REPORT ON THE RELEVANT STATE-OF-THE-ART APPROACHES FOR TRAFFIC PREDICTIONS AND SIMULATIONS

D4.1

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 955273

Deliverable administrative information

Deliverable number	D4.1
Deliverable title	Report on the relevant state-of-the-art approaches for traffic predictions and simulations
Dissemination level	Public
Submission deadline	31/05/2022
Version number	1.0
Authors	Ynte Vanderhoydonc (imec) Mohammadmahdi Rahimiasl (imec) Athina Tympakianaki (Aimsun) Panagiotis Fafoutellis (NTUA) Eleni I. Vlahogianni (NTUA) Juan Sebastian Angarita Zapata (Deusto)
Internal reviewers	Eleni Mantouka (NTUA) Arka Ghosh (Deusto) Antonio Masegosa (Deusto)
Document approval	N/A

Legal Disclaimer

https://twitter.com/TANGENT_H2020

https://www.linkedin.com/company/tangent-project/

https://www.youtube.com/channel/UCjhz4kwEm_sTHj7fE4zXToA

For further information please visit http://www.tangent-h2020.eu/

Executive summary

The main goal of work package (WP) 4 is to develop a framework for real-time traffic monitoring and forecasting under various circumstances (e.g. large/sport event, roadworks, accidents, …) while incorporating novel traffic sensing technologies from smart infrastructure and sensors. It will extend the state-of-the-art regarding traffic forecasting approaches with a focus on modern and future mobility.

Since this WP extends the state-of-the-art, the first task will undertake the state-of-the-art itself, namely a thorough analysis of existing approaches in order to set up the requirements and limitations for the future network flow predictions, both from network demand and supply side. Deliverable 4.1, in the form of this report, explores the state-of-the-art prediction methodologies such as data-driven and simulation-based estimation and prediction approaches, both on supply and demand side. The focus of the report is on deep learning approaches for efficient numerical representations of traffic networks, data availability and granularity, new data sources and graph theory. The analysis is widened with aspects such as transferability and generalizability of the different methods.

The state-of-the-art provides an overview of what has been done in the field of traffic state predictions and simulations and what should be further explored. The results of this exploration will provide the base for the development of the techniques for the remaining tasks of this WP. These tasks include the development of traffic supply and traffic demand predictions, the identification of critical conditions and congestion duration prediction and the development of a framework for real-time traffic monitoring and forecasting.

Key words

Traffic predictions, traffic simulations, deep learning

Table of contents

DELIVERABLE ADMINISTRATIVE INFORMATION 1

EXECUTIVE SUMMARY 3

LIST OF ABBREVIATIONS AND ACRONYMS 6

1 PURPOSE OF THE DELIVERABLE 9

2 INTRODUCTION 10

3 SUPPLY ESTIMATION AND PREDICTION 11

4 DEMAND ESTIMATION AND PREDICTION 22

5 RESEARCH GAPS AND CHALLENGES 28

6 GUIDELINES FOR DEVELOPING ROBUST PREDICTION MODELS 30

7 CONCLUSIONS 35

8 REFERENCES 36

List of figures

Figure 1: Summary key methods traffic state prediction 12

List of tables

Table 1: Challenges and contributions of WP4 of the TANGENT project 29

List of abbreviations and acronyms

Acronym	Meaning
ARIMA	Autoregressive integrated moving average
CA	Cellular Automata
CART	Classification and Regression Trees
CNN	Convolutional Neural Network
CTAGO	Cycle Truncation Accumulated Generating Operation
DCN	Diffusion Convolutional Network
DCRNN	Diffusion Convolutional Recurrent Neural Network
DGM	Discrete Grey Model
DTRP	Dynamic Turn Ratio Prediction Model
DNN	Deep Neural Networks
EKF	Extended Kalman Filter
GCN	Graph Convolution Network
GMAN	Graph Multi-Attention Network
GMM	Gaussian Mixture Model
GNN	Graph Neural Network
GP	Gaussian Process
GPR	Gaussian Process Regression
GPS	Global Positioning System
GRU	Gated Recurrent Unit
KKSW	Kerner–Klenov–Schreckenberg–Wolf
KF	Kalman Filter
KL-divergence	Kullback–Leibler divergence
KNN	k-nearest Neighbors Algorithm
LASSO	Least Absolute Shrinkage and Selection Operator

Acronym	Meaning
LETKF	Local Ensemble Transformed Kalman Filter
LSTM	Long Short Term Memory
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MAP	Maximum a Posteriori
MI	Mutual Information
ML	Machine Learning
MLP	Multi-Layer Perceptron
MSE	Mean Squared Error
NB	Naive Bayes
OD	Origin-Destination
PCA	Principal Component Analysis
PSO	Particle Swarm Optimization
RF	Random Forest
RNN	Recurrent Neural Network
SARIMA	Seasonal ARIMA
SDGM	Seasonal Discrete Grey Model
Seq2Seq	Sequence to Sequence
STARIMA	Space-Time ARIMA
STARMA	Space-Time Autoregressive Moving Average
SVR	Support Vector Regression
SVM	Support Vector Machine
UKF	Unscented Kalman Filter
VAE	Variational autoencoder
VAR	Vector autoregressive

Acronym	Meaning
WP	Work Package

1 Purpose of the deliverable

The main goal of WP4 is the development of a real-time traffic monitor and forecasting framework for traffic management. It will extend the state-of-the-art regarding mobility monitoring and forecasting approaches with a focus on modern and future mobility. The first task is a thorough analysis of existing approaches in order to set up the requirements and limitations for the future network flow predictions, both from network demand and supply side.

1.1 Attainment of the objectives and explanation of deviations

Since WP4 extends the state-of-the-art, this deliverable provides a comprehensive overview of forecasting and simulating traffic on the supply and demand side, including both traditional and state-of-the-art methodologies. The objectives related to this deliverable have been achieved in full.

1.2 Intended audience

This deliverable is mainly intended to the public interested in an overview of the state-of-the-art for traffic predictions and simulations. Mainly academics and students might benefit from this report, since it summarizes the state-of-the-art and the requirements and limitations for traffic state predictions.

1.3 Structure of the deliverable and links with other work packages/deliverables

D4.1 reports on the relevant state-of-the-art approaches for traffic predictions and simulations. Furthermore, it analyses the requirements and limitations for the future network flows predictions and simulations, both from network demand and supply side. This will be used for the development of the techniques in the remaining tasks of this work package. Thus D4.1 is input for the remaining deliverables of WP4: D4.2 and D4.3 assess the different approaches to predict traffic demand and supply, while D4.4 analyses the use of these predictions for the detection and impact analysis of different events. These developments are gathered in the framework for real-time traffic monitoring and forecasting in D4.5.

2 Introduction

As a Research and Innovation Action, TANGENT will address several advances beyond the state-of-the-art. One addressed topic is the field of traffic predictions and simulations.

Traffic prediction has more than 30 years of history, yet the specific field is fragmented in terms of methodologies and tools implemented (Vlahogianni et al., 2014). Traditional approaches for traffic prediction present limitations when capturing the spatial and temporal relationships of traffic patterns (L. Zhao et al., 2020). Both data-driven approaches for the prediction of future traffic supply and demand, and simulation-based techniques to study the evolution of traffic supply and demand, are incorporated in this report. With the ever-growing data collection, deep learning techniques started to improve and surpassed the traditional methods in performance (Angarita-Zapata et al., 2019). These new approaches scale with the amount of available data because of their high representational power.

The two major research tracks related to traffic forecasting with deep learning are: (1) developing efficient representations of multi-model traffic networks (B. Yu et al., 2018) and (2) efficient large-scale forecasts based on historical data on short to medium time (up to 4 hours) horizons, flexible and applied on different transport modes (Bogaerts, 2019). Furthermore, this report focuses on data-driven prediction approaches for the traffic demand which is an essential input for the simulations. Additionally, the prediction of the impact of events (e.g. car accidents, large sport events, …) on the traffic network is important. This report also includes hybrid approaches that combine the strengths of both approaches such as using simulated data to train forecasting algorithms on new traffic modes for which there is no historical data available yet (such as Connected and Autonomous Vehicles).

To address advances beyond the state-of-the-art, it is important to include an analysis of the current state-of-the-art. This is outlined in this report. The fields we tackle in this report concern traffic state supply (section 3) and demand (section 4) prediction and simulation, while focusing on the limitations and requirements. Furthermore, the state-of-the-art of event detection is incorporated since the identification of events and the prediction of e.g. congestion duration is an important objective of WP4. Section 5 lists the research gaps and challenges, while section 6 proposes a 10-step approach for researchers and practitioners in order to develop accurate and robust prediction models.

3 Supply estimation and prediction

This section describes the state-of-the-art of traffic state estimation and prediction, which represents the supply side of transportation forecasting. Forecasting the traffic supply contributes significantly to anticipate and prevent traffic congestion. The supply side can be defined by traffic measures such as the travel speed, travel time, density, and flow of traffic on the transport network. Section 3.1 introduces the main methods used in the state-of-the-art, which are usually classified as parametric and non-parametric data-driven approaches, simulation-based approaches, and hybrid methods. Then, section 3.2 focuses on presenting the scope in which traffic estimation and prediction are typically applied, that is, at a local (single points on roads or corridors) or network level. Lastly, section 3.3 presents a review of the prediction of traffic anomaly detection (e.g., accidents, traffic jams, roadworks, sport events), as this phenomenon has a serious impact on the traffic state.

3.1 Methods for traffic state estimation and prediction

Traffic estimation and prediction have been tackled from different modeling perspectives during the last years. The approaches commonly reported in transportation literature are parametric and non-parametric approaches that include the cases of data-driven methods like statistical learning and machine learning, simulation-based approaches, and hybrid methods. For the latter category of methods, two or more algorithms are combined to find synergies that improve their isolated performance. The key methods that are summarized in this report are shown in Figure 1. Each of these families of modeling approaches are presented with more details as follows.

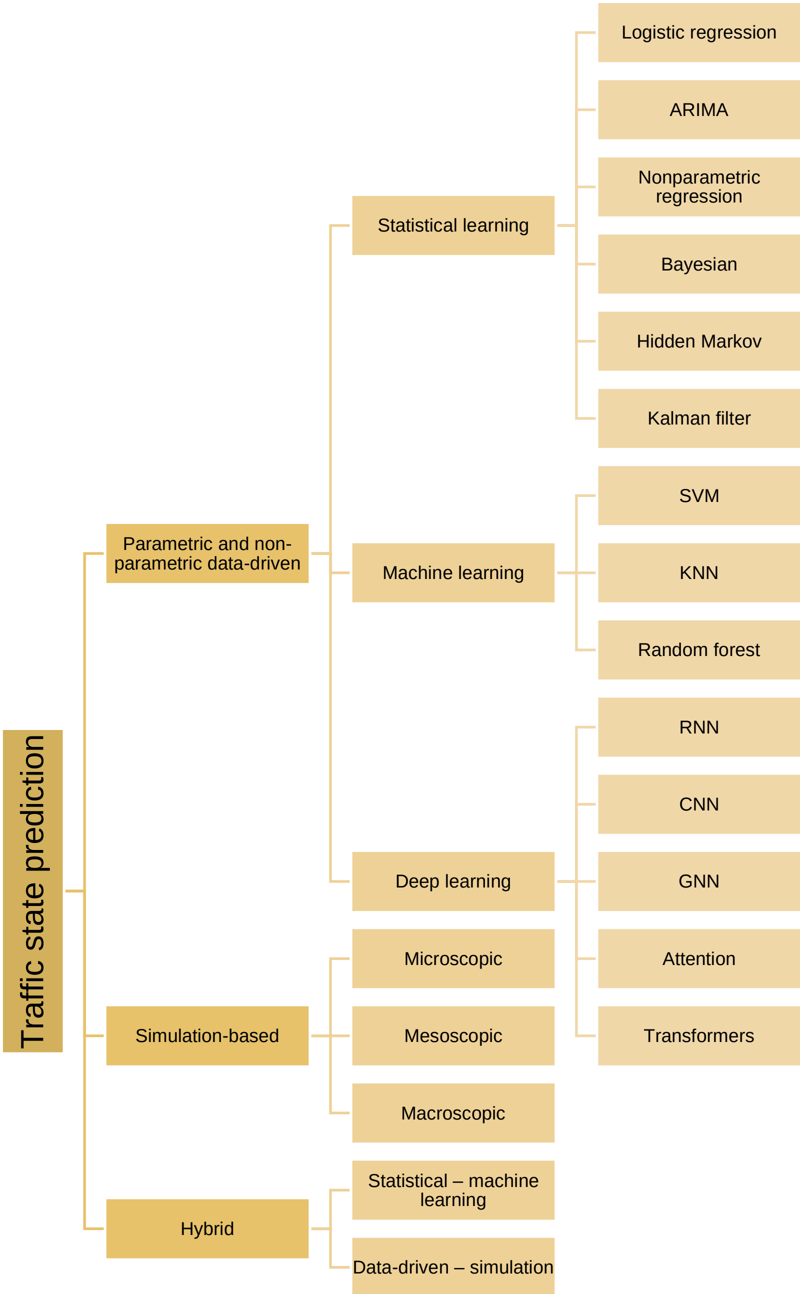

Figure 1: Summary key methods traffic state prediction

3.1.1 Parametric and non-parametric data-driven approaches for mobility estimation and prediction

With the rise of data collection and computational resources, data-driven approaches have been introduced to capture the features present in traffic data and use them to predict the traffic state. Two main categories of data-driven approaches are statistical and Machine Learning (ML) methods. Within the family of statistical approaches, parametric and non-parametric methods are presented.

In general terms, the parametric category of methods assumes the relationship between the explanatory and response variables as known; meanwhile, the non-parametric ones are able to model nonlinear relationships without requiring the mentioned assumptions. Parametric models are described as follows by (Russell & Norvig, 2016). "A learning model that summarizes data with a set of parameters of fixed size (independent of the number of training examples) is called a parametric model. No matter how much data you throw at a parametric model, it won't change its mind about how many parameters it needs." Parametric models include logistic regression, linear discriminant analysis and simple neural networks. In general, these models are simple, fast, and less data-hungry. However, they are constrained to a predefined form, are limited in complexity, and cannot learn the underlying mapping function (Jason Brownlee, 2016).

The autoregressive integrated moving average family of models are more data-hungry parametric statistical approaches for traffic forecasting (ARIMA) (Hamed et al., 1995). It corresponds traditionally with univariate time series prediction using techniques such as autoregressive integrated moving average models. The Seasonal ARIMA (SARIMA) adds seasonality to ARIMA. (Williams & Hoel, 2003) argue that SARIMA is a good model for short-term univariate forecasting, outperforming heuristic forecasts in their experiments. Space-Time ARIMA (STARIMA) is a spatial extension of ARIMA (Pfeifer & Deutsch, 1980). The dynamic turn ratio prediction (DTRP) model was proposed to update upstream traffic flow turn ratios based on household transformation (S. Deng et al., 2009). In (Min et al., 2009), the DTRP model updates the road network and the STARIMA model to forecast short-term traffic patterns. Based on the observation that correlation between traffic at two points may fluctuate, (Duan et al., 2016) propose a modified STARIMA model that incorporates time-varying lags. (Safikhani et al., 2020) apply the spatial-temporal autoregressive moving average (STARMA) model to predict short-term taxi demand. A modified Least Absolute Shrinkage and Selection Operator (LASSO) penalizes parameters. Both vector autoregressive (VAR) and STARIMA are outperformed by this model. However, they cannot capture rapid changes in traffic data. Also, parameter estimation takes a long time for SARIMA models. ARIMA does not capture dynamic traffic conditions but only static traffic conditions (Alsolami et al., 2020). Literature today focuses on machine learning, deep learning, and hybrid methods, since they can handle non-linear traffic flows and large data volumes better. These are also discussed in this section.

Other statistical approaches used for traffic forecasting include nonparametric regression method (Smith et al., 2002), Bayesian model (J. Wang et al., 2014), hidden Markov model (Qi & Ishak, 2014) and Kalman filter model (J. Guo et al., 2014).

(G. Yu et al., 2003) models the traffic flow as a fourth-order Markov chain model. Gaussian Mixture Model (GMM) is employed to model the transition probability between Markovian states, and its parameters are estimated using the Expectation maximization algorithm. Hidden Markov Models can model noisy observation of the traffic data better than Markov Chain models. They can embed mixture models as distributions, enabling them to capture complex traffic relationships better. Because of the reasons outlined above, (Qi & Ishak, 2014) propose a framework to predict traffic conditions using HMMs. In classical SARIMA models, constant parameter values are estimated. These methods cannot model uncertainties found in urban traffic. (Ghosh et al., 2007) propose a SARIMA model whose parameters are optimized using Bayesian Inference. Markov Chain Monte Carlo is employed to estimate the distribution of the parameters of this model. This paper solves the aforementioned problem by providing probability density curves instead of a constant prediction interval.

Gaussian Process Regression (GPR) is a nonparametric model that fits a nonlinear mapping from input to predictions. (Xie et al., 2010) propose a novel kernel-based Gaussian Process (GP) model for short-term traffic forecasting. Its performance outperforms ARIMA models. GP is based on the Bayesian framework, and it also estimates standard deviations in addition to the predictions. To further improve the accuracy of GP, (J. Zhao & Sun, 2016) propose a fourth-order GP dynamical model optimized using Maximum a Posteriori (MAP).

In dynamic systems that contain random noise or noisy real-time measurements, the Kalman filter (KF) (Kalman & Bucy, 1961) is a recursive filter that provides an optimal state estimation. Extended KF (EKF) exploits Taylor expansions to extend KF to nonlinear systems. (Y. Wang & Papageorgiou, 2005) propose a method for real-time traffic estimation using EKF. (Yuan et al., 2012) propose a new estimator using discretized Lagrangian (trajectory data) model as the process equation for Kalman Filtering instead of the previously standard Eulerian (loop detector data) state estimators. They also investigated using Unscented KF (UKF) instead of EKF. A first-order Taylor-series expansion is not a viable approximation for traffic system models. Therefore, EKF estimates around capacity may deviate from reality. The KF has high accuracy but is not compatible with dynamic traffic conditions. It cannot handle nonlinear traffic flows and is not accurate if the data is noisy (Alsolami et al., 2020); (Hu et al., 2016).

Having presented the state-of-the-art of non-parametric and parametric methods within the family of statistical learning approaches, now, we focus on machine learning techniques. Support Vector Machines (SVM), k-Nearest Neighbors (KNN) or Random Forest (RF) (Lana et al., 2018); (Vlahogianni et al., 2014) are examples of traditional ML methods, that are applied to traffic forecasting. A summary of the state-of-the-art at the point where ML and traffic forecasting converge is presented below.

SVM is a popular machine learning algorithm for classification and regression tasks widely applied in different fields. Generally, an SVM is conceived as a mapping of low-dimensional nonlinear data into a high-dimensional space using a kernel function. Two common variations of SVM are C-SVM and v-SVM; C-SVM is a standard SVM implementation in which two parameters control its behavior. v-SVM is a modification of the standard C-SVM in which it requires one parameter instead of two parameters. (Y. Zhang & Xie, 2007) propose using v-SVM for traffic forecasting, making the parameter selection and optimization of the SVM easier. However, although SVM provides good performance for traffic prediction, the kernel function and parameters should be carefully selected to ensure the correct implementation of an SVM (H. Yu et al., 2017). For this reason, many studies utilize SVM in a hybrid manner combined with other algorithms like evolutionary algorithms (Cong et al., 2016); (Hong et al., 2011); (Hu et al., 2016) and chaos wavelet analysis SVMs (J. Wang & Shi, 2013). We discuss some of them in the section on hybrid models at the end of this chapter.

The k-nearest neighbor (KNN) model can be used for both classification and regression tasks, and some studies have applied it to short-term traffic forecasting (Luo et al., 2019). (P. Cai et al., 2016) propose an improved KNN model based on spatiotemporal correlation to increase predicting accuracy and accomplish multistep forecasting. For short-term traffic flow prediction, (Ryu et al., 2018) propose a method for creating traffic state vectors using Mutual Information (MI). The MI evaluates the spatio-temporal correlations between the road sections in the urban road network after the variables with different time delays are produced using historical traffic time series. After that, a greedy search technique generates the traffic state vector by selecting the variables with the strongest correlation to the target traffic flow. The spatio-temporal correlation between road sections was incorporated using an updated KNN model into the prediction model.

(Y. Liu & Wu, 2017) use RF to predict the traffic congestion state. The RF algorithm is a supervised data mining algorithm, which is a classifier that utilizes a large number of Classification And Regression Trees (CART) decision trees. It selects k different sample sets from the data using the bootstrap method. These sample sets are used to construct unrouted decision trees. The decision tree is constructed by using the m candidate attributes randomly selected. Each tree of the decision tree finally gets k classification results. The final classification of the output variable is the one that gets the most votes of the type of output variables for the final category.

These traditional ML methods are suitable approaches when dealing with traffic scenarios where predictions must be made at a single point on the road or in one/more corridors. However, as moving data sources have progressively been incorporated into traffic estimation and prediction (e.g., GPS data), traditional ML methods face limitations when capturing the spatial and temporal relationships of traffic patterns (Yang, 2013); (L. Zhao et al., 2020); (Bogaerts et al., 2020). For this type of network-wide prediction problem, the common approach is to use Deep Neural Networks (DNNs), which allows exploiting the temporal and spatial characteristics of traffic data.

DNNs have been successfully applied in traffic forecasting problems in the past decade. They can capture very complex relations (LeCun et al., 2015); while they guarantee a better performance measured in terms of different accuracy and error metrics (Angarita-Zapata et al., 2019); (Do, Taherifar, et al., 2019); (Ermagun & Levinson, 2018); (Bogaerts et al., 2020). Some types of DNN are Recurrent Neural Networks (RNN) (H. Yu et al., 2017) and Convolutional Neural Networks (CNN) (X. Ma et al., 2017). Over the past few years, the introduction of Graph Neural Networks (GNN) has enabled transportation systems to model road networks, resulting in state-of-the-art performance in various traffic forecasting problems (Jiang & Luo, 2021). Furthermore, a recent attention mechanism that can be applied in GNN, RNN, and transformer architecture which utilizes self-attentions, has enhanced the performance even more. We discuss some of these models below.

(X. Ma et al., 2017) propose modeling a transportation network's spatial and temporal dimensions in two dimensions of an image, retaining spatial and temporal information due to the nearness of the surrounding road portions in the image. The forecasting is done using CNN, making it possible to extract spatiotemporal features without manual feature selection and construction. They suggest that the weight sharing property and pooling layers make it possible to use this model in large networks. This publication's method cannot be used in network-wide predictions. For this reason, (H. Yu et al., 2017) propose a spatiotemporal image-based approach by combining CNNs and Long short-term memories (LSTMs). From a different perspective, (S. Guo et al., 2019) capture spatial and temporal characteristics of the traffic by utilizing 3D convolutions. It considers local patterns and long-term traffic data patterns using a recalibration block.

The attention mechanism showed its potential in incorporating spatial and temporal dependencies in traffic data at an early stage. Attention is an interface connecting the encoder to the decoder. It provides information from the hidden state of the encoder to the decoder. By utilizing the attention mechanism, (Cheng et al., 2018) propose a framework that acquires spatiotemporal information about the transport network topology. They suggest that attention weights are beneficial for daily travel and path planning. In the meantime, (Yao et al., 2019) propose a flow gating mechanism to learn dynamic similarity in the spatial domain and a periodically shifted attention mechanism combining long-term periodic information with temporal shifting. Motivated by the potential of it, (Do, Vu, et al., 2019) utilize an attention mechanism to capture spatial and temporal dependencies using different data resolutions. The analysis of the calculated attention scores provides a posterior interpretation of sensor relationships in space and time. In recent work, (Fang et al., 2022) propose adding the attention mechanism to LSTM for short-term traffic flow forecasting. They argue that canonical LSTM networks cannot focus on long-term traffic flow evolution requirements. The attention mechanism could help solve this problem by allowing the model to assign different weights to various inputs and focus on vital data. They also investigate the regular inner pattern of the traffic flow data to recreate a training dataset by picking a specific range of input lengths.

(Laña et al., 2019) propose a method of obtaining long-term traffic volume forecasts that can be applied in real-time. In most of the methods discussed before, the prediction horizons are under 60 minutes.

Most of the literature consists of short-term predictions (Vlahogianni et al., 2014); (Laña et al., 2019). In (Laña et al., 2019), a long-term estimating technique is presented based on automated pattern recognition and integration with online change detection and adaption mechanisms. The framework uses evolving Spiking Neural Networks to adapt without retraining the model, allowing the entire system to act independently and in real-time.

The traffic prediction problem can also be formulated as a graph analysis problem as the road network is presented as graphs. In traffic forecasting literature, spectral GNNs are prevalent. These are called Graph Convolution Networks (GCN), and they calculate the convolution operator in the Fourier domain. This is done by calculating the eigen decomposition of the Graph Laplacian (Kipf & Welling, 2017); (Y. Li et al., 2018). (Y. Li et al., 2018) represents the spatial correlations between traffic sensors using a directed graph. The Diffusion Convolutional Network (DCN) filter is a spectral convolutional operator derived by modeling traffic flow as a diffusion process defined by a Markov process. The assumption is that the diffusion process will converge to a stationary distribution after a number of time steps, and it describes the transition probabilities between nodes. Then, a Gated Recurrent Units (GRU) is built by replacing matrix multiplication with the diffusion convolutional layer. The sequence to sequence (Seq2Seq) architecture was employed to enable multistep forecasting. The encoder maps the input to a hidden space, and then the decoder employs this hidden space to forecast future values.

(Liao et al., 2018) aims to enhance traffic prediction by incorporating important features stored in auxiliary data. Traffic around a location may increase if many online searches are made due to a public event. The learner uses an encoder-decoder framework that includes offline geo-social traits and information about road intersections, such as road layout and national holidays. This study releases a large-scale dataset with auxiliary data, such as crowd-sourced map queries, road junctions, and geographical and social features. Geographical and social features are fed into the Seq2Seq architecture. They then incorporate the traffic speed of adjacent road segments into the GCN.

In many graph neural network models, future predictions are generated via a step-by-step process, causing error propagation. Therefore, (C. Zheng et al., 2020) propose an attention mechanism that alleviates this problem by incorporating the relationship between historical and future time steps. They propose a Graph Multi-Attention Network (GMAN) which implements an encoder-decoder architecture. Both the encoder and the decoder comprise several spatiotemporal attention blocks. The encoder encodes the properties of the input traffic, while the decoder predicts the output sequence. A transform attention layer is used between the encoder and the decoder to turn the encoded traffic characteristics into sequence representations of future time steps as the decoder's input. A gated fusion algorithm is also proposed to adaptively fuse the information extracted by spatial and temporal attention mechanisms. The Node2vec algorithm is employed to map traffic networks to latent space. Meanwhile, (J. J. Q. Yu et al., 2021) also focuses on the long-term predictions and error propagation. They propose to use a predictor-regularizer design, which splits traffic data into trend, period, and closeness classes, for long-term urban traffic forecasting using GNNs.

(L. Cai et al., 2020)'s work on the other hand is inspired by Google's Transformer framework for machine translation. They propose a hybrid deep learning architecture based on the Transformer and GNNs to model spatiotemporal signals. It presents novel positional encoding strategies and a "similarity-based combination" approach to incorporate position embeddings into the model. This paper also employs Time-series segments; that is, it includes segments within the same timeframe as the predicting timeframe on the last few days. Spatial dependencies are captured by GNNs. GCNConv (Kipf & Welling, 2017) and Diffusion Convolution (Y. Li et al., 2018) are compared in this study. Diffusion convolution improves performance slightly. The transformer architecture resembles the original architecture (Vaswani et al., 2017), but the input signal is fed into a GNN before going to the transformer.

(Bogaerts et al., 2020) also propose a hybrid approach, namely a graph CNN-RNN architecture, since hybrid approaches that combine graph CNNs and RNNs have shown a more superior performance than the RNN and graph CNN separately (Cui et al., 2018); (H. Yu et al., 2017); (Jin et al., 2018); (L. Zhao et al., 2020). They propose a deep neural network that considers traffic's geographical and temporal aspects to create short and long-term predictions, in which the latter was still an open issue in literature since the main focus was on short-term predictions before. Furthermore, they suggest using temporal correlation to reduce the data set and select the most crucial road links. They then propose a data reduction approach to identify the most relevant road links for forecasting. An architecture for trajectory data is presented, with two variations that differ primarily in the CNN's layer. The first uses the standard convolution and max-pooling layers, while the other uses max pooling before convolution to reduce the input data's dimension even more without losing important information.

(M. Li & Zhu, 2021) argue that the standard approach of utilizing a given static adjacency matrix ignores temporal correlations and similarity among traffic flows and may fail when that matrix is incomplete. They also claim that separate modules for capturing spatial and temporal correlation are not a perfect solution for spatio-temporal forecasting. They propose a data-driven graph generation approach that exposes hidden spatial and temporal dependencies. They then propose an algorithm to fuse these dynamic and static networks and a model that simultaneously learns spatial and temporal dependencies of traffic. It is worth mentioning that in this model, matrix multiplication is utilized instead of the standard Laplacian filter in GCNs. In order to address the same problem, (Lee & Rhee, 2022) propose adding spatial relationships that are not Euclidean, such as positional relationships and direction, to the mentioned approaches. It adds an inductive bias to the model that helps the neural network learn patterns better. They also propose a novel deep learning model to incorporate these auxiliary data.

Although attention-based methods include spatial correlation, they ignore the adjacency relationship of the road network structure. Therefore, some papers focus on spatiotemporal dependencies and their dynamic changes. In this regard, (Yin et al., 2021) propose an attention-based GNN that captures multiple time series interactions, the spatial correlations within the same order and between different neighborhoods, and dynamic temporal dependencies. In this method, the attention mechanism dynamically collects spatial connections within the same order region and between different neighborhoods. On the other hand, (Z. Zhang et al., 2021) propose a GCN to model spatial correlation by learning the spatiotemporal dependency and fusing several graph topologies. The fusion is done using an attention mechanism calculating attention between multiple network matrices, and GRU and GCN were integrated to collect spatiotemporal correlations.

The Attribute Augmentation Unit (J. Zhu et al., 2021) combines external factors such as the weather and points of interest as static and dynamic attributes. Historically, traffic prediction methods used historical traffic data to predict traffic states. However, external factors also affect the future states. They embedded the external factors into a knowledge graph and proposed an algorithm to fuse embeddings derived from these graphs with traffic information. As a final step, they utilize well-known GCNs such as Diffusion Convolutional Recurrent Neural Network and temporal graph convolutional networks as the foundation of their method and report superior performance over vanilla GCNs.

Transductive transfer learning is a common approach in deep learning to pretrain neural networks, but it is rarely used in traffic forecasting because the algorithms learn location-specific traffic patterns, while graphs have complex spatial correlations. (Mallick et al., 2021) extract road network subgraphs such that each subgraph resembles a function of network connectivity and temporal patterns. Then, several Diffusion Convolutional Recurrent Neural Network models are applied to predict highway speeds. This way, the trained model can be applied to forecast traffic in similar traffic networks.

3.1.2 Simulation-based estimation and prediction approaches

Traffic simulation models aim to help planners and designers of transportation systems plan and design their routes. Unlike the parametric and non-parametric data-driven methods, simulation models do not rely on historical or real-time traffic data but rather simulate traffic, as no historical data is available when developing a road network. A traffic simulation model can estimate future traffic levels to evaluate the suitability of a transportation system design; however, a model cannot predict the next state of traffic based on historical and real-time data (Nagy & Simon, 2018).

Models for traffic simulation were first developed by (Beckmann et al., 1956). The movement of vehicles is simulated in these models through an origin-destination (OD) matrix, which describes a region's movement. Each cell in the OD matrix represents the number of trips from the origin to the destination. The row expresses the origin, and the column expresses the destination. Traffic simulation models can be classified as microscopic, macroscopic, or mesoscopic (Nagy & Simon, 2018).

In the microscopic perspective, each driver and their interactions are modeled using a multi-agent system where every agent records information or behavior for each trip (Xiong et al., 2015). Cellular automata (CA) is a typical microscopic simulation method in which a road is divided into cells that a vehicle can occupy, and time is discretized into steps. The CA model reproduces a wide variety of traffic phenomena well and is numerically very efficient due to its simplicity. A combination of this model with OD predictions has been used to develop network-wide traffic predictions (Miska, 2007; Nagy & Simon, 2018).

(Kerner & Klenov, 2002) first propose a mathematical model for the three-phase traffic theory. The primary purpose of this theory is to explain how traffic breakdowns occur on highways, resulting in a congested traffic situation. It explains traffic using Freeflow, Synchronized flow, and Wide moving jam phases. Then (Kerner et al., 2013) present the Kerner–Klenov–Schreckenberg–Wolf (KKSW) cellular automaton model, a three-phase CA traffic flow model for real-world scenarios including multilane roadways with passenger cars and trucks.

The macroscopic analysis evaluates each road segment's density, speed, and vehicle count (Ngoduy & Wilson, 2014). Static and dynamic traffic assignment are two methods that allocate traffic on a simulated road network. Temporal OD matrices are necessary for models of traffic flows over time in dynamic assignment models, but static assignment models only require a static OD matrix of overall trip demand of actors (Kemper, 2005). In comparison to older statistical and machine learning models, modern neural networks have superior performance. However, their data-hungry nature makes it challenging to train them in cities and scenarios with little available data. This has led (K. Zhang et al., 2021) to develop a simulator that provides an excessive dataset for training neural networks and statistical methods. In addition to macroscopic models, formulas derived primarily from hydrodynamics can be used to simulate changes in traffic density and vehicle count over time in urban areas (H. Liu et al., 2005); (Nagy & Simon, 2018). Mesoscopic approaches have the advantage of modeling a wide range of phenomena, such as traffic signals, freeway merging, weaving sections, or high-occupancy traveling lanes (Chiu et al., 2010); (Nagy & Simon, 2018).

Several recent publications have examined the new challenges associated with electric and connected vehicles. (Yan et al., 2020) focuses on electric vehicles charging load simulation because the lack of historical data caused by the development of new infrastructures limits forecasting capacity. They have not only taken into account the common factors of the problem, such as traffic congestion and nominal battery capacity but have also considered other impactful factors such as the effect of temperature on air conditioning power consumption and battery capacity as well. Connected and automated vehicles are validated, verified, and tested using scenario-based testing. For obtaining a large number of urban traffic scenarios, (Yue et al., 2020) proposes a low-cost method based on microscopic traffic simulation. They model U-turns and parallel driving, and various kinds of collision accidents and develop a new criticality metric, Scenario Risk Index, to evaluate the risks of connected and automated vehicles. (Shao & Sun, 2021) focuses on modeling short-term traffic prediction in traffic scenarios where both connected and unconnected vehicles are available based on a macroscopic traffic flow model. In recent work, (B. Liu et al., 2022) proposes a dynamic chain propagation model to evaluate the spatiotemporal correlation of road networks via traffic perturbation simulation. The model can be used to design road networks, formulate travel routes, and decrease traffic congestion by integrating dynamic traffic flow and static road network structure.

3.1.3 Hybrid models (statistical-ml, ml/statistical-simulation, ensembles)

The last type of methods to be discussed is the hybrid approach. Within this category, two or more algorithms, from parametric or non-parametric data-driven types, are combined to find synergies that improve their isolated performance (individual performance of either statistical learning or machine learning methods). However, a hybrid model is computationally complex and requires extensive storage capacity (Alsolami et al., 2020). We discuss some of these models in this section.

(Hong et al., 2011) propose a hybrid model using the Support Vector Regression (SVR) model as the predictor and optimize it using the continuous ant colony optimization algorithm for traffic flow prediction. (Hofleitner et al., 2012) propose a statistical-simulation hybrid model. Their model of traffic flow dynamics is modeled as a macroscopic continuum, which is a standard assumption in traffic flow dynamics modeling. By doing so, the model can estimate several static parameters of the streets, such as free flow velocity, that may help the machine learning model forecast traffic. An optimization algorithm based on Expectation-Maximization is used to optimize the model parameters. A particle filter is used in the forecasting algorithm to estimate traffic states based on the available data and learned parameters.

By combining support vector regression and particle swarm optimization (PSO) algorithm, (Hu et al., 2016) propose a Hybrid PSO-SVR forecasting method that optimizes SVR parameters using PSO. (F. Guo et al., 2018) propose a novel ensemble method using Neural Networks, SVMs, and RFs as stand-alone models and compare average, weighted, and KNN-based fusion methods.

(Song et al., 2018) present a model by combining SARIMA and Seasonal discrete Grey model structure (SDGM) for traffic speed prediction. The Discrete Grey Model (DGM) model can predict cross-sectional data, but it cannot capture the oscillations of seasonal data. The SDGM resolves this issue by incorporating the cycle truncation accumulated generating operation (CTAGO) operator into the DGM. The proposed hybrid framework outperforms the SARIMA, SGDM, ANN, and SVR models. (T. Ma et al., 2020) propose a traffic forecasting framework by fusing Multi-Layer Perceptron (MLP) and the ARIMA model. ARIMA model extracts time series residuals from MLP and improves its performance.

In recent work, (Fu et al., 2020) propose a Bayesian spatio-temporal graph convolutional network to exploit both the superior performance of GCNs and the interpretability of Bayesian frameworks.

3.2 Network-wide and local predictions

Traffic forecasting can be performed on a local or network-wide level. Local prediction is the prediction of the future conditions at a single point/road section. On the other hand, network-wide predictions integrate multiple input and output variables, which correspond to points or road sections of an entire road network.

Classic statistical and parametric models can be used efficiently for local predictions. The family of ARIMA (Hamed et al., 1995) and SARIMA (Williams & Hoel, 2003) are two of the most valuable models for local or single-point predictions.

Most of the modern machine learning methods that predict traffic flow, speed, and so on can be exploited for network-wide predictions. For example, the input of temporal CNNs and GNNs are all of the nodes in the road network, and they forecast all of the nodes in a single run (Lee & Rhee, 2022); (H. Zheng et al., 2020); (H. Zhu et al., 2020). One of their most significant characteristics is that they can handle the higher dimensionality of the input, as well as the non-linear correlations.

3.3 Anomaly detection

Traffic anomalies like accidents, traffic jams, roadworks, and sport events have a serious impact on the traffic state. Therefore, early detection and forecasting of these extreme conditions and special events may aid in returning traffic conditions to normal (Jiang & Luo, 2021). Having said that, this section reviews the methods for detecting traffic flow anomalies, which usually capture traffic information using sensors and detect anomalies originating from hardware failure or traffic congestions.

In a data-driven context, traffic anomalies are represented by outliers, which are data points far away from the distribution in which the data originated, while an inlier lies in the interior of that distribution. Considering this distinction, (Djenouri et al., 2019) divide traffic anomaly detection into flow outlier detection and trajectory outlier detection categories.

First, the category of flow outlier detection is discussed. Traffic flow measurement uses equipment like inductive loop detectors, video surveillance, laser, infrared, microwave, acoustic, and ultrasonic sensors.
(Lam et al., 2017) assume that inlier data follow a normal distribution; they propose utilizing GMM and Naive Bayes (NB) for detecting traffic anomalies and outliers resulting from measurement hardware failures. Video surveillance is first converted into spatio-temporal data and then projected into a 2d plane using principal component analysis (PCA). In the first method, this 2d method is fed to NB to find the best inlier bandwidth region, and the test data outside of this region are considered outliers. The same is applied to the GMM model in their second method. This study only focuses on detecting anomalies at a single crossroad and cannot utilize information from adjacent roads and junctions; also, using video surveillance to capture traffic flow introduces some noise into the data. In recent work, (Sofuoglu & Aviyente, 2022) propose an approach for spatiotemporal low-rank tensor decomposition and model anomalies as the sparse part of it. Then observations are scored using standard scoring methods like One-Class SVM (Schölkopf et al., 2000), Local outlier factor (Breunig et al., 2000), and Elliptic Envelope (Rousseeuw & van Driessen, 1999).

There is an increasing trend among researchers to calculate anomaly scores through the use of deep learning and machine learning models. (Kong et al., 2020) define an anomaly score using the difference between historical data and the flow predicted by LSTM. Finally, extreme conditions are detected using this score, and two other anomaly scores and candidates are calculated using Pearson Correlation and One-Class SVM. (Liu et al., 2020) predict traffic anomalies using supervised multi-modal deep learning architecture composed of GCNs for spatial embedding and GRUs for temporal embedding. (L. Deng et al., 2022) use Generative Adversarial Networks for traffic anomaly detection. The generator and the discriminator both employ graph convolutional GRUs, GCNs, and FCs. The generator aims to construct realistic fake data, and the discriminator aims to discriminate those fake data from real ones. Prediction error from the generator and output of the discriminator are combined to form the anomaly score.

Second, the methods for trajectory outlier detection are discussed below. The methodology of trajectory outlier detection detects traffic anomalies based on Global Positioning System (GPS) data from public transportation, navigation software in smartphones, or GPS hardware in private automobiles.

In another work exploiting deep learning models, (Boquet et al., 2020) utilize variational autoencoders (VAE) to project traffic data into a compressed manifold. It is important to note that in VAE, a latent space is a mixture of distributions in contrast to standard autoencoders, which are constant vectors. VAEs optimization is done using Kullback–Leibler divergence (KL-divergence), which measures differences between two probability distributions. They find out that KL-divergence works better than the classical reconstruction loss in autoencoders.

Similar to KL-divergence, the Kolmogorov–Smirnov test quantifies if a sample is drawn from a given probability distribution. In (Huang et al., 2020), speed distributions are derived from GPS trajectory data in the entire dataset as a reference, and the test day data is regarded as an empirical distribution. As part of the Kolmogorov–Smirnov test, P-value statistics are computed. It is proposed that logarithmic values of the P-value serve as anomaly scores, also called Visible Outlier Indexes. Further, the data is aggregated using a proposed Meshed Spatiotemporal Neighborhoods algorithm, and an additional learning method such as SVM and neural networks is utilized for anomaly prediction.

4 Demand estimation and prediction

This section aims to review the state-of-the-art of demand estimation and prediction methods based on the problem formulation and solution algorithms that have been proposed in the literature. Estimation and prediction of travel demand is important since the demand is an essential input to various transport operations and management applications that require knowledge about the spatial and temporal distribution of the demand for a specific transportation network of interest. The demand patterns are commonly expressed as dynamic (time-dependent) origin-destination flow matrices and each cell in the matrix represents the number of trips from an origin to a destination for a specific mode of transport. The term dynamic refers to the temporal variation of trip departures over a specific period (e.g. day, morning-peak). The demand can be also defined as inflows or turning fractions into the network. Section 4.1 defines the dynamic demand estimation problem based on classifications proposed in literature. Then, section 4.2 focuses on simulation-based methods for dynamic demand estimation and prediction methods, while section 4.3 describes the data-driven methods.

4.1 Definition of the Dynamic Demand Estimation Problem (DODE)

The demand estimation and prediction problems have been studied intensively in the literature for several decades. However, due to the complexity and difficulty to observe directly the OD flows along the available paths, the problem still entails significant challenges. The various OD estimation and prediction approaches that have been proposed in the literature can be generally classified based on the demand profile to static and dynamic as well as to direct and indirect based on the observability of OD flows. The dynamic OD estimation methods can be further categorized as **offline** (Cipriani et al., 2011); (Antoniou et al., 2015); (Osorio, 2019a) and **online** (real-time) approaches, depending on the application. The offline estimation refers to the estimation (or calibration) of a set of time-dependent OD matrices given a set of available time-series of link traffic counts (potentially speeds and travel time measurements can be also used) and historical OD matrices. The offline estimation approaches are more relevant to transport planning and evaluation studies (e.g., design and evaluation of traffic management strategies), whereas the online OD estimation and prediction methods are used in the context of real-time traffic management and route guidance, aiming at predicting future OD flows given real-time traffic information from the network as it becomes available. Therefore, in online applications the dynamic OD estimation methods must be able to provide fast estimates for recent time intervals with predictions for future time intervals (Antoniou et al., 2016). OD matrices derived from offline estimation are usually used as input for the online OD estimation problem.

As mentioned above, the estimation of OD matrices can be performed using **direct** and **indirect** methods. In direct estimation methods, OD flows can be derived from travel surveys and interviews for a sample of travelers. However, this approach is difficult and expensive. To overcome the limitation of observing the OD matrices directly, indirect estimation approaches have been proposed that rely on the use of traffic measurements to infer the OD flows. The quality and information provided by the input data (traffic measurements) are critical factors for the accuracy of the estimated OD matrices (Tympakianaki, 2018); (Tympakianaki et al., 2015).

The remainder of section 4 focuses on the review of state-of-the-art methods for dynamic demand estimation and prediction methods. The approaches are divided into simulation-based and data-driven methods that have been proposed in the literature. The advantages and limitations of each approach are emphasized and discussed.

4.2 Simulation-based demand estimation and prediction methods

Traffic models are used to study various complex and large-scale problems, from transport planning to evaluation of traffic management strategies as well as for the estimation and prediction of the transport network traffic state (Ben-Akiva et al., 2001); (Mahmassani, 2001); (Tampère et al., 2010). Dynamic traffic assignment (DTA) and route choice models can capture the traffic dynamics and interactions between the network supply and the demand and representing the evolution of congestion in time and space. Demand is a key input to traffic simulators and DTA models, regardless of the application, along with other parameters (e.g. route choice models) in order to estimate and predict the network performance and capture the travel behaviour of travelers. The demand can be fed into the simulator in the form of OD matrices or inflows. Subsequently, the traffic assignment determines how the traffic demand is loaded onto the network.

Simulation-based approaches are commonly adopted for both the offline and the online applications as simulation models can capture the nonlinear dependence between the travel demand and the traffic measurements through the use of DTA models (Carrese et al., 2017). The literature review in this deliverable focuses on dynamic OD flow estimation methods, which is relevant for the traffic monitoring and forecasting solutions that will be developed in the TANGENT project. The dynamic OD estimation problem, searches for time-dependent OD demand matrices able to best fit measured traffic data (Cascetta & Postorino, 2001). A review of various approaches for offline as well as online demand estimation can be found in (Antoniou et al., 2016); (Djukic, 2014); (Balakrishna et al., 2005). Furthermore, two procedures are proposed for the dynamic OD demand estimation problem: simultaneous or sequential (Cascetta & Marquis, 1993); (Cantelmo & Viti, 2020). The first approach estimates the OD flows for the whole demand period using traffic count information for all intervals simultaneously. In the sequential approach the OD matrices are estimated individually, one interval at a time, using traffic counts and OD estimates of the previous intervals. The latter approach is used in online applications in order to reduce the computational complexity of the problem, which is a critical requirement in the online estimation process.

In the following sections various approaches that have been proposed in the literature for the offline and online DODE are presented, respectively.

4.2.1 Offline DODE approaches

Offline demand estimation models focus on medium-long term planning, while the online are frequently adopted for real-time applications, such as route guidance. The **offline** dynamic OD demand estimation problem can be formulated as an optimization problem aiming to find an OD demand matrix that best reproduces a set of traffic measurements when assigned to a given network (Balakrishna, Antoniou, et al., 2007); (Cipriani et al., 2011) (Antoniou et al., 2015); (Tympakianaki, 2018). Different problem formulations have been proposed in the literature for the offline estimation of dynamic OD demand.

The dynamic demand estimation methods can be categorized based on the mathematical problem formulation to divide them into **assignment matrix-based** (Spiess, 1990); (Cascetta & Postorino, 2001); (Frederix et al., 2011); (Toledo & Kolechkina, 2013) and **assignment-matrix free** algorithms (Cantelmo et al., 2014), through linear transformation (Tympakianaki, 2018; Tympakianaki et al., 2015). The assignment matrix represents proportions of OD flows crossing sensor locations over time. The most common assignment-matrix-based formulation uses observed flows and expresses them as a linear function of the assignment matrix, assuming an analytical representation of the relationship between demand and traffic flows in order to estimate the demand matrix. The estimation of dynamic

OD matrices for large-scale congested networks is a complex and underdetermined problem as the number of equations that captures the relationship between the OD flows and available traffic measurements is less than the number of unknown OD flows. Usually, the traffic observations used in OD estimation are aggregate, time-dependent link counts (e.g. from loop detectors) and historical OD matrices. However, traffic counts alone are unable to provide information about the traffic regime (congested or uncongested network), hence they are not reliable in estimating accurate OD matrices. Hence, the main disadvantage of the OD estimation formulation based on assignment matrices is the limitation to use only traffic count measurements for the estimation due to the use of assignment matrices (Balakrishna et al., 2008), (Tympakianaki et al., 2015). In particular, the assumption regarding a linear relationship between OD flows and traffic count measurements does not hold for other traffic measurements (densities, speeds), as that relationship is complex and non-linear, hence, it cannot be modeled using linear approximations. Specifically, in congested networks the assignment proportions depend highly on the prevailing travel times and route choices, hence, the relationship between demand and traffic counts becomes highly non-linear. Another limitation of the assignment-matrix based OD estimation methods is that they do not allow for simultaneous estimation of the model supply and other demand parameters (e.g. route choices). Furthermore, the computational effort of assignment-based OD estimation methods is high due to the requirement to calculate and store assignment matrices.

On the other hand, other approaches do not rely on the calculation of assignment matrices, assignment-matrix free algorithms (Balakrishna, Ben-Akiva, et al., 2007); (Balakrishna, Antoniou, et al., 2007); (Vaze et al., 2009); (Balakrishna & Koutsopoulos, 2008), but use the output of e.g. a simulation model to capture the complex relationships between the OD flows and any available data. The proposed more general problem formulations and solution approaches can accommodate quite general data sources, as well as the demand and supply inputs and parameters of the DTA models can be jointly estimated. Nevertheless, such approaches may have performance limitations (e.g. computational performance). In (Antoniou et al., 2016) the authors proposed an evaluation and benchmarking framework that can be used to implement and compare various OD estimation approaches for offline and online applications.

Numerous methodologies have been proposed in the literature to address different challenges and issues associated with the estimation of OD matrices. The aforementioned limitations of assignment matrix-based formulations with respect to the addition of traffic measurements besides link counts has been tackled in the literature. In (Antoniou et al., 2004), the authors considered the inclusion of Automated Vehicle Identification information in the OD estimation problem formulation. However, this approach has limitations when applied to congested networks, due to the non-linear relationship between density and flow. (Frederix et al., 2011) proposed a sensitivity-based approximation of the relationship between OD flows and link flows to overcome the influence of non-linearity due to congestion. A marginal computation method is proposed to calculate the sensitivity of the link flows to the OD flows under congested conditions. (Toledo & Kolechkina, 2013) proposed a general solution scheme for the OD estimation problem considering congestion effects in the network. The approach relies on linear approximations of the assignment matrix using a first-order Taylor expansion and different gradient-based solution algorithms (e.g. relative gradient, quasi-Newton methods and meta-heuristic algorithms) were proposed.

The increasing availability of emerging data technologies (e.g. Bluetooth sensors) provides the opportunity for the collection of real-time traffic data (such as travel times) that can be incorporated into the calibration process of the DTA models' inputs and parameters, in particular, the demand input. In that context, several research studies have focused on general formulations and solution approaches that are flexible in accommodating **heterogeneous data sources** (e.g. traffic speeds, travel times) in the estimation process, besides traffic counts, aiming to improve the accuracy and quality of the estimated OD matrices, as better and more accurate information about the real network conditions can be provided (Balakrishna, Ben-Akiva, et al., 2007b);(Balakrishna & Koutsopoulos, 2008); (Cipriani et al., 2011); (Frederix et al., 2011);(Vaze et al., 2009); (Djukic et al., 2014); (Tympakianaki et al., 2015). The benefits from the incorporation of additional traffic information in the demand estimation problem has been demonstrated in different studies (Eisenman & List, 2004); (Cipriani et al., 2011); (Zhou & Mahmassani, 2006), using data from various sources (such as speeds, travel times from vehicle identification data (AVI) and/or probe vehicles) (Tympakianaki, 2018). Several variants of the SPSA algorithm have been proposed in the literature aiming to overcome the limitations for the basic version of the algorithm and improve the accuracy of the estimated OD flows as well as the convergence performance. Focusing on the dimensionality reduction of the OD flows estimation problem, (Cipriani et al., 2014) investigated the effectiveness of incorporating different types of traffic information in the offline dynamic OD estimation problem using the SPSA AD-PI algorithm proposed by (Cipriani et al., 2011). The results demonstrated that the inclusion of path travel times combined with link speeds and flows improved the accuracy of the estimated OD matrix. (Cipriani et al., 2013) and (Kostic et al., 2017) performed sensitivity analyses of different implementation aspects of SPSA and applied techniques for improving the convergence properties of the SPSA algorithm. The results in both studies show the significance of averaging a sufficient number of gradient replications or bounding the magnitude of the approximated gradient values, in order to achieve fast convergence. Finally, both studies have found crucial the selection of the algorithmic parameter values for the effectiveness of SPSA. (Lu et al., 2015) and (Antoniou et al., 2015) proposed a modification of the SPSA algorithm for joint demand and supply offline calibration of DTA models by adding structural information in space and time between the model variables that need to be calibrated and the traffic measurements. The algorithm is called Weighted-SPSA (W-SPSA), where the weight matrix can represent, for example, the correlation between OD flows and network traffic measurements. Depending on the relevance of the different measurements a different weight is assigned, thus, reducing gradient estimation errors. (Tympakianaki et al., 2015) and (Tympakianaki et al., 2018) proposed techniques to enhance the performance and robustness of SPSA. The modifications involved scaling the step size based on the magnitude of the estimated OD flows as well as two clustering strategies for the OD variables based on the OD flow magnitudes and OD pair interactions on common sensor links. The results demonstrate that the modified algorithm, referred to as cluster-based SPSA (c-SPSA), can improve the accuracy of the estimated OD flows.

With respect to the **observability** of OD flows and underdetermines of the problem, several studies have tackled the issue in estimating dynamic OD flows using various approaches. Some authors addressed this issue using various approaches to reduce the high dimensionality of the OD estimation problem using principal component analysis (PCA) (Djukic et al., 2012). Within the context of the OD estimation problem, the PCA dimension reduction technique maps through linear transformation, the authors tackle the problem of non-linearity and dimensionality of OD flows in relation to the calibration and validation approaches for the DTA models. The authors highlight the limitations of the widely used SPSA approach and its variants as the solution algorithm for the calibration problem. They propose a modification of the SPSA algorithm combined with PCA in order to limit the search area of SPSA, hence, reducing the dimension and complexity of the problem. This is achieved by introducing structural relationships from historical estimates in lower dimensions. In (Qurashi, Lu, et al., 2022) the authors use the PC-SPSA (Qurashi, Ma, et al., 2019) to address the scalability issue of dynamic demand estimation on the calibration of large-scale DTA models. The authors propose a novel approach to overcome the limitation of deploying the PCA technique when historical OD estimates are unavailable or irrelevant.

4.2.2 Online DODE approaches

The **online** dynamic OD demand estimation is a very challenging problem due to the non-linearity and observability issues associated with the demand estimation problem as mentioned in the previous paragraphs. For real-time applications, the KF algorithm has been commonly used (Kalman, 1960). The algorithm solves a least-square problem in an incremental way, allowing for the update of the unknown time-dependent OD flows as additional traffic data become available. The Kalman filtering approach uses deviations of OD flows as variables, which incorporates structural information about the relationship between historical OD flow estimates.

Despite the number of research studies that have widely adopted the KF algorithm, its application to the online demand estimation problem is still problematic. The challenges and complexity of the problem lie in several sources as identified in (Castiglione et al., 2021). These are related to the size of the variables to be estimated (OD flows), the non-linearity between variables (traffic measurements and OD flows) as well as the underlying demand structure. The KF algorithm assumes a linear relation between variables, which makes this problem formulation inadequate to represent the traffic dynamics. Hence, there is a need to apply non-linear models (Antoniou et al., 2007). Different formulations for the online applications of the dynamic OD estimation and prediction have been proposed in the literature (Antoniou et al., 2007), several adopting a state-space formulation using Extended, Limiting Extended, and Unscented Kalman filters (Ashok & Ben-Akiva, 2000); (Antoniou et al., 2007); (Prakash et al., 2018); (Cantelmo & Viti, 2020); (Krishnakumari et al., 2019); (Barceló & Montero, 2015); (Liu et al., 2020).

In (Carrese et al., 2017), the authors tackle the problems of both off-line and on-line dynamic demand estimation in the context of recent technology developments and opportunities to collect heterogeneous traffic data that can be incorporated into the estimation problem. For the offline problem, Floating Car Data (FCD) have been used in the demand estimation, including dynamic route choice information and route travel times. The results demonstrated the improvement of the demand estimation solution and the increased accuracy of the estimated OD flows with the inclusion of the OD travel times and route choice probabilities. For the online demand estimation and prediction, an extension of the Kalman filter theory, the Local Ensemble Transformed Kalman Filter (LETKF), has been applied. The benefit of this approach is that no assumption regarding the linear dependency between OD flows and traffic measurements was made, hence the incorporation of new sources such as FCD was feasible. LETKF is found to outperform the common non-linear KF implementations.

With respect to the dimensionality reduction, the concept of principal component analysis has been introduced in different problem formulations that have been proposed. For instance, the PCA was first used with colored Kalman Filter (Djukic et al., 2012), followed by the development of Principal Component–General Least Square (PC-GLS) and Principal Component– Extended Kalman filter (PC-EKF) algorithms (Prakash et al., 2017) for online OD calibration. More recently, in (Qurashi, Ma, Chaniotakis, et al., 2018), the authors applied the PC–SPSA algorithm to online OD demand calibration for DTA models. The results demonstrate the advantages of the PCA technique for dimensionality reduction combined with SPSA in improving the quality of the calibrated OD demand as well as reducing the computational effort. (Castiglione et al., 2021) proposed a non-linear Kalman Filter framework for online dynamic OD estimation that reduces the number of OD flow variables to be estimated using the PCA technique. The proposed approach incorporates heterogeneous data sources (such as link counts and speeds) to capture the spatial correlations between traffic measurements and time-dependent OD-flows, without the need to explicitly map this relationship. The main advantage of this approach over existing methods is the combination of PCA with the LETKF that does not depend on the assignment matrix information.

4.3 Data-driven demand estimation and prediction methods

The literature review presented in this section so far, focused on solution methods for the DODE that use the DTA model to capture the correlation between the demand and supply. While numerous approaches have been proposed that use simulation models to capture the interactions between the demand and supply resulting in the prevailing traffic conditions for the whole network, a research gap is identified regarding data-driven approaches for the demand estimation and prediction problems. The main reason is the complexity of the problem that is severely underdetermined for large-scale congested networks as well as the lack of adequate and complete traffic information in order to infer the demand dynamics in space and time. For example, in (Krishnakumari et al., 2020) the authors propose a data-driven method to estimate OD flows without the need for an iterative dynamic network loading (to result in equilibrium) using link speed and flow measurements and deriving path travel times. Other attempts make use of cellular network data to extract information on travel patterns and infer aggregated OD matrices (for example (Breyer et al., 2017)).

Several data-driven approaches using machine learning techniques have been proposed in the literature to predict OD matrices for public transport planning and operation purposes, based on the available real-time and historical information data for a specific network (for example (C. Li et al., 2020); (Hadjidimitriou et al., 2021)).

The estimation of dynamic demand for simulation-based DTA models is a computationally expensive procedure, especially for large-scale congested networks where it is difficult to connect the OD matrix with the network traffic dynamics at aggregated levels. Recent studies focus on developing metamodel-based methods for the calibration of dynamic OD demand in large-scale networks (Osorio, 2019b). In (Dantsuji et al., 2022) the authors propose a novel approach which combines data-driven approaches with simulation for the calibration of car OD matrices. A metamodel approach is developed that optimizes the demand using the network-wide Macroscopic Fundamental Diagram (MFD), using multiple data sources to derive empirical bi-modal MFDs. The benefit of the metamodel framework is that it accounts for the mismatches between region-based analytical and simulation-based traffic assignments. Such hybrid approaches can be advantageous in order to overcome the computational limitations of the simulation-based demand estimation approaches as well as the observability challenge of the OD matrices. More specifically, the estimation of real traffic conditions at all links is normally infeasible due to lack of complete data for the whole network. Sparse availability of link counts is also problematic for large-scale networks. However, in the presence of network-level ground-truth data the fundamental diagram can be constructed and included in the optimization process reduces significantly the problem dimensionality. The authors demonstrated the transferability of the proposed approach to networks with different topology.

5 Research gaps and challenges

Recently, (Jiang & Luo, 2021) listed some research gaps, which are also of interest in the Tangent project. Many publications consider a limited period of time, most commonly less than a year, in their testing and training procedures. This cannot represent all the traffic patterns. Additionally, many publications analyze cleaned datasets that do not reflect real world scenarios. Furthermore, these do not contain enough anomaly data for deep learning training procedures, resulting in many methods' inability to consider these events in their forecasting. Also, a limited number of publications work on transfer learning which could be beneficial in forecasting traffic in cities without enough traffic data.

Furthermore, two major research tracks related to traffic forecasting with deep learning are:
1. developing efficient representations of multi-modal traffic networks (B. Yu et al., 2018) and
2. efficient large-scale forecasts based on historical data on short to medium time (up to 4 hours) horizons, flexible and applied on different transport modes (Bogaerts, 2019).

The SotA of research using GNNs in various traffic prediction problems focuses on the supply side. This means that it misses a lot of information coming from the demand side, which we will include from transport models in WP4. Furthermore, a focus on the exploitation of event data into spatio-temporal graph convolutional neural networks is included in the development and research of WP4.

With respect to the demand estimation and prediction problem, the main challenges identified in the SotA are the following:
1. *Observability.* The size of the unknown variables (OD flows) to be estimated/predicted is a key factor in the estimation of the OD flows due to the underdetermines of the problem. Dimensionality reduction methods have been proposed in the literature to tackle this challenge that can be explored and extended depending on the quality of the historical OD matrices as well as the availability of traffic measurements.
2. *Non-linear relationship.* The dependency between OD flows and traffic measurements is a critical issue for the estimation of dynamic OD matrices in congested networks. The availability of emerging data sources that can provide more information regarding the prevailing traffic conditions is very important and relevant for the problem. Hence, data availability and adequate algorithms to accommodate any traffic information is a relevant research direction for WP4.
3. *Structural information* of OD flows. The mobility patterns, hence, demand patterns, fluctuate depending on several factors (e.g. seasonal conditions, recurrent and nonrecurrent events, etc.). It is important to be able to identify the variability in the demand structure in order to better estimate and predict the demand in a day-to-day and within-day framework.

Challenges	Contributions WP4
Use information from demand side in traffic supply prediction	Connection between supply and demand
Data resolution, aggregation and quality (limited period of time, no cleaned datasets)	Determine degree of aggregation, check data availability and quality in collaboration with WP2
Without simulation not possible to give predictions on new traffic modes	Hybrid approaches that combine the strengths of forecasting and simulation such as using simulated data to train forecasting algorithm on

Challenges	Contributions WP4
	new traffic modes for which there is no historic data available yet (such as CAVs)
Event detection	Inclusion of event data into spatio-temporal graph convolutional neural networks
Transfer learning	Examine the effectiveness of transfer learning when not enough traffic data is available for a certain city
Limited observability of OD matrices due to lack of adequate data to infer the demand, especially in congested networks	Exploitation of traditional and emerging data sources for estimating the demand to improve the accuracy of the estimated OD matrices
Inference of OD matrices using data-driven as well as hybrid approaches (metamodels)	Investigation and extension of methodologies to estimate and predict multimodal OD matrices based on available data combining simulation-based with analytical and data-driven approaches
Decrease of computational effort for demand estimation and prediction	Investigation of techniques (e.g., cloud computing) to improve the computational cost

Table 1: Challenges and contributions of WP4 of the Tangent project

6 Guidelines for Developing robust Prediction Models

Traffic forecasting is a challenging task that incorporates numerous parameters that must be determined, such as the modeling strategy, data collection, preprocessing, the generalization of the results and the implementation in real-world conditions. In this chapter, we propose a 10-step approach for researchers and practitioners in order to develop accurate and robust prediction models. The following points are, more or less, always considered before developing a prediction model.

6.1 Step 1: Define scope and usage

There are different requirements, for different forecasting applications, in terms of modeling techniques, data, etc. In the previous chapters, the differences between demand and supply side predictions have been extensively discussed. As a first step, the practitioner should decide the scope of the model to be developed and distinguish between supply and demand side prediction, e.g. onset of congestion prediction or origin-destination information predictions respectively, or follow a combining approach.

Furthermore, they should determine the application in which the model is going to be exploited and its exact requirements. Different types of use, such as real-time predictive management in different levels, i.e. operational, tactical and strategic policy making, may also involve very different approaches in terms of modeling and data acquisition. Furthermore, the scope and usage of the model will dictate the selection of the parameters to be predicted, the spatio-temporal extent of the predictions, as well as the modeling tools to be used.

6.2 Step 2: Local or Network level

Traffic forecasting refers to either local predictions or network-wide predictions. Local prediction is the prediction of the future conditions at a single point/road section where a loop detector is usually installed, using as input historical data from the same detector or/and several correlated detectors. On the other hand, network-wide predictions integrate multiple input and output variables, which correspond to points or road sections of an entire road network. The wider the spatial coverage of the model, the more complex the spatial and temporal relations that emerge.

The separation between the two types of predictions will also determine the amount of data that is required, i.e. for network-level predictions a larger amount of data is necessary, as well as more data sources (more loop detectors and GPS-enabled devices). The latter also leads to the deployment of different modeling techniques; Deep learning structures are more suitable for network-wide predictions, as they can handle efficiently the higher dimensionality of the input and output spaces, as well as the dynamic and non-linear dependencies between the traffic conditions of different sections of the road network, but can as well be exploited for local, single point time series predictions. Classic statistical and machine learning models can be used efficiently for local predictions. The practitioner should find a balance between the model's complexity and performance, especially when using complex deep learning architectures.

6.3 Step 3: Prediction Horizon

Another critical consideration for developing a traffic prediction model is the desirable prediction horizon. A horizon from one or more hours up to some days is considered a long-term one, while short-term refers to a horizon from a few minutes up to an hour. In recent literature, short-term forecasting with a horizon of 5 minutes or less is becoming more and more popular, as it is considered a more challenging task with higher research interest.

In real-world applications, the choice of the predicting horizon mostly depends on the application that the model will be used for. Short-term prediction models are suitable for real-time decision making and traffic management, travel time estimation and route planning, while long-term predictions can be exploited for policy-making and planning. The practitioner, therefore, should consider the overall aim of the model and the area of application.

In each of the two cases, data of the corresponding resolution are needed as well, i.e. high-resolution data for short-term forecasting, while data of lower resolution can be used for long-term forecasting. Consequently, the prediction horizon affects the data requirements, which are discussed in the next paragraph.

6.4 Step 4: Define data needs

Data are nowadays a valuable asset for the scientific community, especially for researchers that are involved in model development. The rapid evolution of computing and telecommunication systems, as well as connected smart devices and the Internet of Things, have led to the Big Data era, where vast amounts of data are being collected and made available.

Traffic data can be provided by loop detectors that are installed in numerous road networks or by the GPS signals of vehicles or smart devices of drivers. Although the number of available datasets is always increasing, still no such data are available for a significant number of cities. Therefore, the first limitation to consider is related to the availability of the required data, and, more specifically, the type of data that are needed, their granularity and their quantity.

A network-level prediction model would require data from more sources in the road network (multiple loop detectors), compared to local predictions. Consequently, when a network-level approach is going to be followed, network-level data should also be available. Furthermore, when a short-term prediction is requested, data of high temporal resolution are needed, while for long-term predictions lower resolution data can also be exploited. As far as it concerns the scope of the prediction, supply and demand side prediction models have significant differences in the data needs: for supply prediction, traffic flow or speed data from various points of the road network are required, while for the demand prediction origin-destination data of vehicles (or passengers, bikes, etc.) are needed or number of passenger/ticket validations (for public transport).

Consequently, especially when addressing the demand side, multimodal and multisource data may be necessary, e.g. traffic flow or origin-destination and transit demand data. This kind of heterogeneous data also require special handling and data fusion techniques in order to be exploited for the same model.

Lastly, it is widely understood that complex deep learning structures require huge amounts of data to train, to achieve stable and decent performance. As such, not only the modeling technique defines the data needs, but also data availability determines if certain modeling techniques can be implemented.

6.5 Step 5: Modeling strategy

The modeling strategy has already been mentioned a lot in the previous paragraphs and sections. There are three main categories of modeling techniques that are applied in traffic forecasting: statistical modeling, machine learning, deep learning, simulation-based or even hybrid models. The main differences between them are that deep learning and, secondarily, machine learning models require larger amounts of data and have been observed to have more accurate performance when the data needs are satisfied, but lack the interpretability of statistical models and demand more computational resources and time to train.

Therefore, unless a sufficiently big dataset is available, the practitioner should better avoid developing a deep learning model, especially a very complex one. Also, in real-world applications, if the corresponding authority that will use the model (e.g. road network operator) lacks the computational resources needed to train and maintain the model, complex deep learning models should be avoided. The same applies when the model should operate (get new input data, update-fine tuning, predict) in real-time. In the latter case, simulation-based prediction models should also be ruled out, as their operation may be very time-consuming.

On the other hand, deep learning is superior when making a network-wide prediction using high dimensional input and output space. Different deep learning models can extract the spatial and temporal relations of the road network and model its complex and non-linear mechanisms. Finally, their input data can be represented in more complicated structures such as images or graphs, which are more meaningful and valid, especially for a road network.

Moreover, a sophisticated feature selection strategy is encouraged, in order to reduce the dimensionality of the input space and the overall complexity of the modeling approach, by exploiting the most relevant and correlated input variables. In contrast, a “black-box” approach including all the available inputs, may prevent the convergence of the model. Another way to reduce the complexity of the model is to discretize the output of the model, which usually takes continuous values (e.g. speed, traffic flow), into a few categories. The categories may correspond to free flow, moderate, congested, etc. traffic conditions. The latter depends on the type of output variable values required to fulfill the model’s aim.

Instead of simplifying the input space, model compression techniques (such as pruning, knowledge distillation and quantization) are lately applied in order to reduce the complexity of a Deep Learning model and enhance the probability of operating in real-time. The main idea behind these algorithms varies, but in general they aim to reduce the size of the model by removing insignificant neurons and connections or by replacing some layers or even the entire neural network with linear filters or a simpler network.

As far as it concerns predictive traffic simulation, it can be performed either in macro or micro level, based on the expected outcome. Macro level simulation is suitable for predicting traffic flow levels, delays or average speed of road section in a network level, while micro level simulation goes deeper in a more personalized, agent-based level, allowing, for example, the estimation of travel times of individual users.

The practitioner should consider all the above parameters before deciding on a modeling strategy to follow.

6.6 Step 6: Selection tools

After determining the above specifications, the practitioner proceeds to implement the designed strategy. One of the first steps is to select the tools to use for the model development, usually including a programming language (or other statistical software) and a simulation platform. Python has been the most popular programming language in the last decade, especially for deep learning applications. It provides plenty of relevant libraries and packages that increase functionality and are easy to use. Other programming languages that still maintain a reasonable number of users and are also used for modeling are C and C++, Java, MATLAB and R. The latter is superior to the others when performing statistical analysis or developing statistical models.

6.7 Step 7: Evaluation strategy

Along with the model development, the practitioner should decide on an evaluation strategy for its results. As observed in recent literature, there are various criteria to evaluate a model. First of all, an error metric is used to compare the actual and the predicted values. Among the different existing metrics, the Mean Absolute Error (MAE) is the most popular, because it gives a straightforward estimate of the accuracy of the model, as it has the same units of measurement as the actual value. Moreover, the Mean Squared Error (MSE), which is also very frequently used, sets a higher penalty to instances with extremely high error values and Mean Absolute Percentage Error (MAPE) is used to compare with models fitted on other datasets.

Except for the accuracy of the prediction, the practitioner should also assess the following properties of the model: complexity, how time-consuming its training process is and in which level the results are interpretable and transferable. These properties are very often mentioned in recent literature and are sometimes as important as the absolute value of the error. Therefore, the practitioner should consider the above factors as well.

6.8 Step 8: Fine-tuning

When developing a model, the fine-tuning process, i.e. optimizing the values of its hyperparameters, is crucial for the model’s performance and takes place simultaneously with the model’s training phase. When developing deep learning models, this process may be difficult to converge and time-consuming, as the number of the involved hyperparameters may be extremely high. Also, big amounts of training data are needed in this case.

Similarly, in traffic simulation, before running an instance of a simulation-based prediction model, all the parameters that are integrated into the road network model should be carefully calibrated, so that the latter is representative of the real road network.

6.9 Step 9: Generalization

Generalization has arisen as a major issue of traffic conditions modeling and still remains a quite under-researched area. The effort, time and resources (data etc.) needed to develop a new prediction model, especially a deep learning one, makes it worth for practitioners and researchers to consider and evaluate the possibility of using (or "transferring") an already developed model for predicting the traffic conditions at a point or road network. The concept of transferability in deep learning is usually exploited to cope with the lack of relevant data, e.g. at a road network or section, by using a trained model of a similar road network or section, respectively.

Especially when modeling the demand side of traffic, the concept of transferability can be extended to transferring models across different modes, i.e. converting a model developed for predicting the demand (origin-destination matrix) for private vehicles for public transportation or taxis when those modes do not have available data. After a short fine-tuning process, it is possible that such a model can be transferred, maintaining a decent accuracy. Given the significant importance and the utility of generalization, the practitioner should examine if it's possible.

6.10 Step 10: Implementation (move to production)

Lately, there are several researchers that, without underestimating the effectiveness and the potential of deep learning models, are against its overwhelming use in traffic forecasting. They argue that such complex structures are not efficient for real-world exploitation and that practitioners should focus on less computationally expensive and more interpretable models.

More specifically, when they must be applied in real-world conditions, the models should maintain some specific properties. Firstly, a very complex model that demands high computational resources is not usually suitable for real-world problems. Thus, the practitioner should at least ensure that the aforementioned resources, as well as the required amount of data, are available. This applies to both deep learning and simulation-based models. These models' fitting and operation are also time-consuming, which limits the chances of their deployment in real-time applications.

Moreover, a model that is interpretable and reveals the spatio-temporal relations between road sections can be used for policy- and decision-making, e.g. in case a section closes because of an accident, the administrator would be able to identify the road sections that are going to be immediately impacted and take measures. Most deep learning models are non-interpretable and provide very accurate predictions, indeed, but in normal traffic conditions.

Lastly, the practitioner should ensure a stable data flow within the appropriate time period (depending on the specific application) and the validity and accuracy of the data collection process to use the corresponding model in real-world conditions.

7 Conclusions

As a Research and Innovation Action, TANGENT will address several advances beyond the state-of-the-art. One addressed topic is the field of traffic predictions and simulations. To address advances beyond the state-of-the-art, this report includes an analysis of the current state-of-the-art.

The methods for traffic forecasting vary from parametric-based and data-driven approaches (statistical and machine learning approaches), over simulation-based methods, to hybrid models, in which the latter combines different methods. Furthermore, the supply-side focuses on local/network-wide predictions and traffic anomalies as well. With respect to the demand-side the literature review focused on demand estimation and prediction methods for offline and online applications. The methods are further categorized based on the mathematical problem formulation of the state-of-the-art algorithms as well as the type of traffic information that the different methods can incorporate. Approaches that tackle the challenges associated with the demand estimation/prediction problem. The deliverable summarizes the advantages and limitations of each methodology and highlights the research gaps to be considered in the enhancement and development of adequate methods for travel demand estimation and prediction. Furthermore, 10 steps are proposed for researchers and practitioners in order to develop accurate and robust prediction models: define scope, usage, requirements for the forecasting applications; determine whether they are performed on a local or network-wide level; define the desirable prediction horizon (short-term versus long-term prediction); define the data needs that are necessary to perform the predictions; define the modeling strategy from the different available methods; determine the tools needed for the model development; decide on the evaluation strategy for its results; fine-tune the model (optimize and calibrate parameters); consider generalization concerning the use of an already developed model; move to real-world exploitation by ensuring the previous steps.

The results of the exploration in this report will provide the base for the development of the techniques in the remaining tasks of the WP, namely the development of the traffic supply and demand estimation and prediction approaches, the identification of critical conditions and congestion duration and in the end the development of a framework for real-time traffic monitoring and forecasting. The 10 proposed steps will help us in the development of these tasks.

The approaches outlined in section 3 will be investigated for the development of traffic supply forecasts (Task 4.2). They will be assessed/benchmarked in terms of accuracy, reliability of the outcomes and the computational load (uncertainty analysis/sensitivity analysis of the model scalability).

The work for the development of travel demand estimation and predictions (Task 4.3) will focus on the use of traditional and emerging data to explore and extend the state-of-the-art in methods and techniques to estimate and predict the demand variability in day-to-day and within day scenarios. Furthermore, this task will explore the validity of models and data to estimate and predict multimodal demand. These approaches will be assessed in terms of accuracy, reliability of the outcomes and the computational load for real-time traffic management needs.

The tasks regarding critical conditions and real-time forecasting use the previous established approaches for the detection and impact analysis of various events and the integration of the solutions in a general framework for traffic management.

8 References

1. Alsolami, B., Mehmood, R., & Albeshri, A. (2020). Hybrid statistical and machine learning methods for road traffic prediction: A review and tutorial. EAI/Springer Innovations in Communication and Computing, pp. 115–133.
2. Angarita-Zapata, J. S., Masegosa, A. D., & Triguero, I. (2019). A Taxonomy of Traffic Forecasting Regression Problems From a Supervised Learning Perspective. IEEE Access, 7, pp. 68185–68205.
3. Antoniou, C., Azevedo, C. L., Lu, L., Pereira, F., & Ben-Akiva, M. (2015). W–SPSA in Practice: Approximation of Weight Matrices and Calibration of Traffic Simulation Models. Transportation Research Procedia, 7, pp. 233–253.
4. Antoniou, C., Barceló, J., Breen, M., Bullejos, M., Casas, J., Cipriani, E., Ciuffo, B., Djukic, T., Hoogendoorn, S., Marzano, V., Montero, L., Nigro, M., Perarnau, J., Punzo, V., Toledo, T., & van Lint, H. (2016). Towards a generic benchmarking platform for origin–destination flows estimation/updating algorithms: Design, demonstration and validation. Transportation Research Part C: Emerging Technologies, 66, pp. 79–98.
5. Antoniou, C., Ben-Akiva, M., & Koutsopoulos, H. N. (2004). Incorporating Automated Vehicle Identification Data into Origin-Destination Estimation. Transportation Research Record: Journal of the Transportation Research Board, 1882(1), pp. 37–44.
6. Antoniou, C., Ben-Akiva, M., & Koutsopoulos, H. N. (2007). Nonlinear Kalman Filtering Algorithms for On-Line Calibration of Dynamic Traffic Assignment Models. IEEE Transactions on Intelligent Transportation Systems, 8(4), pp. 661–670.
7. Ashok, K., & Ben-Akiva, M. E. (2000). Alternative Approaches for Real-Time Estimation and Prediction of Time-Dependent Origin–Destination Flows. Https://Doi.Org/10.1287/Trsc.34.1.21.12282, 34(1), pp. 21–36.
8. Balakrishna, R., Antoniou, C., Ben-Akiva, M., Koutsopoulos, H. N., & Wen, Y. (2007). Calibration of Microscopic Traffic Simulation Models: Methods and Application. Https://Doi.Org/10.3141/1999-21, 1999, pp. 198–207.
9. Balakrishna, R., Ben-Akiva, M., & Koutsopoulos, H. N. (2007). Offline Calibration of Dynamic Traffic Assignment: Simultaneous Demand-and-Supply Estimation. Https://Doi.Org/10.3141/2003-07, 2003, pp. 50–58.
10. Balakrishna, R., & Koutsopoulos, H. N. (2008). Incorporating within-Day Transitions in Simultaneous Offline Estimation of Dynamic Origin-Destination Flows without Assignment Matrices: Https://Doi.Org/10.3141/2085-04, 2085, pp. 31–38.
11. Balakrishna, R., Koutsopoulos, H. N., & Ben-Akiva, M. (2005). Calibration and Validation of Dynamic Traffic Assignment Systems. Transportation and Traffic Theory. Flow, Dynamics and Human Interaction. 16th International Symposium on Transportation and Traffic Theory, pp. 407–426.
12. Balakrishna, R., Wen, Y., Ben-Akiva, M., & Antoniou, C. (2008). Simulation-Based Framework for Transportation Network Management in Emergencies. Transportation Research Record: Journal of the Transportation Research Board, 2041(1), pp. 80–88.
13. Barceló, J., & Montero, L. (2015). A Robust Framework for the Estimation of Dynamic OD Trip Matrices for Reliable Traffic Management. Transportation Research Procedia, 10, pp. 134–144.
14. Beckmann, M., Charles B. McGuire, & Christopher B. Winsten. (1956). Studies in the Economics of Transportation. Studies in the Economics of Transportation. https://trid.trb.org/view/91120
15. Ben-Akiva, M., Bierlaire, M., Burton, D., Koutsopoulos, H. N., & Mishalani, R. (2001). Network state estimation and prediction for real-time traffic management. Springer, 1(3/4), pp. 293–318.
16. Bogaerts, T. (2019). Temporal Graph Convolutional Deep Neural Networks for Traffic Forecasting using GPS Data.

17. Bogaerts, T., Masegosa, A. D., Angarita-Zapata, J. S., Onieva, E., & Hellinckx, P. (2020). A graph CNN-LSTM neural network for short and long-term traffic forecasting based on trajectory data. Transportation Research Part C: Emerging Technologies, 112, pp. 62–77.
18. Boquet, G., Morell, A., Serrano, J., & Vicario, J. L. (2020). A variational autoencoder solution for road traffic forecasting systems: Missing data imputation, dimension reduction, model selection and anomaly detection. Transportation Research Part C: Emerging Technologies, 115(March), pp. 102622.
19. Breunig, M. M., Kriegel, H.-P., Ng, R. T., & Sander, J. (2000). LOF: identifying density-based local outliers. Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data - SIGMOD '00, pp. 93–104.
20. Breyer, N., Gundlegard, D., Rydergren, C., & Backman, J. (2017). Trip extraction for traffic analysis using cellular network data. 5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems, MT-ITS 2017 - Proceedings, pp. 321–326.
21. Cai, L., Janowicz, K., Mai, G., Yan, B., & Zhu, R. (2020). Traffic transformer: Capturing the continuity and periodicity of time series for traffic forecasting. Transactions in GIS, 24(3), pp. 736–755.
22. 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, pp. 21–34.
23. Cantelmo, G., & Viti, F. (2020). A big data demand estimation model for urban congested networks. Transport and Telecommunication, 21(4), pp. 245–254.
24. Cantelmo, G., Viti, F., Tampère, C. M. J., Cipriani, E., & Nigro, M. (2014). Two-Step Approach for Correction of Seed Matrix in Dynamic Demand Estimation. Transportation Research Record: Journal of the Transportation Research Board, 2466(1), pp. 125–133.
25. Carrese, S., Cipriani, E., Mannini, L., & Nigro, M. (2017). Dynamic demand estimation and prediction for traffic urban networks adopting new data sources. Transportation Research Part C: Emerging Technologies, 81, pp. 83–98.
26. Cascetta, E., & Marquis, G. (1993). Dynamic Estimators of Origin-Destination Matrices Using Traffic Counts. Https://Doi.Org/10.1287/Trsc.27.4.363, 27(4), pp. 363–373.
27. Cascetta, E., & Postorino, M. N. (2001). Fixed Point Approaches to the Estimation of O/D Matrices Using Traffic Counts on Congested Networks. Https://Doi.Org/10.1287/Trsc.35.2.134.10138, 35(2), pp. 134–147.
28. Castiglione, M., Cantelmo, G., Qurashi, M., Nigro, M., & Antoniou, C. (2021). Assignment Matrix Free Algorithms for On-line Estimation of Dynamic Origin-Destination Matrices. Frontiers in Future Transportation, 2.
29. Cheng, X., Zhang, R., Zhou, J., & Xu, W. (2018). DeepTransport: Learning Spatial-Temporal Dependency for Traffic Condition Forecasting. Proceedings of the International Joint Conference on Neural Networks, 2018-July.
30. Chiu, Y. C., Zhou, L., & Song, H. (2010). Development and calibration of the Anisotropic Mesoscopic Simulation model for uninterrupted flow facilities. Transportation Research Part B: Methodological, 44(1), pp. 152–174.
31. 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(2), pp. 270–282.
32. Cipriani, E., Gemma, A., & Nigro, M. (2013). A bi-level gradient approximation method for dynamic traffic demand estimation: Sensitivity analysis and adaptive approach. 16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013), pp. 2100–2105.
33. 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(2), pp. 139–148.

34. Cong, Y., Wang, J., & Li, X. (2016). Traffic Flow Forecasting by a Least Squares Support Vector Machine with a Fruit Fly Optimization Algorithm. Procedia Engineering, 137, pp. 59–68.
35. Cui, Z., Ke, R., Pu, Z., & Wang, Y. (2018). Deep bidirectional and unidirectional LSTM recurrent neural network for network-wide traffic speed prediction. ArXiv Preprint ArXiv:1801.02143.
36. Dantsuji, T., Hoang, N.H., Zheng, N., & Vu H.L. (2022). A novel metamodel-based framework for large-scale dynamic origin–destination demand calibration. Transportation Research Part C: Emerging Technologies, 136:103545.
37. Deng, L., Lian, D., Huang, Z., & Chen, E. (2022). Graph Convolutional Adversarial Networks for Spatiotemporal Anomaly Detection. IEEE Transactions on Neural Networks and Learning Systems, pp. 1–13.
38. Deng, S., Hu, J., Wang, Y., & Zhang, Y. (2009). Urban road network modeling and real-time prediction based on householder transformation and adjacent vector. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 5553 LNCS(PART 3), pp. 899–908.
39. Djenouri, Y., Belhadi, A., Lin, J. C.-W., Djenouri, D., & Cano, A. (2019). A Survey on Urban Traffic Anomalies Detection Algorithms. IEEE Access, 7, pp. 12192–12205.
40. Djukic, T. (2014). Dynamic OD Demand Estimation and Prediction for Dynamic Traffic Management.
41. Djukic, T., Barceló Bugeda, J., Bullejos, M., Montero Mercadé, L., Cipriani, E., van Lint, H., & Hoogendoorn, S. (2014). Advanced traffic data for dynamic OD demand estimation: The state of the art and benchmark study. Transportation Research Board 94th Annual. https://upcommons.upc.edu/handle/2117/79383
42. Djukic, T., Flotterod, G., van Lint, H., & Hoogendoorn, S. (2012). Efficient real time OD matrix estimation based on Principal Component Analysis. 2012 15th International IEEE Conference on Intelligent Transportation Systems, pp. 115–121.
43. Do, L. N. N., Taherifar, N., & Vu, H. L. (2019). Survey of neural network-based models for short-term traffic state prediction. WIREs Data Mining and Knowledge Discovery, 9(1).
44. Do, L. N. N., Vu, H. L., Vo, B. Q., Liu, Z., & Phung, D. (2019). An effective spatial-temporal attention based neural network for traffic flow prediction. Transportation Research Part C: Emerging Technologies, 108(August), pp. 12–28.
45. Duan, P., Mao, G., Zhang, C., & Wang, S. (2016). STARIMA-based traffic prediction with time-varying lags. IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC, 1, pp. 1610–1615.
46. Eisenman, S. M., & List, G. F. (2004). Using probe data to estimate OD matrices. IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC, pp. 291–296.
47. Ermagun, A., & Levinson, D. (2018). Spatiotemporal traffic forecasting: review and proposed directions. Transport Reviews, 38(6), pp. 786–814.
48. Fang, W., Zhuo, W., Yan, J., Song, Y., Jiang, D., & Zhou, T. (2022). Attention meets long short-term memory: A deep learning network for traffic flow forecasting. Physica A: Statistical Mechanics and Its Applications, 587, pp. 126485.
49. 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, 2263, pp. 19–25.
50. Fu, J., Zhou, W., & Chen, Z. (2020). Bayesian Spatio-Temporal Graph Convolutional Network for Traffic Forecasting. http://arxiv.org/abs/2010.07498
51. Ghosh, B., Basu, B., & O'Mahony, M. (2007). Bayesian Time-Series Model for Short-Term Traffic Flow Forecasting. Journal of Transportation Engineering, 133(3), pp. 180–189.
52. Guo, F., Polak, J. W., & Krishnan, R. (2018). Predictor fusion for short-term traffic forecasting. Transportation Research Part C: Emerging Technologies, 92, pp. 90–100.

53. 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, pp. 50–64.
54. Guo, S., Lin, Y., Li, S., Chen, Z., & Wan, H. (2019). Deep Spatialoral 3D Convolutional Neural Networks for Traffic Data Forecasting. IEEE Transactions on Intelligent Transportation Systems, 20(10), pp. 3913–3926.
55. Hadjidimitriou, N. S., Lippi, M., & Mamei, M. (2021). A Data Driven Approach to Match Demand and Supply for Public Transport Planning. IEEE Transactions on Intelligent Transportation Systems, 22(10), pp. 6384–6394.
56. Hamed, M. M., Al-Masaeid, H. R., & Said, Z. M. B. (1995). Short-Term Prediction of Traffic Volume in Urban Arterials. Journal of Transportation Engineering-Asce, 121, pp. 249–254.
57. Hofleitner, A., Herring, R., & Bayen, A. (2012). Arterial travel time forecast with streaming data: A hybrid approach of flow modeling and machine learning. Transportation Research Part B: Methodological, 46(9), pp. 1097–1122.
58. Hong, W. C., Dong, Y., Zheng, F., & Lai, C. Y. (2011). Forecasting urban traffic flow by SVR with continuous ACO. Applied Mathematical Modelling, 35(3), pp. 1282–1291.
59. Hu, W., Yan, L., Liu, K., & Wang, H. (2016). A Short-term Traffic Flow Forecasting Method Based on the Hybrid PSO-SVR. Neural Processing Letters, 43(1), pp. 155–172.
60. Huang, G. L., Deng, K., & He, J. (2020). Cognitive Traffic Anomaly Prediction from GPS Trajectories Using Visible Outlier Indexes and Meshed Spatiotemporal Neighborhoods. Cognitive Computation, 12(5), pp. 967–978.
61. Jason Brownlee. (2016). Parametric and Nonparametric Machine Learning Algorithms. https://machinelearningmastery.com/parametric-and-nonparametric-machine-learning-algorithms/
62. Jiang, W., & Luo, J. (2021). Graph Neural Network for Traffic Forecasting: A Survey. http://arxiv.org/abs/2101.11174
63. Jin, W., Lin, Y., Wu, Z., & Wan, H. (2018). Spatio-Temporal Recurrent Convolutional Networks for Citywide Short-term Crowd Flows Prediction. Proceedings of the 2nd International Conference on Compute and Data Analysis - ICCDA 2018, pp. 28–35.
64. Kalman, R. E., & Bucy, R. S. (1961). New Results in Linear Filtering and Prediction Theory. Journal of Basic Engineering, 83(1), pp. 95–108.
65. Ke, J., Qin, X., Yang, H., Zheng, Z., Zhu, Z., & Ye, J. (2021). Predicting origin-destination ride-sourcing demand with a spatio-temporal encoder-decoder residual multi-graph convolutional network. Transportation Research Part C: Emerging Technologies, 122, pp. 102858.
66. Ke, J., Zheng, H., Yang, H., & Chen, X. M. (2017). Short-term forecasting of passenger demand under on-demand ride services: A spatio-temporal deep learning approach. Transportation Research Part C: Emerging Technologies, 85, pp. 591–608.
67. Kemper, C. (2005). Dynamic traffic flow model–a new approach with static data. Proceedings of the 5th European Congress and Exhibition on Intelligent Transport Systems (ITS), pp. 1–13.
68. Kerner, B. S., & Klenov, S. L. (2002). A microscopic model for phase transitions in traffic flow. Journal of Physics A: Mathematical and General, 35(3), pp. L31.
69. Kerner, B. S., Klenov, S. L., Hermanns, G., & Schreckenberg, M. (2013). Effect of driver over-acceleration on traffic breakdown in three-phase cellular automaton traffic flow models. Physica A: Statistical Mechanics and Its Applications, 392(18), pp. 4083–4105.
70. Kipf, T. N., & Welling, M. (2017). Semi-Supervised Classification with Graph Convolutional Networks. pp. 1–14.
71. Kong, X., Gao, H., Alfarraj, O., Ni, Q., Zheng, C., & Shen, G. (2020). HUAD: Hierarchical Urban Anomaly Detection Based on Spatio-Temporal Data. IEEE Access, 8, pp. 26573–26582.

72. Kostic, B., Gentile, G., & Antoniou, C. (2017). Techniques for improving the effectiveness of the SPSA algorithm in dynamic demand calibration. 5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems, MT-ITS 2017 - Proceedings, pp. 368–373.
73. Krishnakumari, P., van Lint, H., Djukic, T., & Cats, O. (2019). A data driven method for OD matrix estimation. Transportation Research Procedia, 38, pp. 139–159.
74. Krishnakumari, P., van Lint, H., Djukic, T., & Cats, O. (2020). A data driven method for OD matrix estimation. Transportation Research Part C: Emerging Technologies, 113(May 2019), pp. 38–56.
75. Lam, P., Wang, L., Ngan, H. Y. T., Yung, N. H. C., & Yeh, A. G. O. (2017). Outlier detection in large-scale traffic data by Naïve Bayes method and Gaussian mixture model method. IS and T International Symposium on Electronic Imaging Science and Technology, pp. 73–78.
76. Lana, I., del Ser, J., Velez, M., & Vlahogianni, E. I. (2018). Road Traffic Forecasting: Recent Advances and New Challenges. IEEE Intelligent Transportation Systems Magazine, 10(2), pp. 93–109.
77. Laña, I., Lobo, J. L., Capecci, E., del Ser, J., & Kasabov, N. (2019). Adaptive long-term traffic state estimation with evolving spiking neural networks. Transportation Research Part C: Emerging Technologies, 101(January), pp. 126–144.
78. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), pp. 436–444.
79. Lee, K., & Rhee, W. (2022). DDP-GCN: Multi-graph convolutional network for spatiotemporal traffic forecasting. Transportation Research Part C: Emerging Technologies, 134, pp. 103466.
80. Li, C., Bai, L., Liu, W., Yao, L., & Waller, S. T. (2020). Graph Neural Network for Robust Public Transit Demand Prediction. IEEE Transactions on Intelligent Transportation Systems.
81. Li, M., & Zhu, Z. (2021). Spatial-Temporal Fusion Graph Neural Networks for Traffic Flow Forecasting. Proceedings of the AAAI Conference on Artificial Intelligence, 35(5), pp. 4189–4196.
82. Li, Y., Yu, R., Shahabi, C., & Liu, Y. (2018). Diffusion convolutional recurrent neural network: Data-driven traffic forecasting. 6th International Conference on Learning Representations, ICLR 2018 - Conference Track Proceedings, pp. 1–16.
83. Liao, B., McIlwraith, D., Zhang, J., Chen, T., Wu, C., Yang, S., Guo, Y., & Wu, F. (2018). Deep sequence learning with auxiliary information for traffic prediction. Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 537–546.
84. Liu, B., Long, J., Deng, M., Tang, J., & Huang, J. (2022). Revealing spatiotemporal correlation of urban roads via traffic perturbation simulation. Sustainable Cities and Society, 77, pp. 103545.
85. Liu, H., van Zuylen, H. J., van Lint, H., Chen, Y., & Zhang, K. (2005). Prediction of urban travel times with intersection delays. IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC, 2005(8), pp. 402–407.
86. Liu, J., Zheng, F., van Zuylen, H. J., & Li, J. (2020). A dynamic OD prediction approach for urban networks based on automatic number plate recognition data. Transportation Research Procedia, 47, pp. 601–608.
87. Liu, Y., & Wu, H. (2017). Prediction of Road Traffic Congestion Based on Random Forest. 2017 10th International Symposium on Computational Intelligence and Design (ISCID), pp. 361–364.
88. Lu, L., Xu, Y., Antoniou, C., & Ben-Akiva, M. (2015). An enhanced SPSA algorithm for the calibration of Dynamic Traffic Assignment models. Transportation Research Part C: Emerging Technologies, 51, pp. 149–166.
89. Luo, X., Li, D., Yang, Y., & Zhang, S. (2019). Spatiotemporal traffic flow prediction with KNN and LSTM. Journal of Advanced Transportation, 2019.
90. Ma, T., Antoniou, C., & Toledo, T. (2020). Hybrid machine learning algorithm and statistical time series model for network-wide traffic forecast. Transportation Research Part C: Emerging Technologies, 111, pp. 352–372.

91. Ma, X., Dai, Z., He, Z., Ma, J., Wang, Y., & Wang, Y. (2017). Learning traffic as images: A deep convolutional neural network for large-scale transportation network speed prediction. Sensors (Switzerland), 17(4).
92. Mahmassani, H. S. (2001). Dynamic Network Traffic Assignment and Simulation Methodology for Advanced System Management Applications. Networks and Spatial Economics 2001 1:3, 1(3), pp. 267–292.
93. Mallick, T., Balaprakash, P., Rask, E., & Macfarlane, J. (2021). Transfer learning with graph neural networks for short-term highway traffic forecasting. 2020 25th International Conference on Pattern Recognition (ICPR), pp. 10367–10374.
94. Min, X., Hu, J., Chen, Q., Zhang, T., & Zhang, Y. (2009). Short-term traffic flow forecasting of urban network based on dynamic STARIMA model. IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC, 100084, pp. 461–466.
95. Miska, M. P. (2007). Microscopic online simulation for real time traffic management. In Transport and Planning : Vol. Doctoral (Issue February).
96. Nagy, A. M., & Simon, V. (2018). Survey on traffic prediction in smart cities. Pervasive and Mobile Computing, 50, pp. 148–163.
97. Ngoduy, D., & Wilson, R. E. (2014). Multianticipative nonlocal macroscopic traffic model. Computer-Aided Civil and Infrastructure Engineering, 29(4), pp. 248–263.
98. Osorio, C. (2019a). High-dimensional offline origin-destination (OD) demand calibration for stochastic traffic simulators of large-scale road networks. Transportation Research Part B: Methodological, 124, pp. 18–43.
99. Osorio, C. (2019b). Dynamic origin-destination matrix calibration for large-scale network simulators. Transportation Research Part C: Emerging Technologies, 98, pp. 186–206.
100. Pfeifer, P. E., & Deutsch, S. J. (1980). A STARIMA model-building procedure with application to description and regional forecasting. Transactions, Institute of British Geographers, 5(3), pp. 330–349.
101. Prakash, A. A., Seshadri, R., Antoniou, C., Pereira, F. C., & Ben-Akiva, M. (2018). Improving Scalability of Generic Online Calibration for Real-Time Dynamic Traffic Assignment Systems: Https://Doi.Org/10.1177/0361198118791360, 2672(48), pp. 79–92.
102. Prakash, A. A., Seshadri, R., Antoniou, C., Pereira, F. C., & Ben-Akiva, M. E. (2017). Reducing the Dimension of Online Calibration in Dynamic Traffic Assignment Systems: Https://Doi.Org/10.3141/2667-10, 2667(1), pp. 96–107.
103. Qi, Y., & Ishak, S. (2014). A Hidden Markov Model for short term prediction of traffic conditions on freeways. Transportation Research Part C: Emerging Technologies, 43, pp. 95–111.
104. Qurashi, M., Lu, Q.-L., Cantelmo, G., & Antoniou, C. (2022). Dynamic demand estimation on large scale networks using Principal Component Analysis: The case of non-existent or irrelevant historical estimates. Transportation Research Part C: Emerging Technologies, 136, pp. 103504.
105. Qurashi, M., Ma, T., Chaniotakis, E., & Antoniou, C. (2018). An Alternate Online Calibration approach for O-D demand Calibration in Dynamic Traffic Assignment Systems. Proceedings of 7th Symposium of the European Association for Research in Transportation (HEART).
106. Qurashi, M., Ma, T., Chaniotakis, E., & Antoniou, C. (2020). PC–SPSA: Employing Dimensionality Reduction to Limit SPSA Search Noise in DTA Model Calibration. IEEE Transactions on Intelligent Transportation Systems, 21(4), pp. 1635–1645.
107. Rousseeuw, P. J., & van Driessen, K. (1999). A fast algorithm for the minimum covariance determinant estimator. Technometrics, 41(3), pp. 212–223.
108. Russell, S. J., & Norvig, P. (2016). Artificial intelligence: a modern approach. Malaysia. pp. 1145.
109. Ryu, U., Wang, J., Kim, T., Kwak, S., & Juhyok, U. (2018). Construction of traffic state vector using mutual information for short-term traffic flow prediction. Transportation Research Part C: Emerging Technologies, 96(December 2017), pp. 55–71.

110. Safikhani, A., Kamga, C., Mudigonda, S., Faghih, S. S., & Moghimi, B. (2020). Spatio-temporal modeling of yellow taxi demands in New York City using generalized STAR models. International Journal of Forecasting, 36(3), pp. 1138–1148.
111. Schölkopf, B., Williamson, R., Smola, A., Shawe-Taylor, J., & Piatt, J. (2000). Support vector method for novelty detection. Advances in Neural Information Processing Systems, pp. 582–588.
112. Shao, Y., & Sun, Z. (2021). Energy-Efficient Connected and Automated Vehicles: Real-Time Traffic Prediction-Enabled Co-Optimization of Vehicle Motion and Powertrain Operation. IEEE Vehicular Technology Magazine, 16(3), pp. 47–56.
113. Smith, B. L., Williams, B. M., & Keith Oswald, R. (2002). Comparison of parametric and nonparametric models for traffic flow forecasting. Transportation Research Part C: Emerging Technologies, 10(4), pp. 303–321.
114. Sofuoglu, S. E., & Aviyente, S. (2022). GLOSS: Tensor-based anomaly detection in spatiotemporal urban traffic data. Signal Processing, 192, pp. 108370.
115. Song, Z., Guo, Y., Wu, Y., & Ma, J. (2018). Short-term traffic speed prediction under different data collection time intervals using a SARIMA-SDGM hybrid prediction model. PLoS ONE, 14(6), pp. 1–19.
116. Spiess, H. (1990). A GRADIENT APPROACH FOR THE O-D MATRIX ADJUSTMENT PROBLEM. Publication No. 693, Centre de Recherche Sur Les Transports, Université de Montréal.
117. Tampère, C., Viti, F., & Immers, B. (2010). The Effect of Dynamic Network Loading Models on DTA-based OD Estimation. https://www.researchgate.net/publication/228705272
118. Toledo, T., & Kolechkina, T. (2013). Estimation of Dynamic Origin–Destination Matrices Using Linear Assignment Matrix Approximations. IEEE Transactions on Intelligent Transportation Systems, 14(2), pp. 618–626.
119. Tympakianaki, A. (2018). Demand Estimation and Bottleneck Management Using Heterogeneous Traffic Data.
120. Tympakianaki, A., Koutsopoulos, H. N., & Jenelius, E. (2015). c-SPSA: Cluster-wise simultaneous perturbation stochastic approximation algorithm and its application to dynamic origin–destination matrix estimation. Transportation Research Part C: Emerging Technologies, 55, pp. 231–245.
121. Tympakianaki, A., Koutsopoulos, H. N., & Jenelius, E. (2018). Robust SPSA algorithms for dynamic OD matrix estimation. Procedia Computer Science, 130, pp. 57–64.
122. Vaswani, A., Shazeer, N., Parmer, N., Uszkoreitm, J., & Jones, L. (2017). Attention Is All You Need. Advances in Neural Information Processing Systems, 30(Nips). https://proceedings.neurips.cc/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html
123. Vaze, V., Antoniou, C., Wen, Y., & Ben-Akiva, M. (2009). Calibration of Dynamic Traffic Assignment Models with Point-to-Point Traffic Surveillance: Https://Doi.Org/10.3141/2090-01, 2090, pp. 1–9.
124. 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, pp. 3–19.
125. Wang, J., Deng, W., & Guo, Y. (2014). New Bayesian combination method for short-term traffic flow forecasting. Transportation Research Part C: Emerging Technologies, 43, pp. 79–94.
126. 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, pp. 219–232.
127. 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(2), pp. 141–167.

128. 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(6), pp. 664–672.
129. Xie, Y., Zhao, K., Sun, Y., & Chen, D. (2010). Gaussian processes for short-term traffic volume forecasting. Transportation Research Record, 2165, pp. 69–78.
130. Xiong, C., Zhu, Z., He, X., Chen, X., Zhu, S., Mahapatra, S., Chang, G. L., & Zhang, L. (2015). Developing a 24-hour large-scale microscopic traffic simulation model for the before-and-after study of a new tolled freeway in the Washington, DC-Baltimore region. Journal of Transportation Engineering, 141(6), pp. 1–11.
131. Yan, J., Zhang, J., Liu, Y., Lv, G., Han, S., & Alfonzo, I. E. G. (2020). EV charging load simulation and forecasting considering traffic jam and weather to support the integration of renewables and EVs. Renewable Energy, 159, pp. 623–641.
132. Yang, S. (2013). On feature selection for traffic congestion prediction. Transportation Research Part C: Emerging Technologies, 26, pp. 160–169.
133. Yao, H., Tang, X., Wei, H., Zheng, G., & Li, Z. (2019). Revisiting spatial-temporal similarity: A deep learning framework for traffic prediction. 33rd AAAI Conference on Artificial Intelligence, AAAI 2019, 31st Innovative Applications of Artificial Intelligence Conference, IAAI 2019 and the 9th AAAI Symposium on Educational Advances in Artificial Intelligence, EAAI 2019, pp. 5668–5675.
134. Yin, X., Wu, G., Wei, J., Shen, Y., Qi, H., & Yin, B. (2021). Multi-stage attention spatial-temporal graph networks for traffic prediction. Neurocomputing, 428, pp. 42–53.
135. Yu, B., Yin, H., & Zhu, Z. (2018). Spatio-Temporal Graph Convolutional Networks: A Deep Learning Framework for Traffic Forecasting. Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, pp. 3634–3640.
136. Yu, G., Hu, J., Zhang, C., Zhuang, L., & Song, J. (2003). Short-term traffic flow forecasting based on Markov chain model. IEEE Intelligent Vehicles Symposium, Proceedings, pp. 208–212.
137. Yu, H., Wu, Z., Wang, S., Wang, Y., & Ma, X. (2017). Spatiotemporal Recurrent Convolutional Networks for Traffic Prediction in Transportation Networks. Sensors, 17(7), pp. 1501.
138. Yu, J. J. Q., Markos, C., & Zhang, S. (2021). Long-Term Urban Traffic Speed Prediction With Deep Learning on Graphs. IEEE Transactions on Intelligent Transportation Systems, pp. 1–12.
139. Yuan, Y., van Lint, J. W. C., Wilson, R. E., van Wageningen-Kessels, F., & Hoogendoorn, S. P. (2012). Real-time lagrangian traffic state estimator for freeways. IEEE Transactions on Intelligent Transportation Systems, 13(1), pp. 59–70.
140. Yue, B., Shi, S., Wang, S., & Lin, N. (2020). Low-Cost Urban Test Scenario Generation Using Microscopic Traffic Simulation. IEEE Access, 8, pp. 123398–123407.
141. Zhang, K., Xie, C., Wang, Y., Ángel, S. M., Nguyen, T. M. T., Zhao, Q., & Li, Q. (2021). Hybrid short-term traffic forecasting architecture and mechanisms for reservation-based Cooperative ITS. Journal of Systems Architecture, 117, pp. 102101.
142. Zhang, Y., & Xie, Y. (2007). Forecasting of short-term freeway volume with v-support vector machines. Transportation Research Record, 2024, pp. 92–99.
143. Zhang, Z., Li, Y., Song, H., & Dong, H. (2021). Multiple dynamic graph based traffic speed prediction method. Neurocomputing, 461, pp. 109–117.
144. Zhao, J., & Sun, S. (2016). High-Order Gaussian Process Dynamical Models for Traffic Flow Prediction. IEEE Transactions on Intelligent Transportation Systems, 17(7), pp. 2014–2019.
145. Zhao, L., Song, Y., Zhang, C., Liu, Y., Wang, P., Lin, T., Deng, M., & Li, H. (2020). T-GCN: A Temporal Graph Convolutional Network for Traffic Prediction. IEEE Transactions on Intelligent Transportation Systems, 21(9), pp. 3848–3858.
146. Zheng, C., Fan, X., Wang, C., & Qi, J. (2020). Gman: A graph multi-attention network for traffic prediction. Proceedings of the AAAI Conference on Artificial Intelligence, 34.

147. Zheng, H., Lin, F., Feng, X., & Chen, Y. (2020). Attention-Based Conv-LSTM Networks for Short-Term Traffic Flow Prediction. IEEE Transactions on Intelligent Transportation Systems, pp. 1–11.
148. Zhou, X., & Mahmassani, H. S. (2006). Dynamic origin-destination demand estimation using automatic vehicle identification data. IEEE Transactions on Intelligent Transportation Systems, 7(1), pp. 105–114.
149. Zhu, H., Xie, Y., He, W., Sun, C., Zhu, K., Zhou, G., & Ma, N. (2020). A Novel Traffic Flow Forecasting Method Based on RNN-GCN and BRB. Journal of Advanced Transportation, 2020.
150. Zhu, J., Wang, Q., Tao, C., Deng, H., Zhao, L., & Li, H. (2021). AST-GCN: Attribute-augmented spatiotemporal graph convolutional network for traffic forecasting. IEEE Access, 9, pp. 35973–35983.