; Punzo et al., 2014; Rocha et al., 2015; Toledo et al., 2015).

4 Prospects

4.1 Data requirements and availability

Data quality and availability are fundamental requirements for any high-dimensional modelling/prediction system. In particular, the big data algorithms developed in this WP will rely on data collected and integrated in WP1, WP2 and WP3.

4.1.1 Data requirements

Big data algorithms typically quantify the amount of data to analyse in terms of volume (the total amount), variety (the number of types of data) and velocity (the speed of data processing) - obviously introducing a trade-off among the tractability of these three dimensions. However, there are general data features that strongly correlate with a higher accuracy of the resulting modelling/prediction system, both pertaining to the specific application domain and general criteria.

Data quality

As stated in the SETA project proposal, “Data quality is a multi-dimensional measurement of the adequacy of a particular datum or data sets. In business, data quality is measured to determine whether or not data can be used as a basis for reliable business intelligence and for making organisational decisions. Much has been said about data reliability, and in general, most business intelligence experts agree that to be deemed reliable, data must measure favourably in a number of dimensions, including, but not limited to accuracy, completeness, consistency, and timeliness”.

Concerning SETA, data quality is implemented through an estimation of the accuracy of all deployed sensors / data gathering devices, and timely fault detection and physical replacement for sensors in order to guarantee continuity. Spatial accuracy is relevant when linking mobile sensors (e.g. probe vehicles) with static ones, to smooth the discrete and potentially abrupt pattern. Spatial coverage is also an important factor, for obvious reasons. The noise-tolerance of the system is strongly linked with the time granularity of the collected data: the superposition of different-scale trends (season/day) has to be taken into account as well.

Furthermore, supervised methods (both model-based or data-driven) require to be validated. Typically, the data set is split in three sets, namely the training, validation and test set. The training set is used as a 'gold standard' / 'ground truth' to learn the correlation between inputs and outputs (often called 'features' and 'labels'). The validation is used to measure the accuracy of the learned correlation on new data - this is done to avoid the problem of overfitting. It is often the case that there is noise (small, random errors) in the data: overfitting the data causes the model to fit the noise, rather than the actual underlying behaviour. In these cases, the model will be highly accurate on the training data but perform poorly on validation data. Several models and (hyper)parameters can be tuned and tested on the validation set, and -to avoid overfitting the validation set during the model selection phase, the final accuracy has to be measured in the third and previously unused set: the test set.

Data storage

The massive amount of data collected, from various sources, requires to be handled with care. From the logical point of view, metadata schemas and data integration techniques are necessary to harmonise different sources. The infrastructural perspective requires taking into account data structures for storage and caching and Database Management System (DBMS) architectures; this decision depends on the kind of the most common operations to be performed on data. For instance, hash tables allow for constant-time access but perform relatively poorly for range searches: as data is stored in pseudo-random locations,

accessing the data in a sorted manner is a very time consuming operation. As a second example, Hadoop is a widely used infrastructure for map-reduce-based algorithms, but its load-and-store loop is not ideal for machine learning, where the same data (e.g. weights of a neural networks) go through iterative updates (whereas Spark introduced an in-memory caching abstraction).

There is also a general trade-off between efficiency and responsiveness: for instance, a naive solution would be to collect traffic data during the day and schedule a batch learning phase at night, when the traffic is at a local minimum; this solution, however, would make the system unaware of the current situation (e.g. major weather event / car crash) until the day after. The designed 'wake-sleep' alternation has a big impact on the way data has to be stored and its persistence in the system.

4.1.2 Data availability

Primary data source: traffic monitoring sensors

For traffic monitoring/prediction systems, the core quantities are traffic speed/volume/flow. These data are collected from a wide range of sources: loop detectors, speed cameras, public transport route plans, traffic lights signalling data, probe vehicles, information concerning planned roadworks as well as school locations, bike sharing stations location and saturation, etc.

Secondary data sources and data integration

Many secondary data sources can be integrated to enrich the data modelling system, spanning from parking lots location/saturation to union strikes. Concerning, for instance, social media, scraping traffic-related geo-localised tweets is relatively easy, but the actual semantical interpretation of each tweet requires neuro-linguistic programming and sentiment analysis techniques, and it's prone to false positives/negatives due to the use of slang, sarcasm, etc. Also, aligning a tweet with the mentioned event is not straightforward ('the traffic was terrible this morning' can be tweeted hours later). Another interesting example is weather data: despite the intuitive link between rain and traffic congestion, there is no consensus (and few experimental results) concerning this correlation, so the usefulness of this integration will have to be validated.

Living Labs

SETA is heavily based on real world requirements and data. For this reason the project will implement use cases in three different but complementary metropolitan areas in Europe, all of which have extensive and intense mobility and transport issues: Birmingham, UK; Turin, Italy; Santander, Spain. The information concerning data availability given below is based on inventories provided by case study sites.

Birmingham

Available data: models (network description, simulation software, GIS shapefile), a wide range of traffic demand data and traffic data observations including automatic vehicle identification/location, a parking space availability system (including capacity and use), not much information concerning the traffic control plan (the only available data is Traffic congestion/UTC), public transport data, location of variable message signs and set of possible pre-fixed messages, data concerning pedestrians and (GPS-enabled) bike sharing, meteorological data, time table for road works, locations of schools and term dates.

Turin

Available data: CCTV camera feeds, city traffic models (network description, simulation software, GIS shapefile), traffic demand data and traffic data observations, a rich parking space availability system including priority lanes and parking tickets, rich traffic control plan data (traffic signal coordination, speed camera data), public transport data, location of variable message signs and set of possible pre-fixed messages, information about the bike sharing system, time table for road works, locations of schools and term dates, and air pollution data.

Santander

Available data: CCTV camera feeds, a complete range of traffic models data, traffic demand data and traffic data observations, a simple parking space availability system, traffic control plan data including traffic signal coordination, public transport data, location of variable message signs and set of possible pre-fixed messages, information about the bike sharing system, meteorological data, time table for road works, and air pollution data.

4.2 Prospects for developing mobility predictors in SETA

The main aim of WP4 in the SETA project is to "explore and develop mobility prediction models (in large urban areas) using large scale, highly-dynamic, high-dimensional, high-volume heterogeneous data." This is a challenging objective that involves application of knowledge and tools from a range of disciplines (transport engineering, data and computer science, operations research and many more). The approach taken in this project is to develop and integrate data-driven and simulation-based methods to capitalize on the benefits offered by each one of them. Data-driven approaches offer efficient, robust and well-founded toolkits that can be applied to generate predictions for large high-volume and high-dimensional datasets. The main advantage of simulation-based predictions is that they consider the underlying processes that determine traffic states, which facilitates its integration in decision support systems, part of the SETA vision.

The project will take advantage of existing simulators, AIMSUN and OpenTrafficSim, which are owned and developed by TSS and TU Delft, respectively. Project partners currently examine the possibilities to integrate some of the functionalities of these two models for the purpose of this project. Simulation model based predictors, which are fed with time-dependent OD demand prediction results, will be constructed to obtain the network-wide traffic states. Data-driven generation of OD matrices may involve also implementing clustering techniques to obtain a manageable and meaningful set of origins and destinations. The influence of incidents on travel time predictions and passenger flows are of special interest and will be examined using macroscopic 3D spatio-temporal networks and partitioning these into homogeneous regions using a data-driven approach. These time-space models can be sliced to allow their usage for short-term predictions.

The development and deployment of prediction schemes will take place in two cycles within the course of the SETA project. In the first round, predictors will be formulated for predicting aggregated traffic demand. We intend to deploy both machine learning techniques to combine historical data and real-time traffic observations as well as utilize knowledge from route choice models to forecast downstream conditions. Given that Santander has a calibrated transport model available in AIMSUN and sufficient data availability, it seems to be a promising first candidate for evaluating and testing the predictors. Based on the information gained from the evaluation concerning the accuracy, robustness and efficiency of the predictors, the second round will extend the predictors into a large-scale application with big data streams, including the outputs of the data fusion performed in WP3 (i.e. counts and flows based on object recognition from traffic cameras).

References

Ahmed, M.S., Cook, A.R., 1979. Analysis of freeway traffic time-series data by using Box-Jenkins techniques. Transportation Research Record 1–9.

Aimsun Dynamic Simulators Users Manual v8, 2015. . TSS-Transport Simulation Systems.

Antoniou, C., Balakrishna, R., Koutsopoulos, H.N., Ben-Akiva, M., 2009. Off-Line and On-Line Calibration of Dynamic Traffic Assignment Systems.

Antoniou, C., Ben-Akiva, M., Koutsopoulos, H.N., 2006. Dynamic traffic demand prediction using conventional and emerging data sources. Intelligent Transport Systems, IEE Proceedings 153, 97–104.

Arentze, T.A., 2005. ALBATROSS: A Learning-Based Transportation Oriented Simulation System. Via-Via 36, 49–51.

Arentze, T.A., Timmermans, H.J.P., 2000. ALBATROSS : A Learning-Based Transportation Oriented Simulation System. European institute of retailing and services studies, Eindhoven, Netherlands.

Ashok, K., 1996. Estimation and Prediction of Time-Dependent Origin-Destination Flows.

Ashok, K., Ben-Akiva, M.E., 2000. Alternative Approaches for Real-Time Estimation and Prediction of Time-Dependent Origin-Destination Flows. TRANSPORTATION SCIENCE 34, 21–36.

Ashok, K., Ben-Akiva, M.E., Massachusetts Institute of, T., 1993. Dynamic origin-destination matrix estimation and prediction for real-time traffic management systems. Transportation and traffic theory.

Aw, A., Rascle, M., 2000. Resurrection of“ second order” models of traffic flow. SIAM journal on applied mathematics 60, 916–938.

Axhausen, K.W., Gärling, T., 1992. Activity-based approaches to travel analysis: conceptual frameworks, models, and research problems. Transport reviews 12, 323–341.

Azevedo, J., Santos Costa, M.E.O., Silvestre Madeira, J.J.E.R., Vieira Martins, E.Q., 1993. An algorithm for the ranking of shortest paths. European Journal of Operational Research 69, 97–106.

Bagnerini, P., Rascle, M., 2003. A multiclass homogenized hyperbolic model of traffic flow. SIAM journal on mathematical analysis 35, 949–973.

Balakrishna, R., Ben-Akiva, M., Koutsopoulos, H., 2011. Off-line Calibration of Dynamic Traffic Assignment: Simultaneous Demand and Supply Estimation. Transportation Research Record.

Balakrishna, R., Koutsopoulos, H., 2008. Incorporating Within-Day Transitions in the Simultaneous Off-line Estimation of Dynamic Origin-Destination Flows without Assignment Matrices. Transportation Research Record.

Balakrishna, R., Koutsopoulos, H.N., Ben-Akiva, M.E., 2005. Calibration and Validation of Dynamic Traffic Assignment Systems, in: Mahmassani, H.S. (Ed.), Proceedings 16th International SymposiumTransportation Traffic Theory: Flow, DynamicsHumanInteraction. Elsevier, New York, NY, pp. 407–426.

Barceló, J., Casas, J., 2004. Methodological Notes on the Calibration and Validation of Microscopic Traffic Simulation Models, in: Proceedings 83rd Annual MeetingTransportationResearchBoard. Washington, D.C.

Barceló, J., Casas, J., 2006. Stochastic Heuristic Dynamic Assignment Based on AIMSUN Microscopic Traffic Simulator. Transportation Research Record: Journal of the Transportation Research Board 1964, 70–80.

Barceló, J., Montero, L., Marquos, L., Carmona, C., 2010. Travel Time Forecasting and Dynamic Origin-Destination Estimation for Freeways Based on Bluetooth Traffic Monitoring. Transportation Research Record: Journal of the Transportation Research Board 2175, 19–27.

Barra, T. de la, Perez, B., Anez, J., 1993. Multidimensional path search and assignment, in: Proceedings seminar D heldPTRCTransport,Highways Planning Summer Annual Meeting. PTRC Education and research services Ltd, Manchester, UK, pp. 307–320.
Barrett, C., Beckman, R., Berkbigler, K., Bisset, K., Bush, B., Campbell, K., Eubank, S., Henson, K., Hurford, J., Kubicek, D., others, 2001. TRANSIMS: Transportation analysis simulation system, Los Alamos National Laboratory Unclassfied Report, Tech. Rep. LA-UR-00-1725.
Bayarri, M.J., Berger, J.O., Higdon, D., Kennedy, M.C., Kottas, A., Paulo, R., Sacks, J., Cafeo, J.A., Cavendish, J.C., Lin, C.H., Tui, J., 2002. A Framework for Validation of Computer Models. National Institute of Statistical Sciences.
Bayarri, M.J., Berger, J.O., Paulo, R., Sacks, J., Cafeo, J.A., Cavendish, J., Lin, C.-H., Tu, J., 2012. A Framework for Validation of Computer Models. http://dx.doi.org/10.1198/004017007000000092.
Bell, M.G.H.H., 1991. The real time estimation of origin-destination flows in the presence of platoon dispersion. Transportation Research Part B: Methodological 25, 115–125.
Ben-Akiva, M., Bergman, M.J., DALY, A.J., Ramaswamy, R., 1984. Modelling Inter Urban Route Choice Behaviour, in: Paperspresented Ninth International SymposiumTransportation Traffic Theory. VNU Science Press, Delft, Netherlands, pp. 299–330.
Ben-Akiva, M., Bierlaire, M., Burton, D., Koutsopoulos, H.N., Mishalani, R., 2001. Network state estimation and prediction for real-time traffic management. Networks and Spatial Economics 1, 293–318.
Ben-Akiva, M.E., 1992. Real-time prediction of traffic congestion, in: VehicleNavigation Information Systems Conference (3rd). Vehicle navigation & information systems: conference recordPapers.
Ben-Akiva, M.E., Lerman, S.R., 1985. Discrete choice analysis : theory and application to travel demand. MIT Press.
Bhat, C.R., 2003. Simulation estimation of mixed discrete choice models using randomized and scrambled Halton sequences. Transportation Research Part B: Methodological 37, 837–855.
Bierlaire, M., Crittin, F., 2004. An Efficient Algorithm for Real-Time Estimation and Prediction of Dynamic OD Tables. OPERATIONS RESEARCH 52, 116–127.
Bin, Y., Zhongzhen, Y., Baozhen, Y., 2006. Bus arrival time prediction using support vector machines. Journal of Intelligent Transportation Systems 10, 151–158.
Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. Journal of econometrics 31, 307–327.
Bovy, P.H.L., Fiorenzo-Catalano, S., 2007. Stochastic route choice set generation: behavioral and probabilistic foundations. Transportmetrica 3, 173–189.
Box, G.E., Jenkins, G.M., 1976. Time series analysis: Forecasting and control, Holden-Day series in time series analysis. Holden-Day.
Brackstone, M., McDonald, M., 1999. Car-following: a historical review. Transportation Research Part F: Traffic Psychology and Behaviour 2, 181–196.
Bunch, D.S., Kitamura, R., 1990. Multinomial probit model estimation revisited: testing estimable model specifications, maximum likelihood algorithms, and probit integral approximations for trinomial models of household car ownership. UCD-ITS-RP-90-1, Institute of Transportation Studies, University of California.
Burghout, W., 2005. Mesoscopic simulation models for short-term prediction. Center for Traffic Research, KTH Royal Institute of Technology.
Cai, P., Wang, Y., Lu, G., Chen, P., Ding, C., Sun, J., 2016. A spatiotemporal correlative k-nearest neighbor model for short-term traffic multistep forecasting. Transportation Research Part C: Emerging Technologies 62, 21–34.
Camus, R., Cantarella, G.E., Inaudi, D., 1997. Real-time estimation and prediction of origin–destination matrices per time slice. International Journal of Forecasting 13, 13–19.
Candès, E.J., Li, X., Ma, Y., Wright, J., 2011. Robust principal component analysis? Journal of the ACM (JACM) 58, 11.

Cascetta, E., Inaudi, D., Marquis, G., 1993. Dynamic estimators of origin-destination matrices using traffic counts. Transportation science 27, 363–373.

Cascetta, E., Papola, A., Marzano, V., Simonelli, F., Vitiello, I., 2013. Quasi-dynamic estimation of o-d flows from traffic counts: Formulation, statistical validation and performance analysis on real data. Transportation Research Part B: Methodological 55, 171–187.

Castiglione, J., Bradley, M., Gliebe, J., 2015. Activity-Based Travel Demand Models: A Primer.

Castro-Neto, M., Jeong, Y.-S., Jeong, M.-K., Han, L.D., 2009. Online-SVR for short-term traffic flow prediction under typical and atypical traffic conditions. Expert systems with applications 36, 6164–6173.

Cats, O., West, J., Eliasson, J., 2016. A dynamic stochastic model for evaluating congestion and crowding effects in transit systems. Transportation Research Part B: Methodological 89, 43–57.

Chang, G.-L., Tao, X., 1996. ESTIMATION OF DYNAMIC OD DISTRIBUTIONS FOR URBAN NETWORKS, in: TRANSPORTATIONANDTRAFFICTHEORYPROCEEDINGSOFTHE13THINTERNATIONALSYMPOSIUMONTRANSPORTATIONANDTRAFFICTHEORY,LYON,FRANCE,2426JULY,1996/EDITEDBYJEANBAPTISTELESORT.

Chang, G.-L., Wu, J., 1994. Recursive estimation of time-varying origin-destination flows from traffic counts in freeway corridors. Transportation Research Part B: Methodological 28, 141–160.

Chapin, F.S., 1971. Free Time Activities and Quality of Urban Life. Journal of the American Institute of Planners 37, 411–417.

Chen, H., Grant-Muller, S., 2001. Use of sequential learning for short-term traffic flow forecasting. Transportation Research Part C: Emerging Technologies 9, 319–336.

Chen, L., Chen, C.P., 2007. Ensemble learning approach for freeway short-term traffic flow prediction, in: System Systems Engineering, 2007. SoSE'07. IEEE International Conference. IEEE, pp. 1–6.

Chen, Y., Zhang, Y., Hu, J., 2008. Multi-dimensional traffic flow time series analysis with self-organizing maps. Tsinghua Science & Technology 13, 220–228.

Chien, S., Liu, X., Ozbay, K., 2003. Predicting travel times for the South Jersey real-time motorist information system. Transportation Research Record: Journal of the Transportation Research Board 32–40.

Chien, S.I.-J., Kuchipudi, C.M., 2003. Dynamic travel time prediction with real-time and historic data. Journal of transportation engineering 129, 608–616.

Choudhury, C., Ben-Akiva, M.E., Rao, A., Lee, G., Toledo, T., 2007. State Dependence in Lane Changing Models, in: Allsop, R.E., Heydecker, B.G. (Eds.), Transportation Traffic Theory 2007. Papers SelectedPresentation ISTTT17. Elsevier, London , England, pp. 711–733.

Chu, C., 1989. A paired combinatorial logit model for travel demand analysis, in: Proceedings Fifth World ConferenceTransportationresearch,Transportpolicy,&technology 2001. Ventura, CA., pp. 295–309.

Cipriani, E., Florian, M., Mahut, M., Nigro, M., 2011. A gradient approximation approach for adjusting temporal origin–destination matrices. Transportation Research Part C: Emerging Technologies 19, 270–282.

Cipriani, E., Gemma, A., Nigro, M., 2013. A bi-level gradient approximation method for dynamic traffic demand estimation: Sensitivity analysis and adaptive approach, in: 16thInternationalIEEEConference Intelligent Transportation Systems (ITSC 2013). IEEE, pp. 2100–2105.

Cipriani, E., Nigro, M., Fusco, G., Colombaroni, C., 2014. Effectiveness of link and path information on simultaneous adjustment of dynamic O-D demand matrix. European Transport Research Review 6, 139–148.

Ciuffo, B., Lima Azevedo, C., 2014. A sensitivity-analysis-based approach for the calibration of traffic simulation models. IEEE Transactions on Intelligent Transportation Systems 15, 1298–1309.
Connors, R.D., Sumalee, A., Watling, D.P., 2007. Sensitivity analysis of the variable demand probit stochastic user equilibrium with multiple user-classes. Transportation Research Part B: Methodological 41, 593–615.
Cortes, C., Vapnik, V., 1995. Support-vector networks. Machine learning 20, 273–297.
Cremer, M., Keller, H., 1987. A new class of dynamic methods for the identification of origin-destination flows. Transportation Research Part B: Methodological 21, 117–132.
Csilléry, K., Blum, M.G., Gaggiotti, O.E., François, O., 2010. Approximate Bayesian computation (ABC) in practice. Trends in ecology & evolution 25, 410–418.
Cullen, I., Godson, V., 1975. Urban networks: The structure of activity patterns. Progress in Planning 4, 1–96.
Danech-Pajouh, M., Aron, M., 1991. ATHENA: a method for short-term inter-urban motorway traffic forecasting. Recherche Transports Sécurité.
Davis, G.A., Nihan, N.L., 1991. Nonparametric regression and short-term freeway traffic forecasting. Journal of Transportation Engineering 117, 178–188.
De Freitas, A., Septier, F., Mihaylova, L., 2015. Sequential Markov Chain Monte Carlo for Bayesian Filtering with Massive Data. arXiv preprint arXiv:1512.02452.
Di Gangi, M., Cantarella, G.E., Di Pace, R., Memoli, S., 2015. Network traffic control based on a mesoscopic dynamic flow model. Transportation Research Part C: Emerging Technologies http://dx.doi.org/10.1016/j.trc.2015.10.002.
Ding, A., Zhao, X., Jiao, L., 2002. Traffic flow time series prediction based on statistics learning theory, in: IntelligentTransportationSystems,2002.Proceedings.TheIEEE5thInternationalConference. IEEE, pp. 727–730.
Dixon, M.P., Rilett, L.R., 2002. Real-Time OD Estimation Using Automatic Vehicle Identification and Traffic Count Data. Computer-Aided Civil and Infrastructure Engineering 17, 7–21.
Djukic, T., Barceló, J., Bullejos, M., Montero, L., Cipriani, E., Lint, H. van, Hoogendoorn, S.P., 2015. Advanced traffic data for dynamic OD demand estimation: The, in: Proceeding 94th Annual MeetingTransportationResearchBoard. Transportation Research Board, Washington, D.C., pp. 1–16.
Djukic, T., Flötteröd, G., Van Lint, H., Hoogendoorn, S., 2012. Efficient real time OD matrix estimation based on Principal Component Analysis, in: IntelligentTransportationSystems(ITSC),201215thInternationalIEEEConference. IEEE, pp. 115–121.
Djukic, T., Van Lint, J., Hoogendoorn, S., 2012. Application of principal component analysis to predict dynamic origin-destination matrices. Transportation Research Record: Journal of the Transportation Research Board 81–89.
Dougherty, M., 1995. A review of neural networks applied to transport. Transportation Research Part C: Emerging Technologies 3, 247–260.
Dougherty, M.S., Cobbett, M.R., 1997. Short-term inter-urban traffic forecasts using neural networks. International journal of forecasting 13, 21–31.
Dowling, R., Skabardonis, A., Alexiadis, V., 2004. Traffic Analysis Toolbox Volume III: Guidelines for Applying Traffic Microsimulation Modeling Software. FHWA, Washington, D.C.
Dowling, R., Skabardonis, A., Halkias, J., McHale, G., Zammit, G., 2004. Guidelines for Calibration of Microsimulation Models: Framework and Applications. Transportation Research Record: Journal of the Transportation Research Board 1876, 1–9.
Fearnhead, P., Prangle, D., 2012. Constructing summary statistics for approximate Bayesian computation: semi-automatic approximate Bayesian computation. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 74, 419–474.

Fei, X., Lu, C.-C., Liu, K., 2011. A bayesian dynamic linear model approach for real-time short-term freeway travel time prediction. Transportation Research Part C: Emerging Technologies 19, 1306–1318.
Florian, M., Hearn, D., 1995. Network equilibrium models and algorithms, in: Handbooks Operations ResearchManagementScience. pp. 485–550.
Florian, M., Mahut, M., Tremblay, N., 2001. A hybrid optimization-mesoscopic simulation dynamic traffic assignment model, in: ITSC2001.2001IEEEIntelligentTransportationSystems.Proceedings(Cat.No.01TH8585). IEEE, pp. 118–121.
Flötteröd, G., 2010. A general methodology and a free software for the demand calibration of DTA models, in: InProceedings 3rd International SymposiumDynamicTrafficAssignment. Takayama, Japan, pp. 1–16.
Flotterod, G., Bierlaire, 2009. Improved estimation of travel demand from traffic counts by a new linearization of the network loading map. Proceedings of the European Transport Conference (ETC), Netherlands.
Flötteröd, G., Bierlaire, M., 2009. Improved estimation of travel demand from traffic counts by a new linearization of the network loading map, in: EuropeanTransportConference.
Flötteröd, G., Rohde, J., 2011. Operational macroscopic modeling of complex urban road intersections. Transportation Research Part B: Methodological 45, 903–922.
Fox, K., 2008. Is micro-simulation a waste of time?, in: PROCEEDINGS EUROPEAN TRANSPORT CONFERENCE 2008. Leeuwenhorst, The Netherlands, p. 8.
Frederix R., V.F., Tampère, C.M.J., 2011. A hierarchical approach for dynamic origin-destination matrix estimation on large-scale congested networks. Proceedings of the IEEE-ITSC 2011 conference Washington.
Frederix, R., Viti, F., Corthout, R., Tampère, C.M.J., 2011. New Gradient Approximation Method for Dynamic Origin-Destination Matrix Estimation on Congested Networks. Transportation Research Record: Journal of the Transportation Research Board 2263, 19–25.
Frejinger, E., Bierlaire, M., Ben-Akiva, M., 2009. Sampling of alternatives for route choice modeling. Transportation Research Part B: Methodological 43, 984–994.
Friedrich, M., Hofsaess, I., Wekeck, S., 2001. Timetable-Based Transit Assignment Using Branch and Bound Techniques. Transportation Research Record: Journal of the Transportation Research Board 1752, 100–107.
Friesz, T.L., Bernstein, D., Smith, T.E., Tobin, R.L., Wie, B.W., 1993. A Variational Inequality Formulation of the Dynamic Network User Equilibrium Problem. Operations Research 41, 179–191.
Gentile, G., Meschini, L., 2011. Using dynamic assignment models for real-time traffic forecast on large urban networks, in: InProceedings 2nd International ConferenceModels TechnologiesIntelligentTransportationSystems. Leuven, Belgium., pp. 1–4.
Guo, J., 2005. Adaptive estimation and prediction of univariate vehicular traffic condition series, Ph.D. dissertation. North Carolina State University, Raleigh, NC.
Guo, J., Huang, W., Williams, B.M., 2014. Adaptive Kalman filter approach for stochastic short-term traffic flow rate prediction and uncertainty quantification. Transportation Research Part C: Emerging Technologies 43, 50–64.
Guo, J., Williams, B., 2010. Real-time short-term traffic speed level forecasting and uncertainty quantification using layered Kalman filters. Transportation Research Record: Journal of the Transportation Research Board 28–37.
Hägerstrand, T., 1970. What about people in Regional Science? Papers of the Regional Science Association 24, 6–21.
Hinsbergen, C.I. van, Van Lint, J., Van Zuylen, H., 2009. Bayesian committee of neural networks to predict travel times with confidence intervals. Transportation Research Part C: Emerging Technologies 17, 498–509.

Ho, T.K., 1998. The random subspace method for constructing decision forests. Pattern Analysis and Machine Intelligence, IEEE Transactions on 20, 832–844.
Hollander, Y., Liu, R., 2008. The principles of calibrating traffic microsimulation models. Transportation 35, 347–362.
Hoogendoorn, S., Ossen, S., Schreuder, M., Gorte, B., 2006. Unscented particle filter for delayed car-following models estimation, in: 2006IEEEIntelligentTransportationSystemsConference. IEEE, pp. 1598–1603.
Hoogendoorn-Lanser, S., Ness, R. van, Bovy, P.H.L., 2005. Path size and overlap in multi-modal transport networks: a new interpretation, in: Mahmassani, H.S. (Ed.), Transportation Traffic Theory. Flow, DynamicsHumanInteraction.16thInternationalSymposium TransportationTrafficTheory. Elsevier, Maryland, US, pp. 63–84.
Hu, J., Zong, C., Song, J., Zhang, Z., Ren, J., 2003. An applicable short-term traffic flow forecasting method based on chaotic theory, in: IntelligentTransportationSystems,2003.Proceedings.2003IEEE. IEEE, pp. 608–613.
Janson, B.N., 1991. Dynamic traffic assignment for urban road networks. Transportation Research Part B: Methodological 25, 143–161.
Jayakrishnan, R., Mahmassani, H.S., Hu, T.-Y., 1994. An evaluation tool for advanced traffic information and management systems in urban networks. Transportation Research Part C: Emerging Technologies 2, 129–147.
Jeannotte, K., Chandra, A., Alexiadis, V., Skabardonis, A., 2004. Traffic Analysis Toolbox Volume II: Decision Support Methodology for Selecting Traffic Analysis Tools. Federal Highway Administration, Washington, DC.
Jeong, R., Rilett, L., 2005. Prediction model of bus arrival time for real-time applications. Transportation Research Record: Journal of the Transportation Research Board 195–204.
Jeong, Y.-S., Byon, Y.-J., Mendonca Castro-Neto, M., Easa, S.M., 2013. Supervised weighting-online learning algorithm for short-term traffic flow prediction. Intelligent Transportation Systems, IEEE Transactions on 14, 1700–1707.
Jin, X., Zhang, Y., Li, L., Hu, J., 2008. Robust PCA-based abnormal traffic flow pattern isolation and loop detector fault detection. Tsinghua Science & Technology 13, 829–835.
Jolliffe, I., 2002. Principal component analysis. Wiley Online Library.
Jomaa, D., Yella, S., Dougherty, M., 2016. Speed prediction for triggering vehicle activated signs.
Jovicic, G., 2001. Activity based travel demand modelling -a literature study. Activity Based Travel Demand Modelling.
Kalman, R.E., 1960. A new approach to linear filtering and prediction problems. Journal of basic Engineering 82, 35–45.
Kamarianakis, Y., Kanas, A., Prastacos, P., 2005. Modeling traffic volatility dynamics in an urban network. Transportation Research Record: Journal of the Transportation Research Board 18–27.
Kamarianakis, Y., Prastacos, P., 2003. Forecasting traffic flow conditions in an urban network: comparison of multivariate and univariate approaches. Transportation Research Record: Journal of the Transportation Research Board 74–84.
Karlaftis, M., Vlahogianni, E., 2011. Statistical methods versus neural networks in transportation research: differences, similarities and some insights. Transportation Research Part C: Emerging Technologies 19, 387–399.
Karlaftis, M.G., Vlahogianni, E.I., 2009. Memory properties and fractional integration in transportation time-series. Transportation Research Part C: Emerging Technologies 17, 444–453.
Kessels, F.L.M., 2013. Multi-class continuum traffic flow models: analysis and simulation methods, Ph.D. dissertation, TRAIL Thesis series. TU Delft, Delft University of Technology.

Kirby, H.R., Watson, S.M., Dougherty, M.S., 1997. Should we use neural networks or statistical models for short-term motorway traffic forecasting? International Journal of Forecasting 13, 43–50.
Kitamura, R., 1988. An evaluation of activity-based travel analysis. Transportation 15, 9–34.
Klar, A., Greenberg, J., Rascle, M., 2003. Congestion on multilane highways. SIAM Journal on Applied Mathematics 63, 818–833.
Kohonen, T., 1990. The self-organizing map. Proceedings of the IEEE 78, 1464–1480.
Kotecha, J.H., Djuric, P.M., 2001. Gaussian sum particle filtering for dynamic state space models, in: Acoustics,Speech, Signal Processing, 2001. Proceedings.(ICASSP'01). 2001 IEEE International Conference. IEEE, pp. 3465–3468.
Kotecha, J.H., Djuric, P.M., 2003. Gaussian particle filtering. IEEE Transactions on signal processing 51, 2592–2601.
Kotsialos, A., Papageorgiou, M., 2004. Efficiency and equity properties of freeway network-wide ramp metering with AMOC. Transportation Research Part C: Emerging Technologies 12, 401–420.
Kuwahara, M., Horiguchi, R., Yoshii, T., 2002. Standard Verification Process for Traffic Flow Simulation Model -Draft Version - Standard verification Model for Traffic Flow Simulation Model Verification Manual (. Kochi Technology University, Kochi, Japan.
Kwon, J., Coifman, B., Bickel, P., 2000. Day-to-day travel-time trends and travel-time prediction from loop-detector data. Transportation Research Record: Journal of the Transportation Research Board 120–129.
Lam, S.H., Mahmassani, H.S., 1991. Multinomial probit model estimation: computational procedures and applications,, in: Methods Understanding Travel Behavior1990s,Proceedings International AssociationTravelBehavior. Association Quebecoise du Transp et des Routes Inc, Quebec, Canada, pp. 228–242.
Lampinen, J., Vehtari, A., 2001. Bayesian approach for neural networks—review and case studies. Neural networks 14, 257–274.
Larsson, T., Lundgren, J.T., Peterson, A., 2010. Allocation of Link Flow Detectors for Origin-Destination Matrix Estimation—A Comparative Study. Computer-Aided Civil and Infrastructure Engineering 25, 116–131.
Lebacque, J.-P., Mammar, S., Salem, H.H., 2007. Generic second order traffic flow modelling, in: Transportation Traffic Theory 2007. Papers SelectedPresentation ISTTT17.
Lee, S., Fambro, D., 1999. Application of subset autoregressive integrated moving average model for short-term freeway traffic volume forecasting. Transportation Research Record: Journal of the Transportation Research Board 179–188.
Lighthill, M.J., Whitham, G.B., 1955. On kinematic waves. II. A theory of traffic flow on long crowded roads, in: Proceedings Royal SocietyLondonAMathematical,Physical Engineering Sciences. The Royal Society, pp. 317–345.
Lindveld, K., 2003. Dynamic O-D matrix estimation: a behavioural approach, TRAIL thesis series. Netherlands TRAIL Research School.
Lippi, M., Bertini, M., Frasconi, P., 2013. Short-term traffic flow forecasting: An experimental comparison of time-series analysis and supervised learning. Intelligent Transportation Systems, IEEE Transactions on 14, 871–882.
Liu, H.X., Ma, W., Ban, J.X., Mirchandani, P., 2005. Dynamic equilibrium assignment with microscopic traffic simulation. Proceedings. 2005 IEEE Intelligent Transportation Systems, 2005.
Liu, Y., Bunker, J., Ferreira, L., 2010. Transit Users' Route-Choice Modelling in Transit Assignment: A Review. Transport Reviews 30, 753–769.
Lo, H.K., Szeto, W.Y., 2002. A cell-based dynamic traffic assignment model: Formulation and properties. Mathematical and Computer Modelling 35, 849–865.
Lu, Y., Pereira, F.C., Seshadri, R., O'Sullivan, A., Antoniou, C., Ben-Akiva, M., 2015. DynaMIT2. 0: Architecture Design and Preliminary Results on Real-Time Data Fusion for Traffic Prediction and Crisis Management, in:

IntelligentTransportationSystems(ITSC),2015IEEE18thInternationalConference. IEEE, pp. 2250–2255.
Lv, Y., Duan, Y., Kang, W., Li, Z., Wang, F.-Y., 2015. Traffic flow prediction with big data: A deep learning approach. Intelligent Transportation Systems, IEEE Transactions on 16, 865–873.
Ma, Z., Xing, J., Mesbah, M., Ferreira, L., 2014. Predicting short-term bus passenger demand using a pattern hybrid approach. Transportation Research Part C: Emerging Technologies 39, 148–163.
Mahmassani, H., Fei, X., Eisenman, S., Zhou, X., Qin, X., 2005. DYNASMART-X evaluation for real-time TMC application: CHART test bed, Maryland Transportation Initiative, University of Maryland, College Park, Maryland. Maryland Transportation Initiative, University of Maryland, College Park, Maryland.
Mahmassani, H.S., 2001. Dynamic network traffic assignment and simulation methodology for advanced system management applications. Networks and spatial economics 1, 267–292.
Mahut, M., Florian, M., Tremblay, N., 2002. Application of a simulation-based dynamic traffic assignment model, in: Proceedings IEEE 5th International ConferenceIntelligentTransportationSystems. IEEE, pp. 439–444.
Mahut, M., Florian, M., Tremblay, N., Campbell, M., Patman, D., McDaniel, Z., 2004. Calibration and Application of a Simulation-Based Dynamic Traffic Assignment Model. Transportation Research Record: Journal of the Transportation Research Board 1876, 101–111.
Marjoram, P., Molitor, J., Plagnol, V., Tavaré, S., 2003. Markov chain Monte Carlo without likelihoods. Proceedings of the National Academy of Sciences 100, 15324–15328.
Marzano, V., Papola, A., Cascetta, E., Simonelli, F., 2015. Towards Online Quasi-dynamic o-d Flow Estimation/Updating, in: 2015IEEE18thInternationalConference Intelligent Transportation Systems. IEEE, pp. 1471–1476.
Marzano, V., Papola, A., Simonelli, F., 2009. Limits and perspectives of effective O-D matrix correction using traffic counts. Transportation Research Part C: Emerging Technologies 17, 120–132.
May, A.D., 1990. Traffic flow fundamentals. Prentice Hall, Englewood Cliffs, New Jersey.
McFadden, D., Train, K., 2000. Mixed MNL models for discrete response. Journal of Applied Econometrics 15, 447–470.
McFadden, D.L., 1973. Conditional Logit Analysis of Qualitative Choice Behavior, in: Zarembka, P. (Ed.), Frontiers Econometrics. Wiley, New York.
Meng, Q., Lee, D.-H., Cheu, R.L., Yang, H., 2004. Logit-Based Stochastic User Equilibrium Problem for Entry-Exit Toll Schemes. Journal of transportation engineering 130, 805–813.
Meng, Q., Liu, Z., Wang, S., 2012. Optimal distance tolls under congestion pricing and continuously distributed value of time. Transportation Research Part E: Logistics and Transportation Review 48, 937–957.
Messner, A., Papageorgiou, M., 1990. METANET: A macroscopic simulation program for motorway networks. Traffic Engineering & Control 31, 466–470.
Mihaylova, L., Boel, R., Hegyi, A., 2007. Freeway traffic estimation within particle filtering framework. Automatica 43, 290–300.
Mihaylova, L., Hegyi, A., Gning, A., Boel, R.K., 2012. Parallelized particle and Gaussian sum particle filters for large-scale freeway traffic systems. IEEE Transactions on Intelligent Transportation Systems 13, 36–48.
Milkovits, M., Huang, E., Antoniou, C., Ben-Akiva, M., Lopes, J.A., 2010. DynaMIT 2.0: The next generation real-time dynamic traffic assignment system, in: Advances System Simulation (SIMUL), 2010 Second International Conference. IEEE, pp. 45–51.
Moretti, F., Pizzuti, S., Panzieri, S., Annunziato, M., 2015. Urban traffic flow forecasting through statistical and neural network bagging ensemble hybrid modeling. Neurocomputing 167, 3–7.

Nicholson, H., Swann, C., 1974. The prediction of traffic flow volumes based on spectral analysis. Transportation Research 8, 533–538.
Nie, Y., Zhang, H.M., Recker, W., 2005. Inferring origin–destination trip matrices with a decoupled GLS path flow estimator. Transportation Research Part B: Methodological 39, 497–518.
Nihan, N.L., Davis, G.A., 1987. Recursive estimation of origin-destination matrices from input/output counts. Transportation Research Part B: Methodological 21, 149–163.
Nihan, N.L., Davis, G.A., 1989. Application of Prediction-Error Minimization and Maximum Likelihood to Estimate Intersection O-D Matrices from Traffic Counts. Transportation Science 23, 77–90.
Nihan, N.L., Holmesland, K.O., 1980. Use of the Box and Jenkins time series technique in traffic forecasting. Transportation 9, 125–143.
Oh, S., Byon, Y.-J., Jang, K., Yeo, H., 2015. Short-term Travel-time Prediction on Highway: A Review of the Data-driven Approach. Transport Reviews 35, 4–32.
Okutani, I., Stephanedes, Y.J., 1984. Dynamic prediction of traffic volume through Kalman filtering theory. Transportation Research Part B: Methodological 18, 1–11.
Ossen, S., Hoogendoorn, S., 2008. Validity of Trajectory-Based Calibration Approach of Car-Following Models in Presence of Measurement Errors. Transportation Research Record: Journal of the Transportation Research Board 2088, 117–125.
Ossen, S., Hoogendoorn, S.P., Gorte, B., 2006. Interdriver Differences in Car-Following: A Vehicle Trajectory-Based Study. Transportation Research Record: Journal of the Transportation Research Board 1965, 121–129.
Papageorgiou, M., Papamichail, I., Messmer, A., Wang, Y., 2010. Traffic simulation with METANET, in: Fundamentals traffic simulation. Springer, pp. 399–430.
Pas, E.I., 1997. Recent advances in activity-based travel demand modeling, in: Institute, T.T. (Ed.), ActivityBasedTravelForecastingConferenceProceedings. Texas Transportation Institute, New Orleans, Louisiana, pp. 79–102.
Payne, H.J., 1971. Models of freeway traffic and control. Mathematical models of public systems.
Peeta, S., Ziliaskopoulos, A.K., 2001. Foundations of dynamic traffic assignment: The past, the present and the future. Networks and Spatial Economics 1, 233–265.
Polson, N., Sokolov, V., 2016. Deep Learning Predictors for Traffic Flows. arXiv preprint arXiv:1604.04527.
Prato, C.G., Bekhor, S., 2006. Applying Branch-and-Bound Technique to Route Choice Set Generation. Transportation Research Record: Journal of the Transportation Research Board 1985, 19–28.
Punzo, V., Montanino, M., Ciuffo, B., 2014. Global sensitivity analysis techniques to simplify the calibration of traffic simulation models. Methodology and application to the IDM car-following model. IET Intelligent Transport Systems 8, 479–489.
Punzo, V., Tripodi, A., 2007. Steady-State Solutions and Multiclass Calibration of Gipps Microscopic Traffic Flow Model. Transportation Research Record: Journal of the Transportation Research Board 1999, 104–114.
Raadsen, M., Schilpzand, M., Mein, E., 2009. Applying inter regional shared routes in detailed multiregional dynamic traffic models, in: EuropeanTransportConferenceProceedings. Noordwijkerhout, Netherlands.
Rakha, H., Hellinga, B., Van Aerde, M., Perez, W., 1996. Systematic Verification, Validation and Calibration of Traffic Simulation Models, in: Proceeding 75th Annual MeetingTransportationResearchBoard. Washington, D.C.
Ramming, M.S., 2002. Network knowledge and route choice.
Ran, B., Boyce, D., 1996. Modeling Dynamic Transportation Networks. Springer Berlin Heidelberg, Berlin, Heidelberg.
Rice, J., Van Zwet, E., 2004. A simple and effective method for predicting travel times on freeways. Intelligent Transportation Systems, IEEE Transactions on 5, 200–207.
Richards, P.I., 1956. Shock waves on the highway. Operations research 4, 42–51.

Rocha, T. Vieira da, Leclercq, L., Montanino, M., Parzani, C., Punzo, V., Ciuffo, B., Villegas, D., 2015. Does traffic-related calibration of car-following models provide accurate estimations of vehicle emissions? Transportation Research Part D: Transport and Environment 34, 267–280.

Roorda, M.J., Miller, E.J., Habib, K.M.N., 2007. Validation of TASHA: A 24-Hour Activity Scheduling Microsimulation Model, in: TransportationResearchBoard86thAnnualMeeting. Washington, pp. 1–18.

Saifuzzaman, M., Zheng, Z., 2014. Incorporating human-factors in car-following models: a review of recent developments and research needs. Transportation research part C: emerging technologies 48, 379–403.

Samaras, P., Fachantidis, A., Tsoumakas, G., Vlahavas, I., 2015. A prediction model of passenger demand using AVL and APC data from a bus fleet, in: Proceedings 19th Panhellenic ConferenceInformatics. ACM, pp. 129–134.

Sapankevych, N.I., Sankar, R., 2009. Time series prediction using support vector machines: a survey. Computational Intelligence Magazine, IEEE 4, 24–38.

Shalaby, A., Farhan, A., 2004. Prediction model of bus arrival and departure times using AVL and APC data. Journal of Public Transportation 7, 3.

Sheffi, Y., 1985. Urban Transportation networks: Equilibrium Analysis with Mathematical Programming Methods. Prentice Hall Inc., New Jersey.

Sheu, J.-B., 2005. A fuzzy clustering approach to real-time demand-responsive bus dispatching control. Fuzzy sets and systems 150, 437–455.

Smith, B.L., Demetsky, M.J., 1994. Short-term traffic flow prediction: neural network approach. Transportation Research Record.

Smith, B.L., Williams, B.M., Oswald, R.K., 2002. Comparison of parametric and nonparametric models for traffic flow forecasting. Transportation Research Part C: Emerging Technologies 10, 303–321.

Sohn, K., Kim, D., 2008. Dynamic Origin Destination Flow Estimation Using Cellular Communication System. Vehicular Technology, IEEE Transactions on 57, 2703–2713.

Stathopoulos, A., Karlaftis, M.G., 2003. A multivariate state space approach for urban traffic flow modeling and prediction. Transportation Research Part C: Emerging Technologies 11, 121–135.

Sun, H., Liu, H., Xiao, H., He, R., Ran, B., 2003. Use of local linear regression model for short-term traffic forecasting. Transportation Research Record: Journal of the Transportation Research Board 143–150.

Sun, S., Zhang, C., 2007. The selective random subspace predictor for traffic flow forecasting. Intelligent Transportation Systems, IEEE Transactions on 8, 367–373.

Tahmasbi, R., Hashemi, S.M., 2014. Modeling and forecasting the urban volume using stochastic differential equations. Intelligent Transportation Systems, IEEE Transactions on 15, 250–259.

Tan, H., Wu, Y., Shen, B., Jin, P.J., Ran, B., 2016. Short-Term Traffic Prediction Based on Dynamic Tensor Completion. Intelligent Transportation Systems, IEEE Transactions on PP, 1–11.

Tavana, H., Mahmassani, H.S., 2001. Estimation of Dynamic Origin-Destination Flows from Sensor Data using Bi-level Optimization Method. Presented at the 80th annual meeting of the Transportation Reasearch Board, Washington DC, USA.

Tchrakian, T.T., Basu, B., O'Mahony, M., 2012. Real-time traffic flow forecasting using spectral analysis. Intelligent Transportation Systems, IEEE Transactions on 13, 519–526.

Toledo, T., Kolechkina, T., Wagner, P., Ciuffo, B., Azevedo, C., Marzano, V., Flötteröd, G., 2015. Network model calibration studies, in: Daamen, W., Buisson, C., Hoogendoorn, S.P. (Eds.), TrafficSimulation Data: Validation methodsApplications. CRC Press, Boca Raton, Florida, pp. 141–162.

Tong, C.O., Wong, S.C., 2000. A predictive dynamic traffic assignment model in congested capacity-constrained road networks. Transportation Research Part B: Methodological 34, 625–644.
Tossavainen, O.-P., Work, D.B., 2013. Markov Chain Monte Carlo based inverse modeling of traffic flows using GPS data. Networks & Heterogeneous Media 8.
Tsai, T.-H., Lee, C.-K., Wei, C.-H., 2009. Neural network based temporal feature models for short-term railway passenger demand forecasting. Expert Systems with Applications 36, 3728–3736.
Tsekeris, T., Stathopoulos, A., 2006a. Real-time traffic volatility forecasting in urban arterial networks. Transportation Research Record: Journal of the Transportation Research Board 146–156.
Tsekeris, T., Stathopoulos, A., 2006b. Measuring variability in urban traffic flow by use of principal component analysis. Journal of Transportation and Statistics 9, 49.
Uchida, K., Sumalee, A., Watling, D., Connors, R., 2007. A Study on Network Design Problems for Multi-modal Networks by Probit-based Stochastic User Equilibrium. Networks and Spatial Economics 7, 213–240.
Van Der Voort, M., Dougherty, M., Watson, S., 1996a. Combining Kohonen maps with ARIMA time series models to forecast traffic flow. Transportation Research Part C: Emerging Technologies 4, 307–318.
Van Der Voort, M., Dougherty, M., Watson, S., 1996b. Combining Kohonen maps with ARIMA time series models to forecast traffic flow. Transportation Research Part C: Emerging Technologies 4, 307–318.
Van Der Zijpp, N.J., De Romph, E., 1997. A dynamic traffic forecasting application on the Amsterdam beltway. International Journal of Forecasting 13, 87–103.
Van Hinsbergen, C., Lint, J. van, 2008. Bayesian combination of travel time prediction models. Transportation Research Record: Journal of the Transportation Research Board 73–80.
Van Lint, J., 2008. Online learning solutions for freeway travel time prediction. Intelligent Transportation Systems, IEEE Transactions on 9, 38–47.
Van Lint, J., Hoogendoorn, S., Van Zuylen, H., 2002. Freeway travel time prediction with state-space neural networks: modeling state-space dynamics with recurrent neural networks. Transportation Research Record: Journal of the Transportation Research Board 30–39.
Van Lint, J., Hoogendoorn, S., Zuylen, H.J. van, 2005. Accurate freeway travel time prediction with state-space neural networks under missing data. Transportation Research Part C: Emerging Technologies 13, 347–369.
Van Lint, J., Van Hinsbergen, C., 2012. Short-term traffic and travel time prediction models. Artificial Intelligence Applications to Critical Transportation Issues 22, 22–41.
Van Lint, J.W.C., Hoogendoorn, S.P., Schreuder, M., 2008. Fastlane: New Multiclass First-Order Traffic Flow Model. Transportation Research Record: Journal of the Transportation Research Board 2088, 177–187.
Varia, H.R., Dhingra, S.L., 2004. Dynamic Optimal Traffic Assignment and Signal Time Optimization Using Genetic Algorithms. Computer-Aided Civil and Infrastructure Engineering 19, 260–273.
Vlahogianni, E.I., Golias, J.C., Karlaftis, M.G., 2004. Short-term traffic forecasting: Overview of objectives and methods. Transport reviews 24, 533–557.
Vlahogianni, E.I., Karlaftis, M.G., Golias, J.C., 2014. Short-term traffic forecasting: Where we are and where we're going. Transportation Research Part C: Emerging Technologies 43, 3–19.
Wageningen-Kessels, F. van, Van Lint, H., Vuik, K., Hoogendoorn, S., 2015. Genealogy of traffic flow models. EURO Journal on Transportation and Logistics 4, 445–473.
Wahba, M., Shalaby, A., 2006. MILATRAS: a microsimulation platform for testing transit-ITS policies and technologies, in: IntelligentTransportationSystemsConference,2006.ITSC'06.IEEE. IEEE, pp. 1495–1500.

Wang, J., Deng, W., Guo, Y., 2014. New Bayesian combination method for short-term traffic flow forecasting. Transportation Research Part C: Emerging Technologies 43, 79–94.
Wang, J., Shi, Q., 2013. Short-term traffic speed forecasting hybrid model based on chaos–wavelet analysis-support vector machine theory. Transportation Research Part C: Emerging Technologies 27, 219–232.
Wang, R., Work, D.B., Sowers, R., 2016. Multiple Model Particle Filter for Traffic Estimation and Incident Detection. IEEE Transactions on Intelligent Transportation Systems 10.1109/TITS.2016.2560769, 1–10.
Wang, Y., Papageorgiou, M., 2005. Real-time freeway traffic state estimation based on extended Kalman filter: a general approach. Transportation Research Part B: Methodological 39, 141–167.
Wang, Y., Papageorgiou, M., Messmer, A., 2006. RENAISSANCE–A unified macroscopic model-based approach to real-time freeway network traffic surveillance. Transportation Research Part C: Emerging Technologies 14, 190–212.
Wang, Y., Papageorgiou, M., Messmer, A., 2007. Real-time freeway traffic state estimation based on extended Kalman filter: A case study. Transportation Science 41, 167–181.
Wang, Y., Papageorgiou, M., Messmer, A., 2008. Real-time freeway traffic state estimation based on extended Kalman filter: Adaptive capabilities and real data testing. Transportation Research Part A: Policy and Practice 42, 1340–1358.
Wardrop, J.G., 1952. Some theoretical aspects of road traffic research.
Wei, Y., Chen, M.-C., 2012. Forecasting the short-term metro passenger flow with empirical mode decomposition and neural networks. Transportation Research Part C: Emerging Technologies 21, 148–162.
Williams, B., 2001. Multivariate vehicular traffic flow prediction: evaluation of ARIMAX modeling. Transportation Research Record: Journal of the Transportation Research Board 194–200.
Williams, B.M., Hoel, L.A., 2003. Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: Theoretical basis and empirical results. Journal of transportation engineering 129, 664–672.
Wilson, R.E., 2001. An analysis of Gipps's car-following model of highway traffic. IMA Journal of Applied Mathematics 66, 509–537.
Wu, C.-H., Ho, J.-M., Lee, D.-T., 2004. Travel-time prediction with support vector regression. Intelligent Transportation Systems, IEEE Transactions on 5, 276–281.
Wu, J.H., 1991. Monotone Variational Inequality Problem Study and its Application to Transportation Network Equilibrium problems (No. 801). CENTRE DE RECHERCHE SUR LES TRANSPORTS PUBLICATION.
Wu, J.H., Chen, Y., Florian, M., 1998. The continuous dynamic network loading problem: a mathematical formulation and solution method. Transportation Research Part B: Methodological 32, 173–187.
Xiaoyu, H., Yisheng, W., Siyu, H., 2013. Short-term traffic flow forecasting based on two-tier k-nearest neighbor algorithm. Procedia-Social and Behavioral Sciences 96, 2529–2536.
Xing, X., Zhou, X., Hong, H., Huang, W., Bian, K., Xie, K., 2015. Traffic Flow Decomposition and Prediction Based on Robust Principal Component Analysis, in: IntelligentTransportationSystems(ITSC),2015IEEE18thInternationalConference. IEEE, pp. 2219–2224.
Xu, Y., Kong, Q.-J., Liu, Y., 2013. Short-term traffic volume prediction using classification and regression trees, in: IntelligentVehiclesSymposium(IV),2013IEEE. IEEE, pp. 493–498.
Xu, Y.W., Wu, J.H., Florian, M., 1998. Transportation networks : recent methodological advances : selected proceedings of the 4th EURO transportation meeting, in: TransportationnetworksRecentmethodologicalAdvancesselectedProceedings 4th EURO Transportation Meeting. Pergamon, pp. 51–66.
Xu, Y.W., Wu, J.H., Florian, M., Marcotte, P., Zhu, D.L., 1999. Advances in the Continuous Dynamic Network Loading Problem. Transportation Science 33, 341–353.

Yang, H., 1999. System Optimum, Stochastic User Equilibrium, and Optimal Link Tolls. Transportation Science 33, 354–360.
Yang, H., Iida, Y., Sasaki, T., 1991. An analysis of the reliability of an origin-destination trip matrix estimated from traffic counts. Transportation Research Part B: Methodological 25, 351–363.
Yang, H., Meng, Q., Bell, M.G.H., 2001. Simultaneous Estimation of the Origin-Destination Matrices and Travel-Cost Coefficient for Congested Networks in a Stochastic User Equilibrium. Transportation Science 35, 107–123.
Yang, J.-S., 2005. Travel time prediction using the GPS test vehicle and Kalman filtering techniques, in: AmericanControlConference,2005.Proceedings 2005. IEEE, pp. 2128–2133.
Yuan, Y., 2013. Lagrangian multi-class traffic state estimation, Ph.D. dissertation, TRAIL Thesis series. TU Delft, Delft University of Technology.
Yuan, Y., Scholten, F., Van Lint, H., 2015. Efficient Traffic State Estimation and Prediction Based on the Ensemble Kalman Filter with a Fast Implementation and Localized Deterministic Scheme, in: IntelligentTransportationSystems(ITSC),2015IEEE18thInternationalConference. IEEE, pp. 477–482.
Zhang, H.M., 1999. A mathematical theory of traffic hysteresis. Transportation Research Part B: Methodological 33, 1–23.
Zhang, L., Liu, Q., Yang, W., Wei, N., Dong, D., 2013. An improved k-nearest neighbor model for short-term traffic flow prediction. Procedia-Social and Behavioral Sciences 96, 653–662.
Zhang, X., Rice, J.A., 2003. Short-term travel time prediction. Transportation Research Part C: Emerging Technologies 11, 187–210.
Zhang, Y., Zhang, Y., Haghani, A., 2014. A hybrid short-term traffic flow forecasting method based on spectral analysis and statistical volatility model. Transportation Research Part C: Emerging Technologies 43, 65–78.
Zheng, W., Lee, D.-H., Shi, Q., 2006. Short-term freeway traffic flow prediction: Bayesian combined neural network approach. Journal of transportation engineering 132, 114–121.
Zheng, Z., Su, D., 2014. Short-term traffic volume forecasting: A k-nearest neighbor approach enhanced by constrained linearly sewing principle component algorithm. Transportation Research Part C: Emerging Technologies 43, 143–157.
Zhou, Q., Lu, H., Xu, W., 2007. New Travel Demand Models with Back-Propagation Network, in: NaturalComputation,2007.ICNC2007.ThirdInternationalConference. IEEE, pp. 311–317.
Zhou, X., Mahmassani, H.S., 2006. Dynamic origin-destination demand estimation using automatic vehicle identification data. Intelligent Transportation Systems, IEEE Transactions on 7, 105–114.
Zhou, X., Mahmassani, H.S., 2007. A structural state space model for real-time traffic origin-destination demand estimation and prediction in a day-to-day learning framework. Transportation Research Part B: Methodological 41, 823–840.
Zhou, X., Qin, X., Mahmassani, H.S., 2003. Dynamic origin-destination demand estimation with multiday link traffic counts for planning applications. Transportation Research Record: Journal of the Transportation Research Board 1831, 30–38.
Zuylen, H.J. van, Branston, D.M., 1982. Consistent link flow estimation from counts. Transportation Research Part B: Methodological 16, 473–476.