UNIVERSITY OF LEEDS

Uncertainties and Errors in Predicting Vehicle Exhaust Emissions using Traffic Flow Models

by

Arwa Sayegh

Submitted in accordance with the requirements for the degree of Doctor of Philosophy

in the

Faculty of Environment

Institute for Transport Studies

May 2017

Declaration of Authorship

The candidate confirms that the work submitted is her own, except where work which has formed part of a jointly-authored publication has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

The jointly-authored publication is "Sayegh, AS, Connors, RD and Tate, JE (*In press*) Uncertainty propagation from the cell transmission traffic flow model to emission predictions: a data-driven approach, *Transportation Science*, ." This forms part of my PhD work and contains elements from Chapters 2, 3, 4, and 5. I hereby declare I made the major contribution to the publication. The joint-authors were my PhD supervisors who offered advise at various stages of the research project.

to Karl, Richard, and
of course My Family

Acknowledgements

The work of this thesis[1] was carried out at the Institute for Transport Studies (ITS) of the University of Leeds, UK and was funded by the Leeds International Research Scholarship (LIRS). The primary source of data used was made accessible by Highways England. Some of the work was undertaken on ARC2, part of the High Performance Computing facilities at the University of Leeds, UK.

I would like to first thank my supervisor Richard Connors for all his support and guidance throughout the entire three years and a half. I find it a bit difficult to say how much the time you have spent listening to me, discussing things with me, and ensuring all PhD and post-PhD related aspects are going well, means to me. All I can say here really is thank you for everything.

My thanks are due to my supervisor James Tate for his support throughout the period of my PhD, and, of course, his support of my PhD and funding applications that allowed me to be here in the first place. My thanks are also due to my PhD examiners Francesco Viti and David Watling for their discussion and their input on the final submitted thesis. I would also like to thank Dong Ngoduy for always answering many of my questions on certain topics of this research - it really helped me a lot. Also, special thanks go to Deborah Goddard for her assistance on all PhD-related matters and of course, to Paul Adams - I will definitely miss our time-to-time conversations.

During my PhD, I had the chance to work on other projects with few who have taught me a lot and whom I would like to thank here: Andrew Evans, Roger Timmis, Duncan Whyatt, Laura Smith, and Maria Angeles Solera Garcia. Of course, I thank my dear friends Farah, Mojtaba, Jingyan, Wankun, Haneen, Tatjana, Sanna, Ersilia, Mike, and Mahmoud and my dear friends back home or outside the UK who have kept in touch no matter how far they are: Dima, Heba, Neam, Mona, Christian, Anthony, and Koya.

My mentor Karl Ropkins and my family, I thank you for always being there for me, no matter what, and I dedicate this work to you.

[1] LaTeX template of this thesis was made accessible by Sunil Patel.

UNIVERSITY OF LEEDS

Abstract

Faculty of Environment
Institute for Transport Studies

Doctor of Philosophy

by Arwa Sayegh

Vehicle exhaust emissions predicted based on the outputs of traffic flow models are used directly to calculate traffic-related emissions, but also indirectly as input to 'air quality - human exposure' models. Both of which inform transport and environmental policies aimed at achieving sustainable mobility. To be effective, these must be based on robust modelling approaches that not only provide point-based emission predictions, but also inform these with an interval of confidence that properly accounts for the propagation of uncertainties and errors through the complex chain of models involved. This research develops a data-driven methodological framework to probabilistic average speed-based emission predictions using two widely deployed macroscopic traffic flow models. These are the Cell Transmission Model (CTM), a discretised first-order LWR-type model, and METANET, a discretised second-order Payne-type model. Studying both allows quantitative comparison in their application to predicting emissions. While this research discusses all potential sources of uncertainty in this modelling chain, it focusses on those arising from the traffic flow modelling output. The methodology starts with an ensemble-based optimisation approach to estimate both calibration and validation prediction errors in the traffic flow model, and then proposes a Monte Carlo sampling approach to propagate these to emission predictions. This allows predicting emissions alongside their upper and lower bounds for any time period and road network, at different levels of detail. To ensure transferability of findings, this methodology has been tested on three motorway road networks, one of which operates under Variable Speed Limits (VSL). This permits the quantitative assessment of VSL-modified traffic flow models. In the results of this research, emissions of Oxides of Nitrogen (NO_x) and uncertainty associated with their prediction are specifically reported for each road network under study. Finally, this research argues that the methodological framework developed can (and should) be applied to any other (relatively) simple or complex integrated 'traffic flow - emission' modelling chain used as part of policy and decision making process.

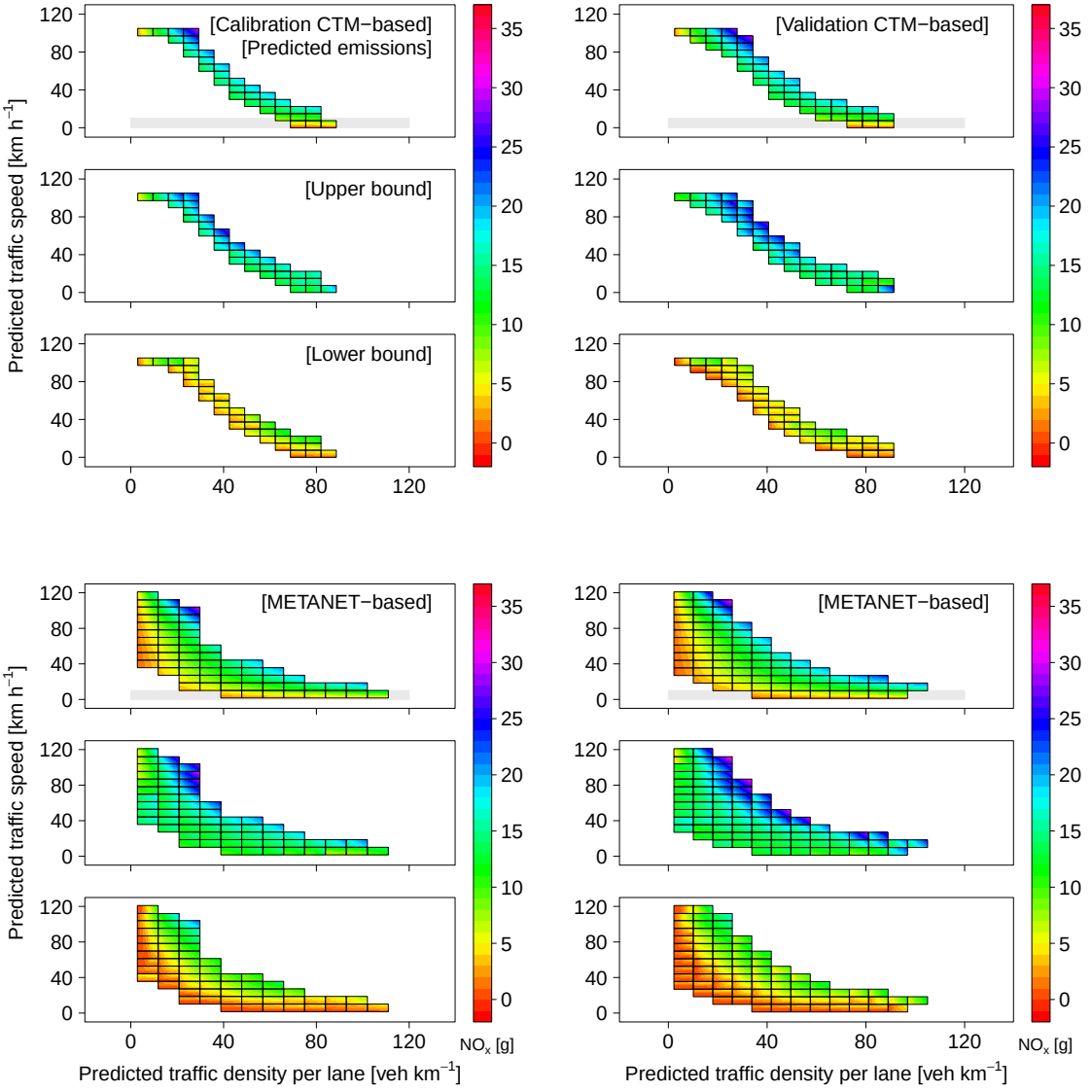
UNIVERSITY OF LEEDS

Graphical Abstract

Faculty of Environment
Institute for Transport Studies

Doctor of Philosophy

by Arwa Sayegh



Contents

List of Figures

List of Tables

Abbreviations

LWR	**L**ighthill **W**hitham **R**ichards.
CTM	**C**ell **T**ransmission **M**odel.
METANET	**M**odéle d'**É**coulement de **T**rafic sur **A**utoroute **NET**works.
VSL	**V**ariable **S**peed **L**imit.
SQP	**S**equential **Q**uadratic **P**rogramming.
N-M	**N**elder-**M**ead.
LS	**L**east **S**quares.
CEM	**C**ross **E**ntropy **M**ethod.
GA	**G**enetic **A**lgorithms.
CS	**C**uckoo **S**earch.
PSO	**P**article **S**warm **O**ptimisation.
SE	**S**quared **E**rror.
MSE	**M**ean **S**quared **E**rror.
RMSE	**R**oot **M**ean **S**quared **E**rror.
CO	**C**arbon **M**onoxide.
HC	**H**ydro**c**arbon.
NO_x	**Ox**ides of **N**itrogen.
PEMS	**P**ortable **E**mission **M**easurement **S**ystem.
RSD	**R**emote **S**ensing **D**evice.
NAEI	**N**ational **A**tmospheric **E**mission **I**nventory.
COPERT	**CO**mputer **P**rogramme to **C**alculate **E**missions from **R**oad **T**ransport.
DTA	**D**ynamic **T**raffic **A**ssignment.
EPA	**E**nvironment **P**rotection **A**gency.
CFL	**C**ourant-**F**riedrichs-**L**ewy.
PC	**P**assenger **C**ar.
LDV	**L**ight **D**uty **V**ehicle.
NSL	**N**ational **S**peed **L**imit.
MIDAS	**M**otorway **I**ncident **D**etection and **A**utomatic **S**ignalling.
RCC	**R**egional **C**ontrol **C**entre.
NTIS	**N**ational **T**raffic **I**nformation **S**ervice.

Symbols

Z	traffic variables	-
r	road network	-
R	total number of road networks	-
m	time period	-
M	total number of time periods	-
f	traffic flow model structure	-
I	traffic flow model input	-
β	traffic flow model parameter vector	-
q	traffic flow	veh h^{-1}
u	traffic speed	km h^{-1}
ρ	traffic density	veh km^{-1}
d	traffic demand	veh h^{-1}
θ	off-ramp split ratio	-
h	queue length	veh
λ	number of lanes	-
i	cell number	-
N	total number of cells	-
t	time step	-
T	total number of time steps	-
T_s	time step length	h
l	cell length	km
u_f	free-flow traffic speed	km h^{-1}
Q_{max}	maximum flow rate	veh h^{-1}
w	backward wave speed	km h^{-1}
ρ_{max}	maximum traffic density	veh km^{-1}
τ	relaxation time parameter	h
η	anticipation parameter	km^2 h^{-1}
κ	model parameter	veh km^{-1}
ϕ	lane-drop parameter	-
δ	on-ramp merging parameter	-
u_{min}	minimum traffic speed	km h^{-1}
$\underline{\rho}_c$	critical traffic density per lane	veh km^{-1}

α	fundamental diagram parameter	-
g	emission model structure	-
a	air pollutant	-
c	vehicle category	-
γ	proportion of vehicles	-
ef	emission factor	g km^{-1}
ω	emission factor's function parameter	-
e	optimisation run	-
E	total number of optimisation runs	-
χ	threshold percentile	-
ξ	significance level	-
ε	prediction error	-
j	traffic density grid	-
J	total traffic density grids	-
k	traffic speed grid	-
K	total traffic speed grids	-
o	traffic speed-density point	-
$\tilde{X}$	*measured* traffic variable	-
$\bar{X}$	*aggregated* traffic variable	-
$\underline{X}$	traffic variable *per lane*	-

Chapter 1

Introduction

In their popular book on *uncertainty*, Morgan et al. (1992) argued that dealing with uncertainty in policy-focussed research has become fashionable but also noted that doing so is indeed essential. Policy-focussed research where uncertainty is being studied, includes, but not limited to, climate systems (Palmer, 2000, for example), hydrological systems (Zappa et al., 2011, for example), geographical systems (Evans, 2012, for example), and waste management systems (Clavreul et al., 2012, for example). Morgan et al. (1992) argued that similar to physical sciences where estimates of uncertainty always accompany experimental results, quantitative-policy analyses or decision-support modelling systems need to do the same for three system-independent reasons. Firstly, dealing with uncertainty helps identify important factors of disagreement in any problem, and anticipate the unexpected and plan accordingly. Secondly, it helps decision makers assess whether results obtained from different entities are actually different or not, and make decisions accordingly. Thirdly, it provides greater confidence in appropriately using any work done at any point in time.

The vehicle exhaust emission modelling system is an example of a decision-support modelling system. It predicts exhaust emissions of all vehicles driving on a road network during a period of time. These predictions are used to feed in traffic-related emission inventories i.e. the total emissions of each traffic-related air pollutant to the atmosphere. Emission inventories are required to be reported to fulfil national and international obligations of, for example, the National Emission Ceilings Directive or the Convention on Long-range Trans-boundary Air Pollution (European Environment Agency, 2016). They are also used as input to physical air quality models giving rise to 'emission - air quality' modelling chain (Namdeo et al., 2002; Borrego et al., 2003, 2004, for example). While emission models represent what comes out directly from a particular source such as vehicle exhaust, air quality models represent the final outcome of emissions from all available sources after being exposed to dispersion and chemical transformation (McCubbin & Delucchi, 1999; Mayer, 1999). Such outcome is required for a number of national and international directives such as the UK Environment Act

(Department for Environment, Food and Rural Affairs, 2011) and the European Ambient Air Quality Directive (European Commission, 2016). Any (expected) violation of such directives or acts implies that adequate policy measures need to be devised at an appropriate scale i.e. local, national, or regional. These can be source-related measures such as improving vehicle exhaust emission standards, road network-related measures such as optimising traffic management systems, or receptor-related measures such as developing technology-based emission dispersion systems or even information systems. Testing any of these measures before implementation as an evidence of expected compliance, relies primarily on emission modelling. Other than their use to directly inform policy, outputs of air quality models are also used to predict human exposure levels (Gulliver & Briggs, 2005; L. Zhang et al., 2013, for example). The latter relies both on air quality and on human activity itself, hence the difference between the two. As such, a complex 'emission - air quality - human exposure' modelling chain arises, founded upon the modelling of emissions.

Emission models depend primarily on emission factors and traffic activity data. Emission factors specify the rate at which emissions are generated, typically per unit activity, and are developed based on emission measurements of sample driven vehicles - the source being the vehicle exhaust. The detail level of emission factors determines the relevant detail level of traffic activity data. This can range from second-by-second individual vehicle dynamics to aggregate traffic dynamics or even static traffic data. Given the spatiotemporal and predictive power limitations of direct traffic activity measurements, emission models typically rely on the outputs of traffic flow models. Such models provide the basis for testing of potential future policy measures at the source or road network level. As such, a more complex 'traffic flow - emission - air quality - human exposure' modelling chain arises, founded upon the 'traffic flow - emission' sub-modelling chain. In fact, the traffic flow models themselves rely on input data that may come not from direct traffic activity measurements, but from traffic assignment models that distributes traffic demand on a road network accounting for the route choice of individual drivers. This gives rise to an even more complex 'traffic assignment - traffic flow - emission - air quality - human exposure' modelling chain.

The above-mentioned complex modelling chain introduces where emission models are situated and how emission predictions are able to inform transport and environmental policy. Therefore, this modelling system is not an exception to the best practices highlighted in Morgan et al. (1992) to dealing with its uncertainty as a rule to better under understand the system under study and ultimately to better inform policy. The goal of this research is to explore uncertainty in the initial sources of such complex modelling chain, particularly, in the 'traffic flow - emission' sub-modelling chain; traffic assignment is outside the scope of this research and direct measurement data are obtained as input. The following three sections provide more details on specific research aims and objectives (§ 1.1), more details on specific research contributions to current knowledge (§ 1.2), and lastly a summary of the structure of this research (§ 1.3).

1.1 Research aim and objectives

In order to explore uncertainty in the 'traffic flow - emission' modelling chain, it is first necessary to select models for both the traffic flow and emission components. A basic outline is included in this section to allow understanding of the aims and objectives of this research. Various approaches have been developed so far to integrate traffic flow and emission models, depending on the detail level desired in emission predictions. Macroscopic traffic flow models, which represent traffic as a continuum and consider only aggregate traffic behaviour, have recently been integrated with different type of emission models in an attempt to improve on the low-fidelity of static traffic flow models while avoiding the high costs associated with microscopic traffic flow models in terms of their computation, input requirements, and parameter calibration and validation. Research studies have developed methodologies for the direct integration of macroscopic traffic flow models with average speed-based emission models. This research focusses on this integration approach for three main reasons.

1. There exist two widely used macroscopic traffic flow models which vary in structure, input, and parameters, although they produce similar spatiotemporal resolution of traffic flow outputs. These are the first-order LWR-type models (Lighthill & Whitham, 1955a, 1955b; Richards, 1956) and second-order Payne-type models (Payne, 1971). Both of which have been integrated with average speed-based emission models. Critical comparison of these models typically focusses on theoretical issues. Their performance has rarely been compared quantitatively using real traffic data. Also, their influence on uncertainty in emission predictions has not been investigated and compared before.
2. Unlike other traffic flow models, real traffic data (albeit not without their own limitations) sufficient for the calibration, validation, and the quantification of uncertainty in the outcome of macroscopic traffic flow models is widely collected both in space and in time and is generally easily accessible.
3. Direct integration of macroscopic traffic flow models with average speed-based emission models provides an initial ground for developing a research methodology to exploring uncertainty in emission predictions, which can further be expanded to more complex integration approaches.

The second key methodological aspect that needs to be determined is an adequate approach to dealing with uncertainty, which primarily depends on how one needs to inform policy (Morgan et al., 1992). Uncertainty can be investigated through *sensitivity analysis* which enables identifying sensitive inputs to emission models or through *uncertainty analysis* (better referred to as *uncertainty contribution analysis* in Clavreul et al. (2012)) which enables investigating where uncertainty in emission predictions originates from. In addition to that, one of the most powerful tools to informing policy is *uncertainty propagation* which allows for quantifying uncertainty in emission predictions given the uncertainty in its inputs. As such, uncertainty propagation enables

predicting emissions in a probabilistic manner. The sources of uncertainty in the 'traffic flow - emission' modelling chain is sufficiently large that reporting probabilistic-based emission predictions rather than point-based emission predictions is essential. One of the early attempts to doing so is that of Kini and Frey (1997). Their work however was focussed on uncertainty in emission factors rather than traffic flow modelling outputs. This research focusses on uncertainty propagation as a tool to exploring uncertainty in the 'traffic flow - emission' modelling chain, while recognising the different possibilities to doing so.

Having identified the two main aspects, the main aim of this research is to quantify uncertainty in emission predictions given the uncertainty in the predictions of macroscopic traffic flow models. In order to achieve that, this research has two main objectives set out as follows.

1. To develop a data-driven methodological framework that enables quantifying and propagating uncertainty in the outputs of macroscopic traffic flow models to emission predictions. The framework needs to be applicable to macroscopic traffic flow models and any road traffic network with sufficient real traffic data.
2. To apply the developed methodological framework on multiple real road networks in order to ensure applicability and transferability of the framework as well as to allow rigorous quantitative comparison of results amongst the utilised macroscopic traffic flow models.

1.2 Research contributions

This research makes *four original contributions* to existing knowledge. One of the critical issues in macroscopic traffic flow models, or in fact any traffic flow model, is the calibration and validation of their parameters against real traffic data. The complexity of the calibration problem leads to difficulties in finding the 'global' optimal solution, reflecting uncertainty around the model parameters. This uncertainty can be a result of not only limitations in the optimisation algorithms but also uncertainty in the model structure and model input. Efforts have been focussed on developing new optimisation algorithms or tailoring existing ones to illustrate the possibility of getting a 'single' good solution for a typical time period with good performance of the models against real traffic data. Good performance is normally assessed using single-valued quantitative and/or qualitative measures. The first *contribution* to current knowledge is developing an ensemble-based approach to the optimisation problem using available real traffic data combined with statistical testing. This allows to better understand the landscape of all possible good solutions on different time periods and to select 'good' solutions objectively. The ensemble of 'good' solutions ultimately allows better estimation of uncertainty in the traffic flow model outputs, mainly traffic density and speed, in terms of

prediction errors. The latter can be based on calibration outputs (i.e. given optimal parameters) or validation outputs (i.e. non-optimal parameters). Calibration prediction errors allow to understand uncertainty of current scenarios, whereas validation errors allow to understand uncertainty in future scenarios, although not completely, due to additional sources in model structure and input that can arise in future scenarios. The second *contribution* directly follows which is the use of a novel method to propagate the quantified uncertainty in the traffic flow model outputs to emission predictions. The method entails the use of both bivariate probability error distributions of traffic density and speed with a grid-based Monte Carlo simulation method. The use of the first reflects the dependence of the two error distributions. The use of the second grid-based approach reflects the dependence of error values on predicted traffic variables.

The first and second contributions constitute the main elements of the developed methodological framework that can be applied to different macroscopic traffic flow models and any average speed-based emission model. As such, the third *contribution* is the application of the methodology to real road networks using CTM, a widely used first-order LWR-type model, and METANET, a widely used second-order Payne-type model. This allows a rigorous quantitative comparison of the two models not only of their influence on uncertainty in emission predictions but also of their performance as traffic flow models. In doing so, the first phases of the developed methodological framework are shown to be useful as a stand-alone approach applicable to other applications of macroscopic traffic flow models, for instance quantifying uncertainty in travel time prediction.

The fourth and last *contribution* is the application of the developed methodological framework to three road networks, one of which is operated under Variable Speed Limits (VSL). This investigation includes two structures of CTM and two structures of METANET: for each model the first is the original, and the second is a modified structure applicable to VSL (from current literature). This research provides a comparison between using the original structure and the modified structure on the same road network. It also draws conclusions on the complexity of the problem where a modified model structure alongside the uncertainty in the model parameters hinders a better understanding of the influence of VSL on traffic dynamics through traffic flow models.

1.3 Thesis structure

The rest of this thesis is divided into five chapters, the content of each is briefly described here. Figure 1.1 below illustrates how these chapters are interlinked.

- Chapter 2 introduces the main developments in the fields of traffic flow modelling and emission modelling. A description of macroscopic traffic flow models and their

two most controversial sub-categories, the first-order LWR-type and second-order Payne-type models, is emphasised. This allows better understanding of currently available integration approaches of macroscopic traffic flow models and emission models, which are also reviewed in the chapter. The chapter ends by introducing the concept of uncertainty in decision-support modelling systems and how it has been dealt with in emission predictions. The aim of this chapter is to provide a critical review of relevant studies and to identify gaps that justify the need to the original contributions of this research described earlier. This eventually justifies the main elements of the research methodology.

- Chapter 3 focusses entirely on the research methodology which is divided into three steps. The first sets out the mathematical preliminaries of selected macroscopic traffic flow models, emission model, and their integration approach. The second discusses the sources of uncertainty in the resulting modelling chain and identifying those which are to be focussed upon. The third and most important step describes a five-phase data-driven methodological framework towards a probabilistic-based emission predictions. The framework consists of a data preparation preliminary phase and four main phases (Phases I, II, III, and IV).
- Chapter 4 is dedicated entirely to the data preparation preliminary phase. It describes in detail the sources and the general extraction and pre-processing of necessary data. It then describes the road networks and time periods selected for this research, for which their corresponding data are extracted and pre-processed. The chapter ends also by providing a summary of the final data sets as a necessary background to the discussion of results obtained for the four main phases.
- Chapter 5 presents and discusses the results of each phase of the methodological framework in four separate sections: the ensemble-based optimisation phase, the prediction error estimation phase, the grid-based Monte Carlo sampling phase, and the final uncertainty propagation phase. The latter presents the probabilistic-based surfaces of emission predictions for each studied road network. The chapter ends by providing an example of how these can be used to calculate probabilistic-based emission predictions of a particular traffic flow modelling scenario.
- Chapter 6 summarises the results obtained, expands on the limitations of this research, particularly of each phase of the developed methodological framework. The chapter ends by recommendations of possible future work in the domain of uncertainty in the complex five-step modelling chain introduced earlier.

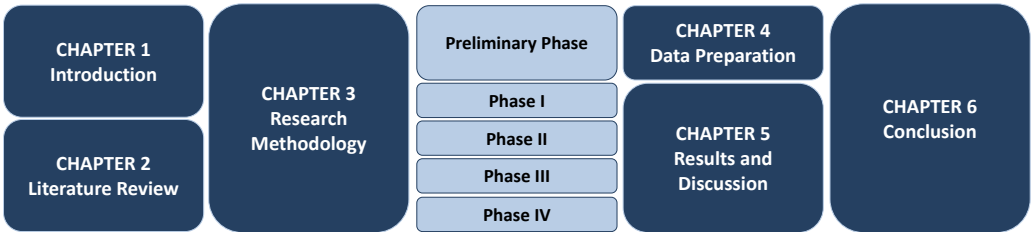


FIGURE 1.1: Thesis structure.

Chapter 2

Literature Review

2.1 Introduction

This chapter provides a summary and critical review of literature on the subjects introduced earlier in Chapter 1. The purpose of this chapter is to both highlight the current state of knowledge as well as to identify gaps which are considered the main motivation behind the research methodology developed in Chapter 3. This chapter attempts to provide:

- necessary background of traffic flow modelling with emphasis on the macroscopic modelling category, particularly the debate between its different classes (first-order and higher-order), and current estimation methods of their parameters;
- necessary background of vehicle exhaust emission modelling, with particular emphasis on vehicle exhaust emission measurements upon which the models are developed, and the models' traffic-related input;
- an overview of current integration approaches of traffic flow and emission models with particular emphasis on those based upon macroscopic traffic flow models; and
- necessary background of the concept of uncertainty in terms of its definition and its different dimensions, with a focus on the studies who have dealt with uncertainty in the context of emission predictions.

Developments in the fields of traffic flow and emission modelling are addressed in § 2.2 and § 2.3, respectively. They provide the necessary background to understanding the integration approaches of the two fields developed so far (§ 2.4), and to broadly identifying the sources of uncertainty in the different modelling components (§ 2.5). By reviewing the integration approaches in § 2.4, the choice of a particular integration approach to be focussed upon in this research is justified. The chapter concludes with a summary of findings and gaps motivating the research methodology in § 2.6. It is important here to note that while the selected traffic flow and emission models are

introduced in this chapter, the mathematical preliminaries behind these are provided as part of the research methodology (§ 3.2).

2.2 Developments in traffic flow modelling

Based upon certain concepts or so-called fundamentals, traffic flow models are mathematical equations which provide a description of the complex traffic system comprising vehicles, drivers, and roadways (H. Zhang, 1998; Treiber & Kesting, 2013). This is made up to a certain level of accuracy and precision in comparison to the real world. The mathematical equations define the traffic flow model structure, input, and parameters; the last two typically need to be determined in accordance to the road network and time period under study. Fundamentals of traffic flow are developed mainly through quantitative observations/data of one or more components of traffic such as:

- trajectory data using cameras and global positioning systems; and
- cross-sectional data collected using (single or dual) loop detectors or radar sensors.

Examples of traffic flow fundamentals are time-space trajectory of vehicles; traffic density, speed, and flow characteristics and their relationships; and spatiotemporal dynamics of traffic state (e.g. Leutzbach (1988) or Hoogendoorn and Knoop (2013), for a background on traffic flow fundamentals). While Bruce Greenshields and colleagues have initiated the work with traffic data and the fundamentals of traffic flow in the 1930s (Greenshields et al., 1934), major breakthrough in the field of traffic flow modelling took place between 1950 and 1960, mainly by Louis Pipes (Pipes, 1953), Michael Lighthill and Gerald Whitham (Lighthill & Whitham, 1955a, 1955b) (their first paper develops the theory with respect to flood movement while the second with respect to traffic flow), Paul Richards (Richards, 1956), and Ilya Prigogine and Frank Andrews (Prigogine & Andrews, 1960). There exist two main aspects to the field of traffic flow modelling:

1. The development of a traffic flow model and its verification - a process to check if the developed model performs as expected qualitatively and quantitatively.
2. The calibration (or optimisation) and validation of the model parameters using real traffic data.

Based on the performance of the calibration and validation models against real traffic data, one can determine if such models can be directly used in further applications, need further improvement/modification before any model application, or perhaps even considered as not viable. Often times, however, the performance of a model needs to be tested through its application. For instance, applying a traffic flow model for travel time reduction policy assessment might require only the replication of congestion phenomenon but its application for emission reduction policy assessment might require also the close approximation of both low and high average traffic speeds. This also

implies that performance measures are very much dependent on the application of a traffic flow model.

The number of traffic flow models developed since the 1950s to date is sufficiently large that it became difficult to keep track of each one. Review studies typically attempt to refer to developed models by categorising them. Such categorisation is typical of mathematical modelling. The large number of models also makes it more difficult for the two aspects (verification or calibration and validation) to be addressed. This section starts by the categorisation of traffic flow models to date based on most recent reviews while providing a description of the broad categories i.e. definitions, advantages, and disadvantages/limitations. This is considered the first step to justifying the selection of macroscopic traffic flow models. This is followed by a focussed review of literature concerning the calibration and validation of macroscopic traffic flow models.

2.2.1 Categorisation of (dynamic) models

Traffic flow models have been categorised differently in literature depending on the criterion being used. Treiber and Kesting (2013) highlighted seven possible criteria one can use, depending on the purpose of the study. These are aggregation level, mathematical structure, conceptual foundation, randomness, identical versus heterogeneous drivers and vehicles, constant versus variable driving behaviour, and single-lane versus multi-lane models. Most commonly referred to criterion in literature is the aggregation level. Following this, the three microscopic, mesoscopic, and macroscopic model categories arise.

Given the large number of models developed by researchers from engineering, mathematics, operations research, and physics (Helbing, 2001), many have focussed on reviewing particular categories of models. This is often limited to certain, key sub-categories. For instance, Papageorgiou (1998) focussed only on reviewing the macroscopic category of models while highlighting the qualitative issues of most commonly used sub-categories: first-order and second-order models. H. Zhang (2001) focussed on the same but rather highlighted the commonalities between certain sub-categories of macroscopic models, mainly, the first-order and second-order models. Brackstone and McDonald (1999) focussed on the microscopic category of models and reviewed a number of historical models under the car-following sub-category by assessing their availability and validity. Orosz et al. (2010) focussed on the description of the dynamical processes of two categories of models which, according to them, have shaped the traffic flow modelling research: continuum-based macroscopic models and car-following microscopic models.

More comprehensive reviews have also been made, for instance by Hoogendoorn and Bovy (2001), Helbing (2001), Bellomo et al. (2002), Maerivoet and De Moor (2005), Bellomo and Dogbe (2011), and Treiber and Kesting (2013). Perhaps the most recent

and comprehensive is that of van Wageningen-Kessels (2013) and van Wageningen-Kessels et al. (2015) who provided not only a classical review of models and their mathematical literature but also a historical overview of the main highlights in the field. Their historical review was summarised in a 'genealogical tree' of traffic flow models. The tree is characterised by 'nodes' under each category which refer to major developments (based on publications) of models and 'links' which represent connections between different models.

The tree highlighted the fact that all the developed dynamic models were established based on the fundamental diagram. The fundamental diagram describes the theoretical relationship of two aggregate traffic variables: traffic speed-density diagram, flow-density diagram, or speed-flow diagram. Different shapes have been proposed throughout the years. Amongst others are the parabolic traffic density-flow relationship (Greenshields et al., 1935) and the triangular and trapezoidal relationship (Daganzo, 1994, 1995a). M. Li (2008) and Carey and Bowers (2012) provide an extensive review of the different relationships proposed.

Fundamental diagrams can by themselves be used as static models. Static models are also referred to as 'whole-link' models in Wilson (2008) since they do not consider within-link traffic structure, as opposed to dynamic models. Most often, traffic speed versus flow relationship is transformed to travel time versus flow relationship on a link. These can be used as cost functions in the traffic assignment step of transport models. Transport models typically consist of four sub-models: trip-generation, trip distribution, mode choice, and traffic assignment. As opposed to traffic flow modelling, transport modelling does not use readily available traffic demand data but are rather modelled internally and iteratively (e.g. de Dios Ortúzar and Willumsen (2011) for more details on transport modelling). Traffic demand data is considered the main input to any traffic flow model. This study deals with scenarios where traffic demand is extracted using real traffic data. While transport modelling is outside the scope of this research, it is noteworthy to highlight the difference. In many cases, static models (and also dynamic models) are embedded in transport models to predict emissions, examples of which are provided in § 2.4.

The development of dynamic traffic flow models has often followed the broad families of microscopic, mesoscopic, and macroscopic i.e. the aggregation level criterion. Since this categorisation can also be followed in the field of vehicle exhaust emission modelling, these three categories of dynamic models are introduced next.

2.2.1.1 Microscopic models

The term 'microscopic' refers, in general, to entities individually identified (Bellomo & Dogbe, 2011). Microscopic traffic flow models describe the dynamics (i.e. position,

speed, and acceleration) of individual vehicles with respect to time under the influence of surrounding vehicles. They are thus considered the most detailed and least aggregated. As such, their modelling structure is a large system of Ordinary Differential Equations (ODEs), i.e. involving functions dependent on a single variable: time. The solution of the system provides a detailed description of traffic conditions on a predefined road network.

In the genealogical tree of van Wageningen-Kessels et al. (2015), there are 25 nodes (models) under this category which have different modelling structures . These nodes (models) mainly correspond to two sub-categories of microscopic models: car-following, first developed by (Pipes, 1953), and Cellular-Automata, first developed by Cremer and Ludwig (1986). Car-following models can further be distinguished to safe-distance models (Gipps, 1981; Newell, 1961, 2002, for example), stimulus-response models (Helly, 1959, for example), and action point models (Brackstone & McDonald, 1999, for example). These models adhere to the 'follow-the-leader' concept but with different assumptions on the process of which a vehicle (follower) follows another vehicle (leader). Thus, they are referred to as longitudinal car-following models. van Wageningen-Kessels et al. (2015) noted that the lateral lane-changing models were not included in their historical review. Recent review of lane-changing microscopic models by Rahman et al. (2013) showed that since the 1980s, 11 further nodes appeared under this category of models. It starts with the lane-changing model proposed by Gipps (1986).

The detailed output of individual vehicle dynamics (also referred to as vehicle trajectories) is advantageous in some aspects. It has been applied in fields which require such information such as the evaluation of in-vehicle measures - for instance, driver assistance systems in Kesting et al. (2008), or roadside measures - for instance, traffic control in Yang and Koutsopoulos (1996). Conceptually, microscopic traffic flow models also allow the derivation of macroscopic quantities. However, this category of models suffers from a number of limitations, listed as follows:

- the large system of equations implies high computational and monetary cost required for data collection as well as model development and simulation;
- detailed input required for the road network and vehicle characteristics as well as for the calibration and validation - although this can be advantageous as it implies they can take into account the influence of multiple vehicle classes and lanes; and
- the large number of parameters, corresponding to individual driving behaviour, implies a challenging and demanding calibration and validation process.

This limits their application to small road stretches/road networks. Their application is also limited to offline rather than online as a result of the absence of an explicit analytical relation between the model input and output as suggested by Hoogendoorn and Bovy (2001). Traffic-related online applications are often targeted to real-time

traffic state estimation and traffic control. Most importantly, their calibration and validation become often based on aggregate quantities, which can be easily observed and measured in comparison to individual vehicle dynamics. As a result, while the output level of detail is high, it becomes much harder to determine if such detail is reliable or not. Although recent studies have been proposed on the calibration and validation of microscopic traffic flow models using sampled vehicle trajectory data, Ossen and Hoogendoorn (2008) suggested that the process is far from trivial.

2.2.1.2 Mesoscopic models

Mesoscopic models are also referred to as 'statistical descriptions' in Bellomo and Dogbe (2011). They provide a description of traffic flow at a medium level of detail. They somehow fill the gap between the least detailed macroscopic (as will be described next) and most detailed microscopic models. Mesoscopic traffic flow models describe traffic behaviour at an aggregate level using, for instance, probability distributions of traffic variables whereby the behavioural rules are defined individually.

The tree of van Wageningen-Kessels et al. (2015) has 11 nodes of mesoscopic traffic flow models with the earliest classical models being the gas-kinetic model proposed by Prigogine and Andrews (1960). In their model, speed distributions are described at small segments and time intervals by taking into account the acceleration and interaction dynamics between vehicles. A number of models were then developed and extensions of the classical gas-kinetic model have been proposed. While such models are less detailed and less parameterised in comparison to microscopic traffic flow models, mesoscopic models such as the gas-kinetic ones have also been criticised for having too many variables. This limits their applicability to: offline applications, and cross sections and stretches rather than motorway/urban networks (Hoogendoorn & Bovy, 2001). Nonetheless, the main reason of not using any model under this category in this study is their scarce integration with emission models, as will be noticed in § 2.4.

2.2.1.3 Macroscopic models

The term 'macroscopic' refers, in general, to averaged gross quantities (Bellomo & Dogbe, 2011). Macroscopic traffic flow models provide a description of traffic flow dynamics at an aggregate level yet at a high spatiotemporal resolution. Thus, individual vehicles/drivers are not modelled. They require aggregate input quantities and they output the dynamics of aggregate quantities as variables (such as traffic density) in both space and time. They are considered the least detailed amongst other dynamic model categories. However, they are normally able to describe collective phenomena of traffic flow such as the evolution of traffic congestion on a road network during a particular time period, which is useful in many applications. Their modelling structure

is typically a system of Partial Differential Equations (PDEs), i.e. involving functions dependent on more than one variable: space and time.

Macroscopic models of traffic flow have first been derived in analogy to continuum models of fluids. So they are sometimes referred to as 'continuum' models or 'hydrodynamic' models. Darbha et al. (2008) or Tyagi et al. (2009) criticised such an analogy of traffic to fluid. This is on the basis that the notion of 'density' in fluids corresponds to a sufficiently large number of molecules at length scales of interest (of the order of 10^{23}) which is definitely not the case in traffic. Papageorgiou (1998) argued that there is hardly any hope for macroscopic traffic flow models to reach levels in other domains such as Newtonian Physics or thermodynamics. So, one needs not to seek such descriptive details in the assessment of this category of models but rather their quantitative and qualitative descriptive power of real-traffic data and real-traffic phenomena, depending on the application.

The tree of van Wageningen-Kessels et al. (2015) has 19 nodes (models) with different modelling structures under this category. All of which emerged as a modification of the basic *L*ighthill, *W*hitham, and *R*ichards (LWR) model. Main sub-categories are the first-order LWR-type models and second-order Payne-type models described below. While higher-order models have also been proposed such as the third-order Helbing-type models (Helbing, 1996, for example), LWR-type and Payne-type models have received more attention in literature and in their real-world application. The foundation of both LWR-type models and Payne-type models is the conservation law of vehicles along with the assumption that the three traffic variables are related by $q = \rho \times u$ ($flow = density \times speed$), such that:

$$\frac{\partial \rho}{\partial t} = -\frac{\partial q}{\partial x} \qquad \text{or} \qquad \frac{\partial \rho}{\partial t} = -\frac{\partial \rho u}{\partial x} \tag{2.1}$$

In physical systems, the conservation law is also referred to as the mass or material balance. It basically means that nothing is created or destroyed spontaneously. Equation 2.1 describes the dynamics (or the temporal evolution) of traffic density ρ as a function of traffic flow differences or gradients. The equation illustrates that the change in traffic density during time interval ∂t is equal to the change in traffic flow over distance ∂x. A macroscopic traffic flow model is then defined by specifying how to calculate traffic flow or speed i.e. by specifying another independent model equation. Papageorgiou (1998) noted that the conservation equation is perhaps the only accurate physical law available for traffic flow, and all other equations reflect either idealisations and analogies or approximations of real data.

First-order LWR-type models: The LWR model was described in two seminal papers by Lighthill and Whitham (1955b) and Richards (1956). It makes the simplest assumption of traffic flow or speed: they satisfy the equilibrium (or steady-state) relationship i.e. they are static functions of traffic density. As such, the equation above

turns into,

$$\frac{\partial \rho}{\partial t} = -\frac{\partial q^e(\rho)}{\partial x} \qquad \text{or} \qquad \frac{\partial \rho}{\partial t} = -\frac{\partial \rho u^e(\rho)}{\partial x} \tag{2.2}$$

where $q^e(\rho)$ or $u^e(\rho)$ represent the equilibrium traffic flow-density or traffic speed-density relationships, respectively. The LWR model and its variations (or extensions) are often referred to as LWR-type models or first-order macroscopic traffic flow models. Variations of the LWR model can be in terms of: the model structure itself (i.e. targeting a particular aspect such as the fundamental diagram shape or deficiency such as heterogeneity of drivers or multiple lanes); or in terms of scale of representation (i.e. whether it is continuous or discrete). In real applications, a space-time discretised LWR-type model needs to be used. By breaking down a road segment into short enough cells and a time period into small enough time steps, it becomes then possible to describe the space-time dynamics of traffic density at each cell and time step rather than continuously at all locations and all time instants. Depending on the problem, the length of cells and time steps can be set to (for example) to 100 metres each and 1 second, respectively. This implies that the discretised model would output aggregate quantities of traffic variables at each 100 metres cell and each 1 second time step.

Daganzo (1994, 1995a) was the first to propose a discretised first-order model, referred to as *C*ell *T*ransmission *M*odel (CTM), as an approximation of the LWR model. Discretised models such as CTM or its variations make the simulation of traffic flow simpler and more efficient, allowing their generalisation to both motorway and urban road networks. This is mainly a result of:

- a simple model structure;
- few parameters involved which are only those of the fundamental diagram; and
- few input requirements which are mainly the simplified characteristics of a road network as well as boundary and initial conditions (at a resolution similar to the discretisation used).

The simplicity of this type of models is accompanied with a number of drawbacks which have induced a lot of debate. These drawbacks originate primarily from the assumption of equilibrium traffic speed-density relationship which implies that:

- traffic speed adjust instantaneously to changes in traffic density without considering any delay. This can lead to unbounded or infinite accelerations. While some have proposed modifications to the LWR model itself to bound the acceleration (Lebacque, 1997, 2002, 2003, for example), the main development to resolve this issue was the rise of higher-order models which are described below; and
- non-equilibrium (or variations around the equilibrium traffic speed) conditions observed in real-life traffic are not taken into account. These are normally a result of traffic flow instabilities and traffic hysteresis. Traffic flow instabilities such as start-stop waves and capacity-drop can occur as a result of on-ramp merging or

lane-drops. Traffic hysteresis can occur as a result of different vehicle acceleration and deceleration processes. Of course many of the variations or extensions of LWR have targeted this drawback in different ways. For instance, Wong and Wong (2002) proposed a multi-class model and argued that their model can predict phenomena such as capacity-drop and traffic hysteresis. Other studies have also proposed stochastic LWR models, to explain the random variations behind the non-equilibrium conditions, occurring due to factors such as weather and lighting conditions or percentage of heavy duty vehicles (Sumalee et al., 2011; Ngoduy, 2011, for example). However, a common attempt to resolving these issues is also through the development and use of higher order-models.

Despite the above-mentioned limitations of first-order LWR-type models, they have been commonly used in a number of fields. While the main interest of this research is their application in the emission modelling field (which is to be expanded upon in § 2.4), examples of other applications are:

- traffic control optimisation on motorways such as ramp metering (Sun et al., 2003; Gomes & Horowitz, 2006; Gomes et al., 2008), Variable Message Signs (VMS) (Shang & Huang, 2007), and VSL for capacity improvements (Hadiuzzaman & Qiu, 2013) or collision risks and injury severity reductions (Chiou et al., 2012; Z. Li et al., 2015);
- traffic signal optimisation on urban roads (Lo, 1999; Feldman & Maher, 2002); and
- evacuation planning (Ben-Tal et al., 2011; Kimms & Maassen, 2011).

Second-order Payne-type models: Higher-order models have also been developed such as the second-order Payne-type models with two equations describing the dynamics of both traffic density and speed. These have been first proposed by Payne (1971) who suggested replacing the equilibrium traffic speed with an equation describing its dynamics (acceleration equation), as follows:

$$\frac{\partial u}{\partial t} = \frac{u^e(\rho) - u}{\tau} - \frac{u\partial u}{\partial x} - \frac{\eta \partial \rho}{\tau \rho \partial x} \tag{2.3}$$

Equation 2.3 above essentially makes the following three assumptions:

1. Traffic speed tends to relax to equilibrium traffic speed represented by the fundamental diagram. This is referred to as relaxation and involves the relaxation term τ. It is represented by the first term on the right side.
2. Traffic speed at the upstream tends to travel downstream (ahead) at the same speed. This is referred to as convection. It is represented by the second term on the right side.
3. Traffic speed tends to anticipate to downstream traffic conditions. This is referred to as anticipation and involves the anticipation term η. It is represented by the third term on the right side.

Many variations around the basic Payne model have been developed giving rise to the generation of second-order Payne-type models (Phillips, 1979; Kuhne, 1991; Kerner et al., 1996, for example). The main difference amongst these models is the acceleration equation, particularly the assumptions taken into account which lead to non-equilibrium traffic conditions. This means that parameters (number and representation) are dependent on the model itself. Variations around the original Payne model are made also to tackle certain aspects such as the fundamental diagram shape or certain deficiencies such as heterogeneity of drivers or multiple lanes.

Similar to LWR-type models, real-world implementation of second-order models requires their space-time discretisation. First discretised model was developed by Payne (1979) which was based on his original continuum-based model i.e. the acceleration equation term remained the same. Papageorgiou et al. (1989, 1990) also proposed a discretised Payne-type model but with the addition of two terms to the acceleration equation which account for: the influence of merging at on-ramps, and the influence of lane-changing at lane-drop areas. The reason behind these extensions was to resolve the poor model performance of the original model in some cases at on-ramps and lane-drop areas (Papageorgiou, 1998). This discretised model was also developed into a commercial software referred to as *M*odéle d'*É*coulement de *T*rafic sur *A*utoroute *NET*works (METANET) (Messner & Papageorgiou, 1990).

With the introduction of an acceleration equation, the main argument against LWR-type models (i.e. the equilibrium traffic speed-density relationship) is resolved. However, this was at the expense of higher complexity in terms of model structure i.e. system of two dynamic equations instead of one, input required, and the number and physical meaning of parameters. Payne's model has also been criticised for a number of reasons on a rather theoretical basis. For instance, Daganzo (1995b) and Aw and Rascle (2000) demonstrated the possibility of getting negative speeds in certain scenarios as a result of the anticipation term and possibly the relaxation term. Del Castillo et al. (1994), Daganzo (1995b), and Aw and Rascle (2000) also showed that information in Payne's model can flow faster than the vehicles themselves, proposing the model is flawed. While some models were proposed to resolve the issue for second-order models (Aw & Rascle, 2000, for example), Papageorgiou (1998) argued that this is not a physical contradiction as long as the speed of information is lower than a maximum individual vehicle speed since traffic flow (particularly on motorways) consists of vehicles and lanes with systematically different average speeds.

Second-order Payne-type models have gained a lot of attention in the traffic control field (Van den Berg et al., 2003; Hegyi et al., 2005a, 2005b; van den Berg et al., 2007; Carlson et al., 2010; Groot et al., 2011a) but also they have been used for incident detection purposes (Willsky et al., 1980, for example). Most importantly is their application in the emission modelling field which is to be expanded upon in § 2.4.

Papageorgiou (1998) argued that macroscopic traffic flow models are, to a large extent, empirical models. This means that sufficient quantitative accuracy of any model needs to be determined through the calibration and validation of a model's parameters with real data. Depending on the application, the outputs' comparison with real data needs to be the ultimate indication of their usefulness. Regardless of the order of macroscopic traffic flow models, they have three attractive properties:

1. Simple model structure implying relatively lower computational time and low cost of development, calibration, and simulation.
2. Few input data required in terms of the road network characteristics and initial and boundary conditions which can be relatively easy to measure and obtain.
3. Few parameters to be calibrated based on data of aggregate quantities which are relatively easy to measure and obtain.

Going back to the scope of this work, the limitations of microscopic and mesoscopic models as well as static models illustrate the attractiveness of macroscopic ones. These are dynamic models taking into account within-link traffic conditions as opposed to static models. They are also simpler in terms of structure, input, and parameterisation as opposed to other dynamic models.

It is argued here that the methodological framework developed in this study based on macroscopic traffic flow models is transferable to other modelling categories, in terms of the general headlines/phases. In fact, it is argued that the methodological framework (with suitable modifications) needs to be applied inherently in any integration approach, as part of policy and decision making process. While macroscopic traffic flow models (which have gained a lot of attention in the emission prediction field as will be described in § 2.4) are to be focussed upon here, other model categories are also used for emission prediction. Nonetheless, there exist two main reasons to focussing on macroscopic traffic flow models.

1. The interesting debate between first-order and second-order models. It is argued here, that by studying uncertainties and errors in emission predictions using each of the first-order and second-order macroscopic traffic flow models, new insights can be found on their advantages and disadvantages of their use, both in isolation (i.e. before their application to predicting emissions), and in their application to predicting emissions.
2. The availability of necessary (and large amount of) data to studying uncertainties and errors arising from macroscopic traffic flow models.

As such, the following section delves more into the most critical aspect to the use of macroscopic traffic flow models: calibration and validation of model parameters. For a detailed review on calibration and validation of microscopic traffic flow models, the reader is referred to Daamen et al. (2014).

2.2.2 Calibration and validation of macroscopic models

Upon the use of any selected macroscopic traffic flow model, input and parameters need to be determined. The model can then be simulated and its outputs can be extracted. While model input can be extracted for a current scenario based on real data, parameters need to be determined through a calibration process, as a first step; a current scenario reflects both the road network and the time period. This is primarily because different road networks can exhibit different driving styles, different classes of vehicles, and different traffic management policies (Treiber & Kesting, 2013). Even for the same road network, different time periods can exhibit different driving behaviour and different vehicle characteristics (Treiber & Kesting, 2013). This implies that default parameters cannot be applied to all scenarios. They need to be adjusted (or optimised) to best fit the data of current scenarios while avoiding the issue of overfitting which prevents their use on scenarios where calibration cannot be performed.

Model calibration refers to the process of finding parameter values which produce the optimal fit between a model's prediction and corresponding real measurements. In the context of macroscopic traffic flow models, the optimal fit needs to reproduce aggregate traffic quantities: traffic density, flow, and/or speed. The calibration process can be either undertaken manually (i.e. trial and error) or using calibration methods (i.e. optimisation algorithms). Manual calibration does not guarantee finding an optimal solution (i.e. optimal parameter values). Here, the focus is on the use of calibration methods. Model calibration requires the following:

- obtaining real measurements for a current scenario for which the model has to be simulated and then compared against - in a perfect scenario, real traffic measurements would be of similar space-time resolution to that of the model's prediction;
- determining the optimisation algorithm;
- determining the model prediction output on which the comparison with real measurements needs to be based - this can utilise a single traffic variable or more; and
- determining the goodness-of-fit measure on which the comparison is to be based upon - this is referred to as the objective function, or sometimes Performance Index (PI).

Depending on the optimisation algorithm, the objective function is evaluated multiple times until the algorithm converges to an optimal solution which minimises the objective function. The optimal parameters obtained are referred to here as 'calibrated parameters'. The outputs of a model simulation based on these parameters (and the same input on which the calibration was based upon) are referred to as 'calibration model outputs'.

The calibration of macroscopic traffic flow models is a non-linear optimisation problem. Depending on the traffic data, the mathematical properties of the selected objective function can be (Treiber & Kesting, 2013):

- smooth and unimodal i.e. the function is differentiable with a single global optimum which can be found using gradient-based optimisation algorithms;
- smooth but multi-minimal i.e. the function does not have a single global optimum and it typically requires derivative-free methods to find a 'good' minimum solution; or
- fluctuating and multi-minimal i.e. the function is not smooth and have multiple minimum solutions (or multiple local solutions) which can be either 'false' or 'good' minimum solutions. It typically requires derivative-free methods to find a 'good' minimum solution for such problems.

As such, there is not a single method that can solve all non-linear optimisation problems but rather the method depends on the problem in hand. Calibrating macroscopic traffic flow models to minimise the difference between its outputs and measurements from loop detectors normally generates a fluctuating and multi-minimal surface (Treiber & Kesting, 2013). This phenomenon has been demonstrated, for example, in Ngoduy and Maher (2012) for a second-order model. Figure 2.1 below shows a fluctuating and multi-minimal surface of the objective function used in Ngoduy and Maher (2012) when all parameters of a second-order Payne-type model (Treiber et al., 1999), except two, are fixed.

In the case of fluctuating and multi-minimal objective function surface, Treiber and Kesting (2013) (and others as will be demonstrated in the review) suggest that the best calibration methods are heuristic derivative-free ones to avoid 'false' minimum

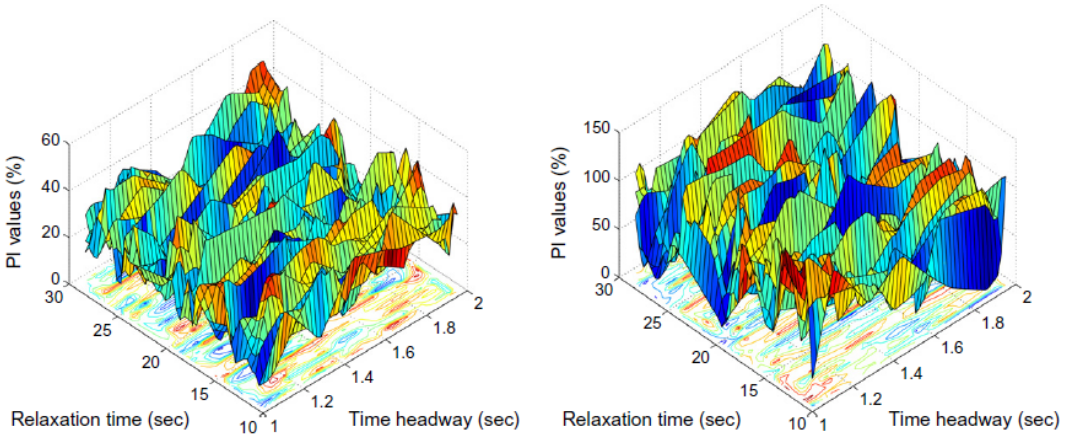


FIGURE 2.1: Fitting surface of the total relative mean squared errors of the predicted and measured speed and density objective function (referred to here as Performance Index or PI) based on a second-order Payne-type model (Treiber et al., 1999). The model has five parameters: two of which are varied in these simulations and the others are held fixed. Left side figure shows simulation results based on synthetic traffic data. Right side figure shows simulation results based on real traffic data. These are **Fig. 1** in Ngoduy and Maher (2012).

solutions. Even with such methods, a 'good' minimum depends on the initial parameter estimates (referred to as initial guesses or initial points), that the optimisation algorithm starts its iterations with.

The main purpose of obtaining calibrated parameters is to use them for predicting future scenarios. Even when such parameters are optimal for a current scenario, it does not mean they can automatically be used for future scenarios (with a change in time periods and/or road network characteristics). This can be a result of, for instance, different driver and driving behaviour, different vehicles and vehicle fleet compositions, and different traffic conditions overall. This can also be a result of model structure deficiencies which the calibrated parameters have accommodated for or over fitted on. Validation is thus used to test the reliability of such parameters in predicting traffic flow using data that has not been used for the calibration. Efforts in the calibration and validation of macroscopic traffic flow models are quite limited. Studies which meet the following three criteria are reviewed.

1. Calibration and validation made using real networks and real traffic data only.
2. Calibration and validation studies which have either used existing calibration algorithms or developed new ones.
3. Calibration and validation of models targeted to offline rather than online applications.

Table 2.1 below provides a summary of the review of both types of models. They summarise eight aspects of each study:

1. The model type used: first-order or second-order.
2. The calibration approach used: '1' indicates the fundamental diagram-based method and '2' indicates the model-based method.
3. The optimisation algorithm(s) used: *S*equential *Q*uadratic *P*rogramming (SQP), *N*elder-*M*ead (N-M), *L*east *S*quares (LS), kernel-based *C*ross *E*ntropy *M*ethod (CEM), *G*enetic *A*lgorithms (GA), *C*uckoo *S*earch (CS), or *P*article *S*warm *O*ptimisation (PSO).
4. The objective function used: *S*quared *E*rror (SE), *M*ean *S*quared *E*rror (MSE), or *R*oot *M*ean *S*quared *E*rror (RMSE). Some studies have used weighted errors, especially when more than one traffic variable is included, but that was dropped for convenience.
5. The number of real road networks employed.
6. The number of time periods used for calibration and for validation, separately.
7. The number of solutions reported per model type, calibration approach, optimisation algorithm, and/or real road network. Although, some studies who use multiple calibration time periods do not report the final solutions of all time periods, they just mention that they do use multiple time periods.
8. The criteria used for assessing the calibration and/or validation performance: '1' indicates qualitative assessment and '2' indicates quantitative assessment.

TABLE 2.1: summary of reviewed calibration and validation studies of first-order LWR-type and second-order Payne-type models[1]

Reference	Model Type	Calibration Approach	Optimisation Algorithm	Objective Function	Real Road Networks	Cal. Time Periods	Val. Time Periods	NR. of Solutions[2]	Assessment Criteria
Muñoz et al. (2004)	first-order	1	LS	SE	1	3	-	1	1/2
Dervisoglu et al. (2009)	first-order	1	LS	SE	1	1[3]	-	1	1/2
Gkiotsalitis and Chow (2014)	first-order	1	LS	SE	1	4	-	1	2
Zhong et al. (2015)	first-order	1	LS	SE	1	2	17	1	1/2
		2	SQP	MSE					
Spiliopoulou et al. (2014)	first-order	2	N-M	RMSE	1	1	3	1	1/2
	second-order								
Cremer and Papageorgiou (1981)	second-order	2	Box	error	1	> 1	NA	1	1/2
Sanwal et al. (1996)	second-order	2	Box	MSE	2	2	1	1	1/2
Frejo et al. (2012)	second-order	1 then 2	SQP	SE	1	5	5	1	1/2
Kotsialos et al. (2002)	second-order	2	Box	SE	1[4]	4	NA	1	1
Ngoduy and Maher (2012)	second-order	2	Kernel-based CEM	relative-MSE	1	1	1	1	1/2
Poole and Kotsialos (2012)	second-order	2	GA	SE	1	5	5	1	1/2
Spiliopoulou et al. (2015)	second-order	2	N-M, GA, CEM	RMSE	1	1	3	1	1/2
Poole and Kotsialos (2016)	second-order	2	GA, CS, PSO	SE	2	6	6	1	1/2

[1] Description of each column is provided in the previous page.
[2] The first four studies distinguish the fundamental diagram based on cells and the last study considers automatic assignment of fundamental diagrams to cells.
[3] Multiple time periods were used for parameter calibration but a single time period used for model simulation.
[4] One large road network divided into four stretches.

CTM, a discretised LWR-type model introduced earlier, is the most commonly deployed first-order model in the reviewed studies of Table 2.1. There exist two main approaches to the calibration of its (only) fundamental diagram-related parameters: conventional fundamental diagram-based or model-based. These have been best highlighted in a recent study by Zhong et al. (2015) who compared the calibration and validation performance of both. The first approach does not consider traffic dynamics. It is merely a calibration problem of the assumed fundamental diagram equation on the traffic flow-density data available for a road network and time period. It does not involve the simulation of the traffic flow model itself. Once the optimal parameters of the fundamental diagram are obtained, they are used for the simulation of the first-order macroscopic traffic flow model. Alternatively, the second approach considers traffic dynamics and calibrates the parameters through the simulation of the model itself. Zhong et al. (2015) concluded that a better model performance is obtained, during both calibration and validation, when using the model-based approach. Using the conventional method does not guarantee that the solution obtained is an optimal solution when it is used in the context of a dynamic traffic flow model. In addition to Zhong et al. (2015) who have deployed a gradient-based SQP algorithm to minimise their objective function, Spiliopoulou et al. (2014) have also followed the model-based calibration approach but have used N-M heuristic derivative-free search algorithm. A number of other studies have followed the conventional method as is shown in Table 2.1 (Muñoz et al., 2004; Dervisoglu et al., 2009; Gkiotsalitis & Chow, 2014). The typical calibration algorithm used is the non-linear LS to obtain the best fit of the traffic flow-density data.

Overall, the number of studies suggests limited efforts into the calibration of widely used models in transport-related applications. Studies involving model-based calibration, the most adequate approach to such complex non-linear optimisation problem, are even scarcer. It is worth mentioning here two of the earliest attempts to calibrate two first-order models by W. Lin and Ahanotu (1995) and Hurdle and Son (2000). Both of which calibrated the fundamental diagram parameters (without any validation against other time periods) in a rather 'crude fashion'. For instance, W. Lin and Ahanotu (1995) calibrated only one parameter (backward wave speed) using traffic measurements while making assumptions on other parameters. They used their 'calibrated' parameters for CTM model simulation. Alternatively, Hurdle and Son (2000) made manual adjustment of one parameter (maximum traffic flow) to get good performance in the simulation of a simplified variant of the LWR model, referred to as Newell's model (Newell, 1993a, 1993c, 1993b). However, unlike any other calibration and validation study reviewed in this section, Hurdle and Son (2000) were the only ones who presented an analysis of individual errors in predicted and measured traffic density. They showed both the distribution of individual errors obtained for a road network and time period, and the relationship between the individual errors and their corresponding measurements. An illustrative example is provided in Figure 2.2 below. This is of importance

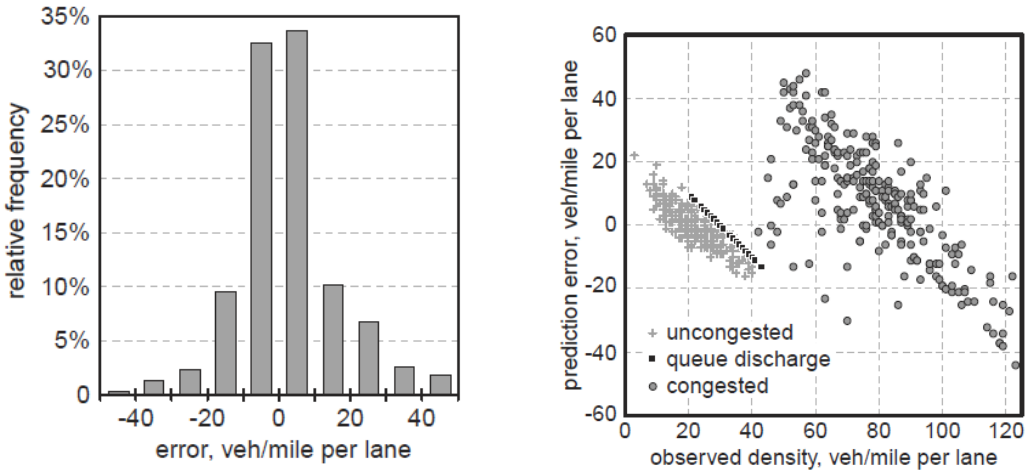


FIGURE 2.2: Analysis of individual errors in predicted and measured traffic density based on Newell's model (Newell, 1993a, 1993c, 1993b). Left side figure shows a histogram of individual errors. This is **Fig. 11** in Hurdle and Son (2000). Right side figure shows the relationship between individual errors and their corresponding measurements. This is **Fig. 12** in Hurdle and Son (2000).

since this research also considers analysing individual errors rather than a single-valued error term. The study of Hurdle and Son (2000) was also followed by that of Ni et al. (2004) who proposed a systematic approach to validating the model's accuracy (defined as zero prediction error mean) and precision (defined as small enough prediction error variance), using a statistical inference approach. Although their study did not report any calibration effort, it also motivates part of this research to use statistical testing in identifying 'good' parameter solutions.

Since second-order models involve parameters other than those of the fundamental diagram, the most commonly used calibration approach is model-based. The only exception is a study by Frejo et al. (2012) who used both to calibrate the parameters of the second-order model, METANET. Frejo et al. (2012) first calibrated the assumed fundamental diagram parameters (comparing three different shapes in the same study) using an SQP algorithm to minimise their objective function. This was followed by model-based calibration using an SQP algorithm to calibrate both the fundamental diagram and other parameters. The parameters obtained from the first calibration exercise were fed as initial guesses for the second step optimisation algorithm. Frejo et al. (2012) did not comment on the use of a gradient-based versus derivative-free methods.

Other model-based calibration studies have used different derivative-free algorithms and objective functions to calibrate the parameters of second-order models, as shown in Table 2.1. The models vary from one study to another in terms of the speed dynamics equation and the fundamental diagram assumed. Recently, the METANET model has gained popularity (Kotsialos et al., 2002; Poole & Kotsialos, 2012; Spiliopoulou et al., 2014, 2015; Poole & Kotsialos, 2016). Both Cremer and Papageorgiou (1981)

and Sanwal et al. (1996) used the derivative-free Complex Method of Box (Box, 1965) to calibrate their second-order model parameters. Cremer and Papageorgiou (1981) argued that gradient-based methods tend to find false optimum and require higher computation cost as a result of gradient computations. Kotsialos et al. (2002) also used the same algorithm to find a solution which can be used for a large network. They approached the problem by dividing their network into four stretches and applying the algorithm on each separately. This was followed by fine-tuning the results for the entire road network simulation. Alternatively, Ngoduy and Maher (2012) proposed a kernel-based CEM necessary to solve their continuous multi-extremal optimisation problem. They argued that their method can find the global optimal solution using a test scenario and that it provides good calibration results for real-world traffic dynamics. GA, CS, PSO algorithms have also been proposed as heuristic search methods for a similar problem, in two papers by Poole and Kotsialos (2012, 2016). While Poole and Kotsialos (2012) tested the performance of a GA algorithm on a single road network, Poole and Kotsialos (2016) tested and compared the performance of ten different GA, CS, and PSO algorithms on two road networks. They showed the superiority of two PSO implementations in comparison to others. This was accompanied by a study for Spiliopoulou et al. (2015) who tested and compared three algorithms: GA, CEM, and N-M. Their study reported a single solution for each calibration algorithm tested. The results showed similar parameter values, albeit not exactly the same. They also demonstrated the superiority of N-M in terms of computational time. Spiliopoulou et al. (2015) also highlighted the need to rerun the calibration several times with different initial guesses before selecting a good solution. However, a systematic approach to selecting such a solution has not been proposed. Spiliopoulou et al. (2014) also previously published a paper comparing the use of N-M algorithm for the calibration of the first-order CTM model and second-order METANET model. This is one of the studies which systematically compared both models in terms of their calibration and validation performance using real data. In each of these studies, a single fundamental diagram was assumed for their selected road network, except for Poole and Kotsialos (2016) who suggested an objective function where an automatic assignment of different fundamental diagrams along the discretised sections of their road networks, where necessary, is achieved. Irrespective of the model type used, the four following points have been observed.

1. Validation is almost always performed when using a model-based calibration approach but not when using the fundamental diagram-based approach.
2. All studies, except for Sanwal et al. (1996) and Poole and Kotsialos (2016), used a single road network to test their calibration and validation methodologies, neglecting the transferability of their methods to other road networks.
3. Except for a study by Hurdle and Son (2000), calibration and/or validation performance has always been assessed using:
    - qualitative methods such as visual inspection of measured and predicted time series data and/or space-time contour plots of measured and predicted

data; and/or

- quantitative methods using single-valued goodness-of-measures such as total percentage error or total mean absolute error.

Both of which are not rigorous enough to assess the validity of a model for further applications. Qualitative methods are often subjective. Single-valued quantitative measures fail to consider the distribution and the variability of errors in traffic variables.

4. A single solution has been reported whether the parameters involved are distinguished amongst cells or not. This indicates that even when studies acknowledge the solution obtained might not be a global solution but rather a 'good' local solution, studies do not exactly identify how many calibration runs were conducted to get to the solution, what criteria were followed to report a solution amongst other 'good' solutions, and why a single 'good' solution is reported if multiple runs can result with more than one 'good' solution.

2.3 Developments in vehicle exhaust emission modelling

Vehicle exhaust emission modelling is the process of estimating or predicting emissions of vehicle exhaust air pollutants for a particular road network and time period. Examples of vehicle exhaust air pollutants include *C*arbon *M*onoxide (CO), *H*ydro*c*arbon (HC), *Ox*ides of *N*itrogen (NO_x), and Particulate Matter. Emission models (in terms of structure, input, and parameterisation) are determined by emission factors and road traffic activity data. Emission factors specify the rate at which emissions are generated (Barth et al., 1996) typically as mass of an air pollutant per unit distance (e.g. grams per vehicle kilometre), mass per unit time (e.g. grams per second), or mass per unit fuel consumed (e.g. grams per kilograms of fuel consumed). They are parameterised functions typically derived for each air pollutant and vehicle category. However, they can also be derived for single vehicles or the entire vehicle fleet (Franco et al., 2013). As such, when they are derived for each vehicle category, necessary road traffic activity data needs to be specified for each vehicle category. When they are derived for single vehicles or the entire vehicle fleet, necessary road traffic activity data needs to be specified for single vehicles or the entire vehicle fleet. Depending on the emission model, a vehicle category can be defined by its class (e.g. passenger car, light duty vehicle, heavy duty vehicle...), fuel type (e.g. petrol, diesel, hybrid...), engine size, and emission control technology or standard (e.g. Euro 1, 2... for European vehicles). The dependent variables of emission factors' functions are often driving conditions-related (e.g. speed, acceleration, engine load...), and ambient and road conditions-related. These define the detail level of emission factors. Consequently, the required detail level of road traffic activity data is determined. For instance, functions developed based on second-by-second vehicle speed and acceleration data are highly detailed implying highly detailed

emission factors and emission models, and consequently highly detailed input requirements. Alternatively, functions developed based on trip-based vehicle speed data are relatively less detailed implying less detailed emission factors and emission models.

Functions of emission factors are developed (or updated), calibrated, and validated based on emission measurement data, typically for a sample of vehicles (not necessarily from a single source). Emission measurement data are collected based on different measurement techniques that keep developing with technological advancements. These are either laboratory-based or real-world-based; each of which has its own advantages and disadvantages, as will be discussed in § 2.3.1. Measurement data are also collected by different institutions in different countries and during different time periods. This implies that measurements are:

- updated regularly (e.g. more vehicles, more air pollutants, different and more realistic measurement techniques...); and
- dependent on local conditions and thus cannot be totally generalised (Johnson et al., 2000).

This also implies that values of emission factors (even of the same detail level) are dependent on where, when, and how the data used for their development were measured. More importantly, emission factors (and their detail level) are not normally calibrated by independent users unless they have access to measurement data (which is not always the case) or are able to collect their own (which is often time consuming and costly). As a result, while users can broadly highlight the uncertainty arising from measurement data, it is difficult to quantitatively study uncertainty arising from the calibration of emission factors and consequently, emission models. This is, of course, different from macroscopic traffic flow models for which their parameters need to be calibrated and validated (or at least assumed) by users.

Four factors that influence the emission factors' functions to be developed are:

1. The spatiotemporal resolution and detail level of emission measurements available for their development, calibration, and validation. Noting that, the more detailed they are, the more measurements they typically need.
2. The application or reason of development. This determines the spatiotemporal resolution of emission estimates or predictions. For instance, emission factors can be used for vehicle category-approval tests or as part of emission models. In the first case, emission factors are normally constants, comparable to vehicle-category emission standards. In the second case, the assessment of local or in-vehicle traffic management schemes requires a more detailed functional form in comparison to the assessment of regional traffic management schemes.
3. The development method used. This can be, for example, a simple averaging over all driving conditions, statistical modelling, clustering analysis, or even complex machine learning algorithms.

4. The availability of road traffic activity data necessary for either highly parameterised or low parameterised emission factors.

This indirectly means that there is not a single emission factor or emission model that gives definitive estimates for all places and times. The fact that emission factors (and their detail level) vary depending on several factors has given rise to numerous emission models. As such, emission models are often categorised. This section starts first by a brief review of currently available measurement techniques that any emission model is based upon. This is also followed with a brief review of the different categories developed so far. The review provides a background to the possible traffic flow and emission modelling integration approaches discussed in § 2.4.

2.3.1 Emission measurements

Emission measurement techniques are broadly categorised based on the location of measurements. These are either in laboratory controlled environment, or in real-world ambient environment. The first gives rise to laboratory-based emission measurements and the second gives rise to real-world-based emission measurements.

Laboratory-based emission measurements are performed under controlled conditions on either engine or chassis dynamometer facilities. Test operators hold control over the driving cycle being followed as well as over environmental conditions and other test-related conditions/parameters (Franco et al., 2013). A driving cycle is is mainly a laboratory-based terminology; it is a pre-defined speed profile over a time period typically representing different road type driving conditions. Official driving cycles mainly reflect country or region-specific driving pattern characteristics. Examples include the USA Federal Test Procedure (FTP) series of driving cycles and the New European Driving Cycle (NEDC) series of driving cycles. A comprehensive list of driving cycles can be found in Barlow et al. (2009). In 2011, global driving cycles started to emerge aiming at a more dynamic, realistic, and global test procedures such as the Worldwide harmonised Light vehicles Test Cycle (WLTC). Given 'continuous measurements' throughout a driving cycle, the output of a laboratory-based measurement technique is high-resolution emissions time-aligned with high-resolution driving characteristics (based on the driving cycle) and associated with vehicle characteristics (e.g. engine/vehicle type, size, standard, fuel type...). This indicates that there is a wide range of possibilities of the detail level emission factors can have for each vehicle category.

Laboratory-based techniques remain the most widely used technology for emission measurement and emission model development, as suggested by many such as Yu et al. (2009) and Smit and McBroom (2009). In fact, Franco et al. (2013) suggested that such measurements are expected to remain the major source of data used in emission models for the upcoming years. This is mainly as a result of:

- controlled conditions, standardisation, and reproducibility of laboratory testing; and
- complete emission measurements over a range of driving behaviour, for each engine/vehicle available in the market, and for relevant vehicle exhaust air pollutants.

However, such techniques have their own limitations (St-Denis et al., 1994; Barth et al., 1996; Yu et al., 2009; Ropkins et al., 2009; Smit et al., 2010; Franco et al., 2013). These are summarised below:

- equipment is costly and the test itself is too expensive to conduct and repeat which implies a relatively limited number of vehicles get tested giving rise to sampling bias;
- reluctance of high-emitter owners to provide their vehicles to testing, deliberate screening out of high-emitter vehicles, and/or deliberate testing of new or well-maintained vehicles which do not represent the maintenance conditions of the on-road vehicle fleet are all factors giving rise to sampling bias; and perhaps most importantly
- driving cycles and controlled conditions do not necessarily reflect real-world driving behaviour as well as ambient and road conditions. All of which are essential parameters to explaining vehicle emissions.

The development of more realistic driving cycles can partly resolve the last issue. However, recent research has shown evidence of the use of new and complex defeat devices in Volkswagen (VW) vehicles, giving rise to the VW Emission Scandal. Such devices "detect laboratory-based conditions and switch on emission control technologies which are otherwise, under real-world driving conditions, bypassed, disabled, or rendered inoperative" (US EPA, 2015). This adds up to the limitations of controlled laboratory-based measurement techniques that facilitate vehicle tampering acts, whether these are deliberate or by mistake/error. It also offsets the development of more realistic driving cycles.

Real-world-based emission measurements can further be distinguished into 'in-situ' and 'in-vehicle' (Ropkins et al., 2009). 'In-situ' measurements involve deploying measurement equipment at fixed roadside points with emissions of passing vehicles throughout a period of time being measured. They are typically obtained from *R*emote *S*ensing *D*evices (RSD) (referred to as across-road measurements in Ropkins et al. (2009)) (Bishop et al., 1989; Stedman, 1989; Guenther et al., 1991; FEAT, 2017, for more details on RSD), road tunnels, and inverse dispersion. The latter exists but is not widely used (Ropkins et al., 2009). RSD offers a better technology to laboratory testing in terms of considering real-world driving conditions and ambient conditions. It also permits for direct measurements from a large vehicle sample in a relatively short period of time. Characteristics of each passing vehicle can also be obtained based on video or camera recordings of license plates (at the time of data collection) cross checked with vehicle registration data. The technology has its own limitations as suggested by many

(Guo et al., 2007; Ropkins et al., 2009; Smit et al., 2010; Franco et al., 2013). These are summarised below:

- conversion assumptions need to be made as the RSD does not directly measure emission factors;
- emission measurements are spatially limited since RSD cannot perform measurements over the entire driving cycle of a particular vehicle but rather measures emissions at a single point of time (when the vehicle passes the equipment) - this implies that limited amount of data (driving conditions) is collected for any one vehicle as opposed to laboratory-based driving cycle tests;
- drivers decelerating (or changing their driving behaviour) when passing the equipment can induce measurement bias; and
- emissions cannot normally be measured at all times of the year since RSD is a fair weather technology (i.e. water sensitive). This can induce sampling bias since weather can have an influence on driving behaviour (and ultimately emissions produced).

Alternatively, road tunnel measurements are used to derive emission factors based on measured concentration levels given the advantage of being a relatively controlled environment. Using road tunnels involve calculating the difference in pollutants' concentration levels at the entrance and exit bores and correlating it with traffic flow, vehicle classes, tunnel features, wind conditions, and measured tunnel air flow. Road tunnels also allow measuring emissions from a large sample of on-road vehicles which represent real-world operation conditions at the location of the tunnel (El-Fadel & Hashisho, 2000). Despite that, this technique has its own limitations (John et al., 1999; El-Fadel & Hashisho, 2000; Corsmeier et al., 2005; Smit et al., 2010; Franco et al., 2013). These are as follows:

- fluctuating traffic flow parameters within the tunnels are normally restricted to short time periods;
- emission factors are indirectly calculated and so unaccounted external factors (e.g. the occurrence of 'piston effect' or the ventilation process within the tunnel which reduce the aerodynamics drag and thus reduce emissions substantially) can play a role in the concentration measurements increasing the level of uncertainty in the calculated emission factors; and
- emission factors derived from such measurements are restricted to aggregate levels as individual driving conditions are not incorporated in the measurements.

'In-vehicle' measurements involve deploying equipment, mainly *P*ortable *E*mission *M*easurement *S*ystem (PEMS), in instrumented vehicles with emissions being measured (typically of the same vehicles but can also be of nearby vehicles) as vehicles are driven with road traffic. PEMS measurements are analogous to laboratory-based ones. This means that they can be used to develop emission factors of different detail levels. At the same time, PEMS enables data collection under real-world conditions at any location travelled by the vehicle and under any weather condition as opposed to other

measurement techniques (Frey, Unal, Rouphail, & Colyar, 2003). Being a relatively new technology, PEMS has been questioned in terms of detection limits, data quality and accuracy, size and weight, and labour time and cost (Yu et al., 2009; Smit et al., 2010). Nevertheless, considering the potential of such system in developing, calibrating, and validating emission factors' functions, new and improved systems are regularly being developed and deployed.

Generally, real-world-based measurements are not standardised and reproducible in comparison to laboratory-based ones. While this can be a limitation, it reflects the complexity of vehicle exhaust emission performance in real-world scenarios. Nonetheless, measurements based on both are being used in the development and calibration of emission factors used for emission modelling. The categories of which are highlighted below.

2.3.2 Categorisation of models

A number of studies has been made to review and categorise existing emission models as well as to identify the advantages and disadvantages of each (Barth et al., 1996; Esteves-Booth et al., 2002; Boulter et al., 2007; Yu et al., 2009; Smit et al., 2010; L. Wismans et al., 2011; L. J. J. Wismans, 2012; Shorshani et al., 2015). While different criteria can be used as the base for categorisation, most commonly used ones are the the main input variables and/or scale of representation; both of which are also relevant to their integration with traffic flow models (since traffic variables and their scale of representation are their main inputs).

TABLE 2.2: summary of emission model categories and their respective modelling components

Model Type	**Main Input Variable**	**Scale of Representation**	**Emission Factor Detail Level**	**Model Example**
Aggregated	Road type	Static	Per road	NAEI[1]
Traffic situation	Distribution of traffic situation	Static	Per traffic situation	HBEFA[2]
Average speed	Average speed per trip/link	Dynamic macroscopic	Per average speed	COPERT[3]
Driving mode	Distribution of driving modes	Dynamic mesoscopic	Per driving mode	MOVES[4]
Instantaneous	Second-by-second speed-acceleration profiles	Dynamic microscopic	Per speed-acceleration value	PHEM[5]

[1] *N*ational *A*tmospheric *E*mission *I*nventory (Department for Environment Food and Rural Affairs, 2016).
[2] HandBook Emission FActors for road transport (HBEFA) (HBEFA, 2015).
[3] *CO*mputer *P*rogramme to calculate *E*missions from *R*oad *T*ransport (COPERT) (Ntziachristos et al., 2009).
[4] MOtor Vehicle Emission Simulator (MOVES) (Frey, Unal, Chen, & Li, 2003).
[6] Passenger car and Heavy duty vehicle Emission Model (PHEM) (Zallinger et al., 2008).

Table 2.2 above summarises the five main categories based on the 'main input variable' and their corresponding 'scale of representation' criterion. Each of these categories is briefly described below. The description is not meant to give an exhaustive review of currently available models. It rather attempts to draw upon aspects of interest which are the traffic flow data input and the advantages and disadvantages of each modelling category with particular emphasis on average speed-based emission model category.

Aggregated models are the least detailed and most basic. Emission factors used for such models are dependent only on the air pollutant, vehicle category, and road type (e.g. urban, rural, or motorway). The simplicity of such models results in their application to large (regional) road networks. This is at the expense of ignoring all the observed variations in emission measurements throughout the driving cycle itself and from one driving cycle to another. This of course prevents their use on a local level and for the assessment of local measures. These issues are partially resolved, the more complex and detailed the model is.

Traffic situation models are slightly less aggregated. Emission factors for such models are dependent on the air pollutant, vehicle category, and traffic situation. A traffic situation is a combination of road type, speed limit, as well as level of service (e.g. free flowing, stop and go, saturated...). The definition of what traffic situation encompasses is dependent on the model itself. Users need to be able to define each road according to the criteria set out for the model, which are often based on qualitative assessment of road traffic conditions. Boulter et al. (2007) suggested that the use of qualitative description to identify traffic situation is not ideal and is open to many different interpretations. Similar to aggregated emission model, they fail to directly consider traffic and vehicle dynamics and have a low spatiotemporal resolution. This restricts their use in assessing local measures or in-vehicle measures.

Average speed-based models are based on the concept that the average emission factor for a certain air pollutant and vehicle category vary according to trip-based average speed. These are typically constructed as follows, for a certain air pollutant and vehicle category:

1. Emission measurements are collected for a number of driving cycles.
2. For each driving cycle, average emission level in grams per distance travelled is calculated. Alternatively, a driving cycle can be divided in sub-cycles and their corresponding average emissions are calculated.
3. For each driving cycle (or sub-cycle), an average speed is also calculated.
4. Statistical modelling is used so average speed-based emission factors' functions are obtained.

Hence, the main input for such models is the average speed per vehicle trip and the distance travelled under each value. While the average speed input to this category of models is often trip-based, it is normally more convenient to use link-based average

speed. This is because link-based average speeds can be easily obtained from all traffic flow models (or loop detector measurements) while trip-based average speeds are restricted to microscopic traffic flow models (or only sample vehicle trajectory measurements). Hence, link-based average speeds are commonly used. Bai et al. (2007) studied the impact of this mismatch in the emission model requirements and available data. They showed that using link-based average speeds can result in higher emissions in comparison to trip-based ones, depending on the scenario. They also suggested that greater care is needed in studying the impact of such mismatch.

The use of average speed emission models (especially using link-based ones) is simple and efficient. This allows their application on large road networks and long time periods. They consider the dynamics of traffic in terms of trip-to-trip or link-to-link, unlike the static aggregated and traffic situation emission models described earlier. Hence, their scale of representation is dynamic macroscopic. Most importantly, in the absence of direct activity data, any type of traffic flow models can produce the necessary input requirements for emission calculations. However, the method suffers from certain limitations which can be resolved by more detailed mesoscopic and microscopic emission models described next. These limitations are mostly related to the model structure itself i.e. to the use of average speed only to define the emission performance, particularly:

- a single average speed can encompass different driving conditions of a vehicle within a trip or of vehicles on a link, which can possibly lead to different emission performance;
- the averaging does not allow considering high resolution individual-based emission prediction;
- there is a general lack of understanding of the spatiotemporal resolution required to define trip-based averaging (i.e. the length of vehicle trips on which the aggregation is applied) or link-based averaging (i.e. the length of the link and the time period used as the basis for aggregation). Only a recent study by Samaras et al. (2014) showed the limitation of using COPERT (in particular) for predicting fuel consumption using speeds averaged on less than 400 metres trips. They also noted that such conclusion can be different for other vehicle exhaust air pollutants, and similarly if other average speed-based models are used.

Driving mode models consider not only the average speed of a trip but also different modes during a vehicle trip. Emission factors are typically developed for each air pollutant, vehicle category, and driving mode in grams per second. Hence, the main input for such models is the distribution of driving modes and the time spent in each. Frey et al. (2002) and Frey, Unal, Rouphail, and Colyar (2003) argued that both second-by-second laboratory-based measurements and real-world-based measurements indicate that vehicle emissions are episodic in nature. They added that while previous categories of models are used to developing regional road traffic emission estimates, they lack the spatiotemporal resolution to properly characterise such microscopic episodic

events. As such, they developed a new driving mode model using real-world-based emission measurements data. Other studies have also proposed similar type of models such as Cernuschi et al. (1995), Frey, Unal, Chen, and Li (2003), and Zhu et al. (2013)[1]. Different studies use different data (laboratory-based or real-world-based) and different definitions of driving modes. This category of models improves upon the static and dynamic macroscopic emission model by incorporating essential parameters to the calculation of emissions, and improving the spatiotemporal resolution of calculated emissions. This is at the expense of higher detail level of input required for the emission models. It also aggregates data to driving modes and thus it still lacks the detail of the second-by-second instantaneous models.

Instantaneous models are the most detailed. They consider vehicle driving patterns at the highest resolution. Speed-based instantaneous models relate emissions to vehicle speed and acceleration during a driving cycle, typically on a second-by-second basis. Emission factors in speed-based models are defined for different combinations of instantaneous vehicle speed and accelerations. Alternatively, power-based instantaneous models relate emissions to engine power, in addition to vehicle speed and acceleration. The main input requirement for such models is the speed and acceleration profiles of individual vehicles. However, as the complexity of the model increases, more input becomes required. Generally, instantaneous emission models produce high spatiotemporal resolution of emissions. This is advantageous for the assessment of local policy scenarios. However, this is at the expense of a complex modelling structure, large input requirements, and highly parameterised model.

2.4 Emission predictions using traffic flow models

Road traffic activity data is the main input to emission models. While such input can be obtained directly from traffic measurements, the most common approach until now is the use of traffic flow models. The use of traffic measurements is often restricting in terms of their predictive power, and in terms of their spatial and temporal availability. They are used to estimate emissions of current scenarios or for real-time applications (Nyhan et al., 2016, for example). Even then, certain assumptions have to be made to allow their use for an entire road network and particular time period. The integration of traffic flow models with emission models is referred to as a 'modelling chain' or particularly, the 'traffic flow - emission' modelling chain.

The different categories of both traffic flow models and emission models described earlier have given rise to various integration approaches leading to such modelling chain. There exist two main considerations to the development of any integration approach. These are the detail level of emission predictions required, and the compatibility of

[1]Developed an average speed-based emission model dependent on the driving mode. While not entirely a driving mode model, driving modes were constructed based on emission measurement data.

the interface between the output of traffic flow models and the traffic flow input of emission models (Shorshani et al., 2015). Given the four categories of traffic flow models described in § 2.2, emissions can be predicted using any of the below categories:

- static traffic flow models;
- dynamic microscopic traffic flow models;
- dynamic mesoscopic traffic flow models; and
- dynamic macroscopic traffic flow models.

Static and dynamic microscopic traffic flow models have been widely used for emission predictions. Static models are often integrated with average speed-based emission models. This is a relatively straightforward approach since the average link-based traffic flow and speed outputs of static models can directly be used as input in average speed-based emission models. Static models are typically used within a transport modelling framework (Namdeo et al., 2002; Borrego et al., 2003, 2004; Gulliver & Briggs, 2005; Mumovic et al., 2006; Nejadkoorki et al., 2008; Bandeira et al., 2011). The main purpose of these studies is mapping emissions (Nejadkoorki et al., 2008), predicting air quality levels (Namdeo et al., 2002; Borrego et al., 2003, 2004; Mumovic et al., 2006; Bandeira et al., 2011), or even predicting exposure levels (Gulliver & Briggs, 2005). As a result of their low detail level and low development cost, they are often applied to large-scale road networks (Shorshani et al., 2015).

Alternatively, predicting emissions using dynamic microscopic traffic flow models result in the most detailed, yet most costly. Microscopic traffic flow models can directly be integrated with (speed-based or power-based) instantaneous emission models. Although in principle, they can be integrated with less detailed emission models. Both Fontes et al. (2015) and Osorio and Nanduri (2015) provided a detailed review of studies which have followed such integration approach. Fontes et al. (2015) suggested that studies which used microscopic traffic flow models are often limited to predicting emissions with very few ones following up to predict air quality levels. They argued that this is because of difficulties (or additional effort required) in deriving aggregate link-based rather than vehicle-based emission predictions typically required by air quality models. Osorio and Nanduri (2015) also commented on the complexity and computational inefficiency of such integration approach that limits their use to 'what-if' analysis rather than optimisation problems (such as traffic control or signal optimisation problems which consider environmental objectives).

Using mesoscopic traffic flow models to predict emissions is rarely achieved. Perhaps the only example is that of Bai et al. (2007). They used a *D*ynamic *T*raffic *A*ssignment (DTA) framework relying on a mesoscopic traffic flow model. The nature of the outputs of a mesoscopic traffic flow model allowed them to obtain both trip-based and link-based average speeds. This served their purpose in deriving and comparing trip-based average speed-based emissions with link-based average speed-based emissions.

In the following two sections, a comprehensive review is provided on studies which have focussed on the last approach which is emission predictions using dynamic macroscopic traffic flow models. Their use to predict emissions normally attempts to improve on the low-fidelity of static models while avoiding the high costs and inefficiencies associated with microscopic ones. Based on § 2.2, the reviewed studies are divided into those which deployed first-order LWR-type models (§ 2.4.1), and those which deployed second-order Payne-type models (§ 2.4.2). Regardless of the dynamic macroscopic traffic flow model used, the review focusses on three key points of interest:

1. The objective of predicting emissions using a macroscopic traffic flow model.
2. The calibration and validation efforts of the utilised macroscopic traffic flow model.
3. The case study used to examine the integration approach (i.e. real/hypothetical and urban/motorway/regional).

The first point allows to appreciate the wide range of applications which depend primarily on the 'traffic flow - emission' modelling chain. The last two points allows to understand to what extent research studies have taken into account the real-world representation and application of their methodologies. Both of which illustrate the motivation behind this study; the need to better understand uncertainties and errors in the modelling chain.

2.4.1 Using first-order LWR-type models

Table 2.3 below summarises ten studies found to have integrated first-order LWR-type models with emission models. In general, each study focusses on a single air pollutant (e.g. CO, HC, NO_x...) when presenting the methodology or the results. However, any proposed methodology is typically applicable to other air pollutants. Once the category of emission model is specified, it is only the functional form of emission factors which would change with the change in the air pollutant. As shall be seen, this reflects the fact that most of the reviewed studies have focussed on developing the methodology rather than examining the performance of traffic flow and emission models with respect to the real-world. The objective behind using first-order LWR-type models differs from one study to another. One study by Zhou et al. (2015) was found to have used Newell's model, a simplified variant of the LWR model, for emission predictions. However, it is clear that CTM is the most commonly used model for predicting emissions.

Two studies have used a multi-class CTM for predicting emissions. For instance, Chiou et al. (2013) used a multi-class (passenger cars and motorcycles) model which was developed and validated earlier in Chiou and Hsieh (2012) [this study was written in Chinese and thus not reviewed here]. Also, Liu et al. (2014) used FASTLANE, a software implementation of multi-class first-order LWR-type model (Van Lint et al., 2008).

TABLE 2.3: review of first-order LWR-type traffic flow - emission modelling chain studies

Reference	Integration Approach	Traffic Flow Model	Emission Model Category	Air Pollutant	Study Objective	Calibration and Validation	Case Study
J. Lin and Ge (2006)	Direct	CTM	Average speed-based	CO	Air quality prediction	-	Hypothetical (1) Urban
Aziz and Ukkusuri (2012)	Direct	CTM	Average speed-based	CO	DTA[1]	-	Hypothetical (2) Urban
Chiou et al. (2013)	Direct	Multi-class CTM	Average speed-based	CO	Air quality prediction	-	Hypothetical (1) Urban
Rehimi and Landolsi (2013)	Direct	CTM	Average speed-based	CO	Air quality prediction	-	Hypothetical (1) Urban
Zhu et al. (2013)	Indirect	CTM	Driving mode and average speed-based	CO, HC, NO_x	Models' Comparison[2]	-	Real (1) Urban
L. Zhang et al. (2013)	Indirect	CTM	Driving mode	CO	Traffic control optimisation[3]	-	Real (1) Urban
Samaranayake et al. (2014)	Direct	CTM	Average speed-based	Any	Real-time air quality prediction	-	Real (1) Motorway
Liu et al. (2014)	Indirect	Multi-class CTM with VSL	Speed-based instantaneous	CO, HC, NO_x	Traffic control optimisation[4]	-	Hypothetical (1) Motorway
Zhou et al. (2015)	Indirect	Newell's model	Driving mode	CO, CO_2, HC, NO_x	DTA[5]	-[6]	Real (2)[7] Regional
Han et al. (2016)	Indirect	CTM	Power-based instantaneous	HC	Traffic control optimisation[3]	-	Hypothetical (1) Urban

[1] DTA with travel time and environmental objectives.
[2] CTM, shockwave model, point queue model, and car-following model.
[3] Signal timing with travel delay and environmental objectives.
[4] Ramp metering and VSL with travel time and environmental objectives.
[5] DTA for environmental assessment.
[6] Validation of derived accelerations only.
[7] One motorway and one regional.

There exist a 'direct' approach and an 'indirect' approach to predicting emissions using first-order LWR-type models. The 'direct' approach is defined here when the output of the traffic flow models used can directly be used as input for the selected emission model. In the case of macroscopic traffic flow models, this is when they are integrated with average speed-based emission models. However, this is under the assumption that the emission model allows for high spatiotemporal resolution of link-based average speeds as input.

Alternatively, the 'indirect' approach requires an intermediary step so that the aggregate (yet high spatiotemporal) outputs of first-order LWR-type models can be used as input to emission models. The 'indirect' approach is required when the emission model is either a driving mode or an instantaneous one. A review of studies using the 'direct' and 'indirect' approach is provided here.

Perhaps the earliest study to directly predict space-time varying average speed-based emissions using a first-order LWR-type model is that of J. Lin and Ge (2006). They illustrated the possibility of using the 'direct' approach in order to predict not only emissions but also air quality levels on urban road networks at a high spatiotemporal resolution. To predict air quality levels, they used emission predictions as input to their utilised dispersion (or air quality) model. Chiou et al. (2013) and Rehimi and Landolsi (2013) used the same approach for the same purpose of predicting air quality levels on urban road networks. However, the average speed-based emission model utilised is dependent on the study. While Chiou et al. (2013) also studied the influence of different signal timing plans on air quality levels, Rehimi and Landolsi (2013) studied the influence of different wind scenarios on air quality levels. In addition, Rehimi and Landolsi (2013) compared hourly air quality levels obtained from CTM with those obtained when traffic dynamics were not considered (simply this meant that traffic is equal to the traffic demand along the entire road). They suggested that during traffic congestion, there are larger differences in air quality levels especially when the wind is parallel to road traffic. Although, they referred to such differences as 'errors', this can be problematic since a comparison with 'ground-truth' air quality levels is not made. Each of these studies demonstrated their approach using a single hypothetical road network. They also have not made any calibration or validation of the traffic flow model i.e. the parameters were always pre-defined.

Samaranayake et al. (2014) developed the same 'direct' integration approach for the real-time estimation of air quality levels on motorways. They demonstrated their method through a prototype implementation for the San Francisco Bay Area in the US. Although an example of results for a real network and real data is provided, the calibration and validation process of the traffic flow model deployed was not provided. They did mention, however, that their system aimed at using models which have been technically validated and accepted by practitioners. This also implied that they did

not attempt to derive acceleration or driving modes from the traffic flow model due to difficulties in measuring and validating them in practice.

Aziz and Ukkusuri (2012) also integrated CTM with an average speed-based emission model but for a different purpose. They developed a DTA framework whereby not only travel time but also emissions predicted by the CTM are to be minimised. They also compared three different scenarios. These are DTA with total travel time minimisation objective, DTA with total emissions minimisation objective, and DTA with both total travel time and total emissions minimisation objective. By converting total travel time and emissions to monetary values, they showed that including both produces the minimum cost in comparison to the other two scenarios. In using such DTA framework, they highlighted the possible use of their methodology to predicting and assessing air quality levels at high spatiotemporal resolution, based on the optimal solution. However, their framework was only applied to two hypothetical scenarios without any real road networks or real traffic data. These are a small hypothetical network, and another commonly used one referred to as the 'Nguyen and Dupuis' network (Nguyen & Dupuis, 1984).

'Indirect' integration approaches have been developed in Zhu et al. (2013), L. Zhang et al. (2013), Liu et al. (2014), and Han et al. (2016). Zhu et al. (2013) developed their own average speed-based emission model using real-world-based emission measurements. Their emission measurements were split into four different driving modes before developing the emission model. This resulted in average speed-based emission factors' functions for each of the idling, cruising, acceleration, and deceleration driving modes. As such, traffic outputs from the CTM needed to be split according to the driving mode before their utilisation as input to their average speed-based emission model. In order to do so, Zhu et al. (2013) suggested that the time evolution of traffic speed on each cell can be used as an indicator of the driving mode. L. Zhang et al. (2013) also extracted driving modes from the CTM outputs. They directly used a driving mode emission model developed by Frey et al. (2002) and Frey, Unal, Rouphail, and Colyar (2003). One main difference between these two studies is their derivation method of the driving mode from CTM output. While Zhu et al. (2013) used the time evolution of traffic speed at each cell, L. Zhang et al. (2013) assumed that vehicles are moving downstream in time and thus driving modes need to be determined by the difference in traffic speed in space and time rather than only time. Liu et al. (2014) combined both concepts together. They considered that acceleration (or deceleration) depend on whether vehicles remain in the same location or move downstream with time. So they distinguished between cross-cell and inter-cell acceleration. However, they used traffic speed and acceleration directly with a speed-based instantaneous emission model without converting the output to a driving mode. In each of the three studies, deriving acceleration (or driving mode) based on a macroscopic traffic flow model implies that these are not individual accelerations but rather for a number of vehicles (i.e. group of vehicles is assumed to have the same driving mode).

Perhaps the most complex integration approach is the one developed in Han et al. (2016). They used a power-based emission model to calculate CTM-based emissions. In addition to cell-based road gradients, they derived aggregate-based accelerations from the CTM traffic speed outputs. However, they used multiple test simulations in order to construct linear relationships between link occupancy and total power-based emissions on a link. Han et al. (2016) suggested that using such relationships can make it easier to incorporate emission objectives in signal optimisation in comparison to the direct use of instantaneous emission predictions in the optimisation problem. However, they acknowledged the uncertainty of using such an approximation because of the inevitable variation around the linear fit. They thus included such uncertainty within their robust optimisation framework.

The main issue in these four studies is neglecting the possibility of obtaining unbounded accelerations from first-order LWR-type models, especially when applied to real road networks with real traffic data. This is also particularly critical when using instantaneous emission models (Liu et al., 2014; Han et al., 2016). This issue might not necessarily arise in the case of using driving modes instead of absolute acceleration values.

Each of these four studies had different purposes. Zhu et al. (2013) calculated both total delays and emissions of different air pollutants at a signalised intersection using the CTM and compared the results with those obtained from a shockwave model, point-queue model, and car-following model. They considered the car-following model as the 'ground-truth'. They concluded that CTM (unlike the shockwave and point-queue models) provided close estimates to those based on the car-following model. L. Zhang et al. (2013) used driving mode emission predictions to predict air quality levels and human exposure. They formulated a bi-objective optimisation framework to determine signal timings which minimise both traffic delay and the risk associated with human exposure to traffic emissions. L. Zhang et al. (2013) argued that while such framework can be developed based on microscopic traffic flow models, these are computationally intensive especially for real networks. Han et al. (2016) also formulated a bi-objective signal optimisation to simultaneously minimise traffic delay and linearised instantaneous power-based emission predictions. Liu et al. (2014) used instantaneous emission predictions for traffic control optimisation purposes. Their traffic control represented both ramp metering and VSL strategies. Accordingly, a VSL version of the utilised multi-class CTM was proposed. Their optimisation problem involved the minimisation of both total travel time and emissions over a road network. Liu et al. (2014) compared their optimal solution based on the multi-class CTM with the optimal solution from uni-class CTM as well as the no-control scenario. They concluded that further work can involve the comparison of such framework with one which utilises a second-order Payne-type model instead. In each of these four studies, only test networks have been used to demonstrate the developed methodology. This implied that calibration and validation of the utilised first-order LWR-type models were not

necessary. Although Zhu et al. (2013) used a real road network, their traffic data was assumed.

In addition to that, Zhou et al. (2015) developed an integration methodology to predict driving mode-based emissions using Newell's model, similar to the one used in Hurdle and Son (2000). They proposed its use with a simple linear car-following model (Newell, 2002) to generate vehicle trajectories (i.e. vehicle speed and acceleration). Vehicle trajectories then allowed them to calculate Vehicle Specific Power (VSP) to be used as input to a driving mode emission model. Although they did not show any calibration and validation efforts of the utilised traffic flow models, they did highlight preliminary results of a validation exercise of their derived acceleration output (based on vehicle trajectories). Their preliminary results showed the distribution of acceleration values in comparison to a sample distribution of real-world acceleration data obtained for one of the real road networks used as their case study.

2.4.2 Using second-order Payne-type models

Fewer studies have attempted to predict emissions using second-order Payne-type models in comparison to first-order LWR-type models. Table 2.4 below provides a summary of the reviewed studies. Most commonly used second-order Payne-type model for emission prediction is METANET. It has been used to directly predict average speed-based emissions (Csikós et al., 2014, 2015) and to indirectly predict instantaneous emissions (Zegeye et al., 2009, 2010, 2011, 2013; Groot et al., 2011b, 2013; Liu et al., 2013). The main purpose of the integration approach in Csikós et al. (2015) is the prediction of air quality levels under different motorway traffic control scenarios. Their framework has only been verified using a test network and through a sensitivity analysis. The sensitivity analysis involved studying the effect of different ramp metering and VSL input scenarios on air quality levels. In Csikós et al. (2014), a traffic control (mainly of ramp metering) optimisation problem with environmental constraints has been proposed using the same air quality modelling framework discussed in Csikós et al. (2015).

Alternatively, a research group from Delft University of Technology has published a number of papers aimed at developing a model-based motorway traffic and emission control using either METANET (Zegeye et al., 2009), multi-class METANET (Liu et al., 2013) originally developed in Deo et al. (2009), or piece-wise-affine approximation of METANET (Groot et al., 2011b, 2013). The latter is used in order to deal with the computational complexity of the original non-linear and non-convex model-based control framework. Traffic control was in terms of ramp metering and VSL strategies. This implied that their METANET model structure was modified to take the influence of VSL into consideration.

TABLE 2.4: review of second-order Payne-type traffic flow - emission modelling chain studies

Reference	Integration Approach	Traffic Flow Model	Emission Model Category	Air Pollutant	Study Objective	Calibration and Validation	Case Study
Zegeye et al. (2009, 2010, 2011) and Groot et al. (2011b, 2013)	Indirect	METANET with VSL	Speed-based instantaneous	CO, CO_2, HC, NO_x[1]	Traffic control optimisation[2]	-	Hypothetical (1) Motorway
Liu et al. (2013)	Indirect	Multi-class METANET with VSL	Speed-based instantaneous	CO, HC, NO_x	Traffic control optimisation[2]	-	Hypothetical (1) Motorway
Zegeye et al. (2013)	Indirect	METANET with VSL	Speed-based instantaneous	CO, HC, NO_x[1]	Emission prediction[1]	-[3]	Real (1) Motorway
Csikós et al. (2014)	Direct	METANET	Average speed-based	CO, HC, NO_x	Traffic control optimisation[4]	-	Hypothetical (1) Motorway
Csikós et al. (2015)	Direct	METANET with VSL	Average speed-based	CO, HC, NO_x	Air quality prediction	-	Hypothetical (1) Motorway

[1] With fuel consumption.
[2] Ramp metering and VSL with travel time and environmental objectives.
[3] Validation of emission predictions with those from a microscopic traffic flow model.
[4] Ramp metering with traffic and environmental objectives.

Zegeye et al. (2010, 2011) took a step further and included air quality levels in their model-based control framework. The integration approach used for emission prediction has been explained in details in Zegeye et al. (2013). Their indirect integration approach involved deriving useful acceleration data from the aggregated traffic speed output of METANET. These are then used as input to a speed-based instantaneous emission model. Deriving acceleration data from a second-order model resolves the issue of unbounded acceleration that could otherwise be obtained when using a first-order LWR-type model. The main concept of their integration approach is that acceleration data cannot be retrieved for individual vehicles but only for aggregate number of vehicles. Zegeye et al. (2013) studied analytically the propagation of uncertainty from such an approximation to emission predictions. They also particularly noted that the errors induced by the traffic flow model itself are not included in their analysis.

Amongst all the above-mentioned studies, only Zegeye et al. (2013) attempted to compare their integration approach with average speed-based emission predictions using METANET traffic flow model, and speed-based instantaneous emission predictions using a microscopic traffic flow model. Emission predictions from the microscopic traffic flow model were considered as the 'ground-truth'. Zegeye et al. (2013) showed that their integration approach provides better emission predictions than those obtained from the use of an average speed-based emission model for the case study considered. There exist two main issues in such a conclusion, reflecting the complexity of the problem:

1. The microscopic traffic flow model and instantaneous emission model deployed are not entirely free-of-errors, although they are often considered as the 'ground-truth'.
2. The average speed-based emission model is entirely different from the instantaneous emission model, not only in terms of the category of each emission model but also in terms of measurement data used for their development. This means that there are multiple factors playing a role in the obtained differences in emission predictions between each of the two approaches and the assumed 'ground-truth'.

Going back to the scope of this research, four points are highlighted here based on the above review of both first-order LWR-type and second-order Payne-type models:

1. Numerous integration approaches have been developed for predicting emissions using macroscopic traffic flow models; some of these approaches are direct ones, and other are indirect (i.e. they require an intermediary step for the output of the traffic flow model to match the input of the emission model).
2. Most commonly used macroscopic traffic flow models are the first-order CTM and second-order METANET.
3. Average speed-based emission modelling is a typical category for models integrated with both first-order CTM and second-order METANET without any study-dependent intermediary steps.

4. Emission predictions using traffic flow models are not only useful on their own but also for various further applications such as traffic control optimisation, air quality predictions, and exposure level prediction.

As indicated in Table 2.3 and Table 2.4, all the studies have focussed on developing the integration approach itself and the possible applications. None of the studies have attempted to calibrate and validate the parameters of the traffic flow models (or at least reported such work). The main reason is their use of hypothetical networks and thus assuming certain parameters was considered sufficient. Where real networks are considered, calibration and validation process is typically referred to other studies. In addition to that, many of the reviewed studies have used a VSL modified version of first-order and second-order models. Although few studies have focussed on studying the impact of VSL on traffic conditions, none of the reviewed studies in § 2.2.2 involved the calibration and validation of such modification to the models. Since one of the road networks selected in this study is VSL-operated, this will be expanded upon when describing the mathematics behind selected traffic flow models in Chapter 3.

It is clear, that in any developed methodology, the results were limited to single-valued emission predictions without any consideration of uncertainties and errors arising in the modelling chain that has been developed. Initial attempts to consider certain sources of uncertainty has been made by Zegeye et al. (2013) and Han et al. (2016). However, this was rather focussed on sources arising from their intermediary steps to integrate a macroscopic traffic flow model with an emission model. Section § 2.5 expands more on the concept of uncertainty in emission predictions.

2.5 Uncertainty in emission predictions

Before highlighting the main work in the study of uncertainty in emission predictions, it is essential to provide an overview of the concept of uncertainty. This section thus starts by introducing the terminology and different dimensions of uncertainty followed by a review of main studies which have contributed to the field.

Defining the concept of uncertainty is not entirely straightforward, as definitions of relevant terminologies and concepts depend on the modelling system under study but also on whether it is the point of view of modellers or policy makers (Walker et al., 2003). In fact, Refsgaard et al. (2007) added that there is a philosophical aspect to it that assists in such system or even study-dependent definitions. In an effort to determine a common terminology and typology of the concept of uncertainty in a model-based decision support system, Walker et al. (2003) defined uncertainty from the modeller's point of view (of interest to this study) as 'any deviation from the unachievable ideal of completely deterministic knowledge of the relevant system'. The 'traffic flow - emission' modelling chain is of course an example of a model-based decision support system. The

latter is mainly defined as the use of models and their outcomes to support policy and decision making. Based on their review, Walker et al. (2003) identified three dimensions to the concept of uncertainty. These are the location, level, and nature of uncertainty. The location of uncertainty is basically the source which can be categorised into five:

1. *Context* which reflects any ambiguity in the definition of the boundaries of the system to be modelled.
2. *Model structure* uncertainty.
3. *Model implementation or technical* uncertainty.
4. *Model input* uncertainty.
5. *Model parameter* uncertainty which can be an exact, fixed, a priori, or calibrated parameter.

The last four reflect any misspecification in the functional form, implementation, input, or parameter. Walker et al. (2003) highlighted that uncertainty in model structure and model parameters are not entirely independent. For instance, a simple model with few parameters can have higher uncertainty in structure but not in parameters, while a complex model with many parameters can have higher uncertainty in parameters but not in structure. This can also be the case for model input where the uncertainty can be higher with more complex and input-demanding models. The total uncertainty from the different sources and their interactions gives rise to *model outcome* uncertainty. In the case where the true outcome values are known, such uncertainty is referred to as *prediction error*. This is mainly defined as the deviation of a predicted outcome from a true value, and is of particular interest in this study. Of course, these sources can arise in a single model but also in a chain of models. The additional complexity of a modelling chain is that the sources of uncertainty in the first model are reflected in the input uncertainty of the second model of the chain, and so forth, depending on how many models are involved. An additional source of uncertainty that arise in a modelling chain is the compatibility of the output of the first model with the input of the second model, and so forth.

The other two dimensions of uncertainty are its level and nature. While the level of uncertainty can range from the (hardly achievable) complete understanding to total ignorance, the nature of uncertainty can be related either to lack of knowledge or to inherent variability. Lack of knowledge is often referred to in literature as epistemological uncertainty, while inherent variability is sometimes referred to as aleatory (or ontological) uncertainty.

In addition to these three dimensions, the fourth dimension which was not the focus of Walker et al. (2003) is the tools available to studying uncertainty. Morgan et al. (1992) suggested three tools to achieving that in a model-based decision support system, each of which serves different purposes. As briefly highlighted in § 1.1, these are:

1. Sensitivity analysis which computes the effect of changes in the sources of uncertainty on model predictions.
2. Uncertainty propagation which computes the uncertainty in the model outputs induced by the uncertainty in its sources.
3. Uncertainty (contribution) analysis which compares the importance of the sources of uncertainties in terms of their relative contributions to the model output uncertainty.

Cullen and Frey (1999) highlighted three main purposes to sensitivity analysis and uncertainty analysis. These are providing awareness about the inputs which influence the results, promoting improved understanding and interpretation of the results and the modelling system, and prioritising future research on aspects which needs further investigation. Uncertainty propagation, however, is the main tool to probabilistic-based modelling, which is the focus of this research. Probabilistic-based modelling is expressed as probability distributions in both the sources and the modelling outcome. Such tool allows evaluating uncertainty in the final outcome by identifying the possible range of its value and possibly its extremes, depending on the deployed method (Cullen & Frey, 1999). This ultimately leads to better targeting of data collection and further research that can reduce the modelling outcome uncertainty (Cullen & Frey, 1999).

Dealing with uncertainty, in the context of emission modelling, has indeed been recognised. For instance, the air pollutant emission inventory guidebook (European Environment Agency, 2016) recommends that wherever possible, an estimate of uncertainty associated with emission factors or activity needs to be reported; this is of course not only for the vehicle exhaust emission source, but any other emission source (such as industrial sources, agricultural sources...). While some efforts has been made in using sensitivity analysis and uncertainty contribution analysis to better understand uncertainty in vehicle exhaust emission predictions (Kioutsioukis & Tarantola, 2003; Kioutsioukis et al., 2010, and references therein), one of the early attempts to probabilistic-based emission predictions is a study by Kini and Frey (1997), which is also summarised in Cullen and Frey (1999) as a case study. Kini and Frey (1997) developed a probabilistic version of an emission model referred to as MOBILE5 which enables the propagation of model input uncertainty to the model outcome. MOBILE5 is a commonly used average speed-based emission model developed by the US *E*nvironment *P*rotection *A*gency (EPA) using laboratory-based emission measurements. The probabilistic version allows expressing model input as probability distributions rather than point-estimates. Using numerical simulation techniques such as Monte Carlo simulation or Latin Hypercube Sampling (LHS), a probability distribution of the model outcomes can be obtained. Such powerful techniques allow for empirical multivariate distributions and are not restricted to parametric univariate distributions. The main disadvantage of such methods, however, is the sampling size and the implied model simulations required to study uncertainty. For probabilistic-based modelling, it is essential to be able to quantify the uncertainty in model inputs to be propagated, preferably

using quantitative analysis rather than qualitative judgment. Kini and Frey (1997) and Frey and Zheng (2002) quantified the empirical distributions of inter-vehicle variability and uncertainty in emission correction factors used in their emission model (e.g. ambient temperature correction factor). Their quantification was based on laboratory emission measurement data of LDVs only, which they have acknowledged its limitation in terms of their real-world representativeness.

Alternatively, Kühlwein and Friedrich (2000) highlighted that uncertainty in predicted emissions from road traffic using complex methods and models has largely been unknown. They have thus focussed on studying the influence of model input uncertainty on emission predictions of a traffic situation emission model referred to as IER. Uncertain model input included traffic flow, driving pattern mix, fleet mix, emission factors, and road gradients. However, their traffic data was not the output of a traffic flow model but rather traffic count data. So, uncertainty in traffic data is basically related to measurement errors and extrapolation errors from sample to annual data. For uncertainty propagation, Kühlwein and Friedrich (2000) proposed an approximation method based upon Taylor series expansion, which they referred to as 'statistical error propagation' method. Their approximation was based on propagating the mean and the variance (the second central moment) of each model input to calculate the mean total error in emission predictions. More details on the method in the context of uncertainty propagation can be found in Cullen and Frey (1999). Unlike numerical simulation methods, the main disadvantage of such method is the assumption of independence between the model inputs which is often not the case. Also, they do not allow for characterising the probability distribution of model outcome but only for its mean (and possibly its central moments).

The above studies show that while probabilistic-based emission modelling has been conducted, it is often based on a single model rather than a modelling chain. Traffic data is thus based on measurements and expert assumptions rather than outputs of a traffic flow model which leads to more complex interactions and possibly larger uncertainty in the modelling outcome of emission predictions. Shorshani et al. (2015) provided a comprehensive review of the 'traffic flow - emission - air and water quality' modelling chain whereby they highlighted the categories of traffic flow models, emission models, and possible integration approaches between the two. This recent study also highlighted the limitations and general sources of uncertainty in emission modelling and traffic flow modelling. Shorshani et al. (2015) concluded that while uncertainty in model outcomes have been investigated for isolated models (i.e. either emission models or traffic flow models), its propagation throughout the modelling chain is yet to be investigated.

2.6 Summary

This section summarises the findings of this chapter in relation to the four points introduced in § 2.1 and summarises the gaps which this research aims to contribute to.

To date, numerous mathematical models of traffic flow have been developed. In addition to static models, mathematical models describing the dynamics of traffic flow are often categorised under either the microscopic, mesoscopic, or macroscopic models. These correspond to the detail level considered in the modelling structure. Generally, the more detailed the model, the more input is required and the more parameters need to be calibrated and, of course, validated. Under each category, many models are developed in an attempt to resolve certain issues found in other models. Macroscopic traffic flow models are advantageous in terms of the input required for their simulation and for the calibration of their parameters, which are relatively few. Although they do not describe the dynamics of traffic at individual levels, they are characterised by their high spatiotemporal resolution. While the concept behind macroscopic modelling category (i.e. the analogy of traffic flow to fluids) has been challenged by few, the main debate or controversy has been focussed around the superiority of its sub-categories: first-order or higher-order models. While a lot of theoretical work has been seen around the debate, limited efforts have targeted the issue quantitatively in relation with their real-world performance. This is despite the availability of measurement data required to do so (although not without its limitations), and their wide use in various applications. Examples of such applications are traffic control/signal optimisation, evacuation planning, incident detection, and most importantly in this research, emission prediction. As such, the interesting debate between the first-order and second-order traffic flow models, the availability of traffic data for their calibration and validation, and their application in emission prediction are considered the main reasons to selecting macroscopic traffic flow models in this research.

Studies found on the calibration and validation of macroscopic traffic flow models have been reviewed. Different studies typically use different models, different optimisation algorithms, and different case studies. Most commonly used models are the first-order CTM and second-order METANET. Many optimisation algorithms have been suggested, mostly heuristic search methods. These are preferable in solving non-linear optimisation problems where real traffic data is used to avoid falling into a 'false' solution. However, even with the use of such methods, the solution most often depends on the starting point. The solution is also not guaranteed to be a 'global' one, if it exists. Normally, qualitative assessment such as visual inspection of results and/or quantitative assessment using single-valued goodness-of-fit measures are used to demonstrate the validity of the solution. Two of the main gaps found in this review motivates two of the main contributions in this research. First, studies have often reported a single solution selected based on non-rigorous assessment, despite the uncertainty around the reported solution. As such, an ensemble-based approach to the optimisation problem

with statistical inference is proposed in the research methodology. Second, studies often focussed on a single non-VSL road network as their case studies. As such, the research methodology is applied to three road networks, one of which is a VSL-operated road network, in order to both ensure transferability of results and to compare between the basic model structures of macroscopic traffic flow models and their VSL-modified versions.

In order to better understand the application of macroscopic traffic flow models in predicting emissions, emission models were reviewed. Emission models mainly depend on emission factors and traffic data. The detail level of traffic data depends on the detail level of emission factors. For instance, those developed for each trip-based average speed are different from those developed for each second-by-second vehicle speed and acceleration. As such, emission models normally fall under five different categories. These are aggregated, traffic situation, average speed, driving mode, or instantaneous emission models. Many models are developed under each category. Each model utilises different emission factors. This is because they are dependent on the location, time, and technique used to measure emissions of sample vehicles which form the basis of their development, calibration, and validation.

Different studies have shown that there are different approaches to integrating the outputs of macroscopic traffic flow models with emission models. These can be categorised to direct and indirect ones. Direct approaches imply compatibility between the outputs of traffic flow models and the inputs of emission models. Indirect approaches require an intermediary step for these to match. Commonly used models in the direct approach are the CTM/METANET and average speed-based models, while those used in the indirect one are the CTM/METANET and driving mode or instantaneous models. While the integration approaches between macroscopic traffic flow models and emission models are well-established, the main gap found in research studies is their disregard to the uncertainty arising in their real-world application. Only two studies attempted to study uncertainty arising from their intermediary steps rather from the deployed traffic flow models. As such, all the reviewed studies develop methods which can report point-based prediction of emissions at the resolution of the traffic flow model used.

This has been also noted in § 2.5 which introduced the concept of uncertainty, its definition and typology, and highlighted main studies which attempted to provide probabilistic-based predictions of emission. A number of insights have been found from this review. First, sources of uncertainty can generally be context-wise, model structure, model implementation, model input, model parameters, and model outcome (referred to as prediction error when the true values exist). The main tool to probabilistic-based modelling is uncertainty propagation. In the context of emission modelling, uncertainty propagation are often dealt with through numerical simulation methods or approximation methods, such as Taylor series expansion. Numerical

simulation methods are advantageous as they allow the use of empirical multivariate distributions, but are also flexible in terms of the complexity and nonlinearity of the models involved. It was found from the review of § 2.5 that uncertainty in emission predictions were often focussed on uncertainty in emission factors or direct measurements of traffic data, rather than those obtained from traffic flow models, despite the common integration of traffic flow models with emission models. As such, this gap motivates developing a research methodology which quantifies the uncertainty in macroscopic traffic flow model outcomes and propagates these to emission predictions. Because the first-order CTM, second-order METANET, and average speed-based emission models have been found to be widely used in § 2.4, this research focusses on average speed-based emission predictions using CTM and using METANET, separately. Starting with these three models provide the basis to extending such work to more complex models and more complex integration approaches.

Chapter 3

Research Methodology

3.1 Introduction

In Chapter 2, the reasoning behind studying uncertainty in the modelling chain of two widely used, yet debated, macroscopic traffic flow models with an average speed-based emission model has been set. The present chapter develops a research methodology to studying and comparing uncertainty in emission predictions using first-order CTM and using second-order METANET models. The developed research methodology is divided into three main steps, as illustrated in Figure 3.1 below. The first step sets out the mathematical preliminaries of the modelling chain, as described in § 3.2. It builds on the literature review provided for first-order LWR-type models and second-order Payne-type models of § 2.2.1.3, and average speed-based emission models of § 2.3.2. The outcome of the first step is a clear description of the final 'traffic flow - emission' modelling chain.

The second step, described in § 3.3, then follows by specifying the sources of uncertainty in emission predictions based on the final modelling chain. It builds on the literature review of Chapter 2. In this step, focussing on propagating uncertainty in the model outcome of macroscopic traffic flow models is justified, while acknowledging the existence of other sources of uncertainty. The third and final step, described in

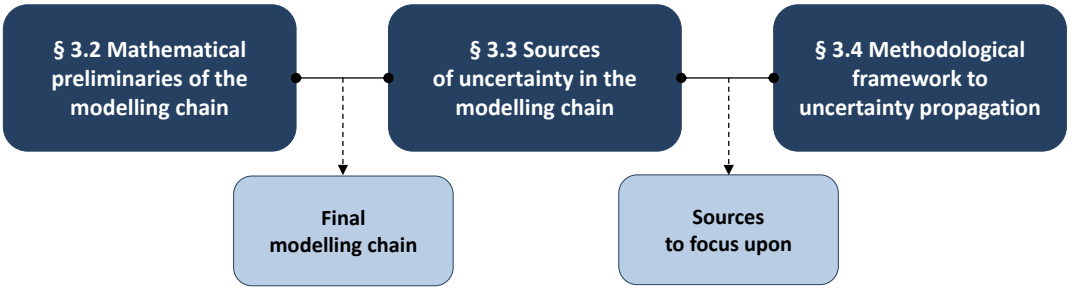


FIGURE 3.1: The three main steps of the proposed research methodology.

§ 3.4, develops the data-driven methodological framework to uncertainty propagation. A summary of the chapter is provided in § 3.5.

3.2 Mathematical preliminaries

Generally, given certain input data and model parameters, the main outputs of a discretised macroscopic road traffic flow model is *predicted* traffic flow (q) or traffic flow per lane ($\underline{q}$), traffic speed (u), and traffic density (ρ) or traffic density per lane ($\underline{\rho}$). Each of these predicted outputs is a 2-D matrix capturing the space-time dynamics of these variables at every pre-defined discretised cell of a particular road network and every time step over a particular period of time. Cells and time steps are defined as short segments (not necessarily of equal lengths) of a road network link and short time periods (of equal lengths) within the entire time period. Their lengths define the level of discretisation of the model. Let a road network be denoted by r and a modelling time period be denoted by m, this can be summarised by either Equations of 3.1 where Z comprises the five output matrices of traffic variables; f is the model function which defines the model structure; I or $\tilde{I}$ is the model input data which is either predicted (by for e.g. traffic assignment models) or measured, respectively; and β is the model parameter vector. Typically, β is not known and needs to be calibrated (or optimised) using certain measured traffic variables, denoted here as $\tilde{Z}$. Notice that from now on, any predicted traffic variable X *per lane* is denoted as $\underline{X}$; and any *measured* traffic variable is denoted as $\tilde{X}$. This study focusses on using measured traffic data for the calibration of β and for the preparation of input data. Hence, it rather builds on Equation 3.1b.

$$Z_{r,m} = (q, \underline{q}, u, \rho, \underline{\rho})_{r,m} = f(I_{r,m}, \beta_{r,m}) \tag{3.1a}$$

$$Z_{r,m} = (q, \underline{q}, u, \rho, \underline{\rho})_{r,m} = f\left[\tilde{I}_{r,m}, \beta(\tilde{Z}_{r,m})\right] \tag{3.1b}$$

There are three components to be defined clearly in any such models: structure, input, and parameters, as these are directly related to the sources of uncertainty. The following sections define these for each of the CTM (§ 3.2.1) and METANET (§ 3.2.2). It is worth noting here that since the implementation of models has been integrated with a particular data source, scripts of each model have been written here in R (R Core Team, 2015); a flow-diagram and a description of the developed functions for data extraction and pre-processing, and traffic flow modelling are provided in Appendix A. In § 3.2.3, the average speed-based emission model and its integration with the CTM and METANET model outputs are then defined simultaneously. Since the outputs of both CTM and METANET are the same in terms of their spatiotemporal resolution, there is not any difference in their integration approach.

3.2.1 First-order CTM

First-order LWR-type models were briefly described mathematically in their continuous form. Here, a detailed description of the discretised first-order CTM is provided. The original CTM model has been first developed by Daganzo (1994, 1995a) to describe the space-time dynamics of occupancy (in terms of number of vehicles) rather than traffic density. The original formulation also required the cell configuration of a road network to have equal lengths. Muñoz et al. (2004) then proposed a modified CTM which describes the dynamics of traffic density rather than occupancy in order to allow for flexibility in the cell configuration of a road network, i.e. to accept uneven cell lengths. This study uses the modified CTM (sometimes referred to as MCTM but this study keeps referring to it as CTM since there is no substantial difference in the structure).

Under both the CTM and METANET models, a road network is partitioned into short cells with number of lanes λ_i $\{i = 1, 2, ..., N\}$ where $i - 1 = 0$ represents the mainline origin and $i + 1 = N + 1$ represents the mainline destination of the road network. Figure 3.2 below provides an example of a road network partitioned to N cells, with the origin and destination mainline cells as cells 0 and $N + 1$, respectively. Cell interfaces are denoted by $i \pm 1/2$. Using only two main recursive equations, CTM allows for the calculation of traffic density at each cell i of length l_i $\{i = 1, 2, ..., N\}$, in kilometres [km], and time step t $\{t = 1, 2, ..., T\}$ of length T_s, in hours [h], where T is the total number of time steps. These are the equation of conservation law of vehicles, and the equation of fundamental diagram (i.e. traffic flow-density relationship).

CTM can be described as a combination of a link model and node model: a node is located wherever there is a change in geometry of a road network, i.e. at on-ramps

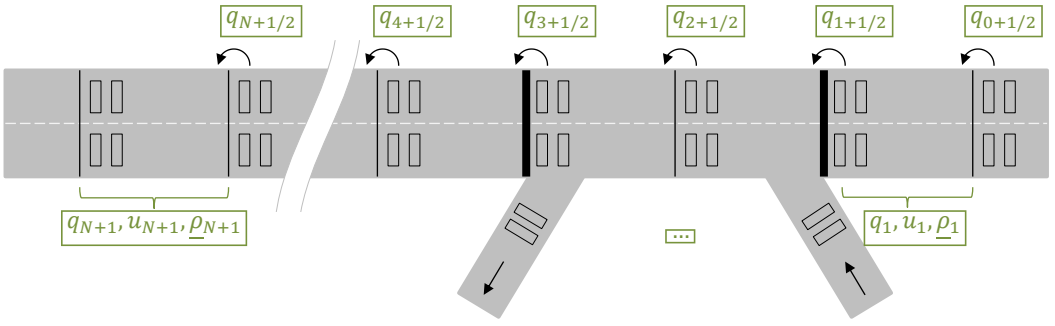


FIGURE 3.2: Example schematic of a two-lanes road network partitioned into N cells with cell 0 ($i - 1 = 0$) representing the mainline origin and cell N+1 ($i + 1 = N + 1$) representing the mainline destination. Vertical lines indicate cell interfaces ($i \pm 1/2$). Thick vertical lines represent the presence of a node. The road network has an on-ramp (i.e. merge node) at the start of cell 2 (at cell interface $1 + 1/2$) and an off-ramp (i.e. diverge node) at the end of cell 3 (at cell interface $2+1/2$). Rectangles show typical locations of dual loop detector sites providing measurements of traffic variables ($\tilde{Z}$). One can distinguish the number of a cell based on the link number (in this example, 3 links). Here, cell numbers are distinguished for the entire road network, for simplification.

requiring merge nodes and off-ramps requiring diverge nodes; and a link connects such nodes. The CTM model structure for cells (comprising a link) without any merge or diverge nodes is described by Equation 3.2.

$$\underline{\rho}_i(t+1) = \underline{\rho}_i(t) + \frac{T_s}{l_i \lambda_i}\Big[q_{i-1/2}(t) - q_{i+1/2}(t)\Big]; \underline{\rho}_i(0) = \tilde{\underline{\rho}}_i(0) \tag{3.2a}$$

$$q_{i+1/2}(t) = \min\Big\{S_i(t), R_{i+1}(t)\Big\} \tag{3.2b}$$

$$S_i(t) = \min\Big\{u_f \underline{\rho}_i(t), \underline{Q}_{max}\Big\}\lambda_i; S_0(t) = \tilde{d}_0(t) \tag{3.2c}$$

$$R_{i+1}(t) = \min\Big\{w\Big[\underline{\rho}_{max} - \underline{\rho}_{i+1}(t)\Big], \underline{Q}_{max}\Big\}\lambda_{i+1}; \underline{\rho}_{N+1}(t) = \tilde{\underline{\rho}}_{N+1}(t) \tag{3.2d}$$

such that $(u_f, \underline{Q}_{max}, w, \underline{\rho}_{max}) > 0$;

$$0 \leq \underline{\rho} \leq \underline{\rho}_{max};$$

$$w \leq; u_f;$$

$$\underline{Q}_{max} \leq \frac{\underline{\rho}_{max}}{1/u_f + 1/w};$$

$$u_f \leq \frac{\min(l_i)}{T_s}.$$

The CTM parameter vector is limited to those defining the trapezium fundamental diagram, as shown in Figure 3.3A below. While a triangular fundamental diagram was first proposed by Daganzo (1994), this was then relaxed to a trapezium fundamental diagram in Daganzo (1995a) which is the one considered in this study. The CTM parameter vector is denoted by $\beta = (u_f, \underline{Q}_{max}, w, \underline{\rho}_{max})$ and is comprised of: free-flow traffic speed (u_f) in kilometres per hour [km h^{-1}]; maximum traffic flow per lane ($\underline{Q}_{max}$) in vehicles per hour [veh h^{-1}]; backward wave speed (w) in kilometres per hour [km h^{-1}]; and maximum traffic density per lane ($\underline{\rho}_{max}$) in vehicles per kilometre [veh km^{-1}]. The parameter vector can either be cell-specific or uniform across all cells. While some studies have assumed cell-specific parameter vector (especially CTM-based studies with reference to Table 2.1), it is assumed here to be uniform across all cells for two main reasons (Spiliopoulou et al., 2014):

- for a simple road network, cells are assumed to share the same (static) infrastructural characteristics; and
- while flexibility, in terms of cell-specific parameters, can reflect real differences between cells, it can also lead to fake results from the calibration of high numbers of parameters based on limited data.

To initialise the recursive Equations of 3.2, model input data in the form of initial conditions and boundary conditions is required and is assumed here to be based on traffic measurements. For a simple road network (i.e. without any on-ramps and off-ramps), this is: traffic density for each cell at time step 0, $\tilde{\underline{\rho}}_i(0)$; traffic demand from mainline origin cell 0, $\tilde{d}_0(t)$; and traffic density at mainline destination at each time

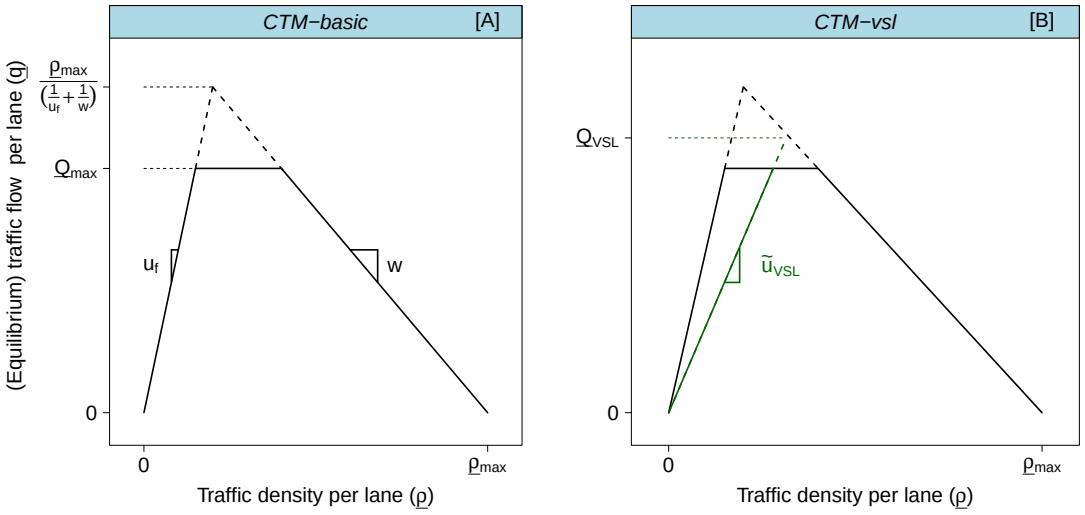


FIGURE 3.3: Traffic flow per lane [veh h^{-1}] as a function of traffic density per lane [veh km^{-1}] considered as [A] the CTM-basic fundamental diagram and [B] the CTM-vsl fundamental diagram[1].

step t, $\tilde{\underline{\rho}}_{N+1}(t)$. Given these, Equation 3.2a shows that traffic density $\underline{\rho}_i(t)$ can be calculated according to the conservation law of vehicles: in time interval $[t, t+1]$, the traffic density per lane depends on the inflow into cell i from cell $i-1$ (or equivalently, the traffic flow at cell interface $i-1/2$), and the outflow from cell i to cell $i+1$ (or equivalently, the traffic flow at cell interface $i+1/2$). These are denoted as $q_{i-1/2}(t)$, and $q_{i+1/2}(t)$, respectively. Traffic inflow and outflow (or traffic flow at cell interfaces) are calculated based on the fundamental diagram using Equation 3.2b: for cell interface $i+1/2$, it is the sending capacity of the upstream cell i under free-flow traffic conditions (Equation 3.2c) limited by the receiving capacity of the downstream cell $i+1$ under congested conditions (Equation 3.2d). In the case of $q_{0+1/2}(t)$ (i.e. traffic outflow from mainline origin), it is the input traffic demand at cell 0 limited by the receiving capacity of cell $i=1$ of the road network. The advantage behind assuming traffic flow measurements at origin cell 0 to be traffic demand (rather than actual traffic outflow) is ensuring that traffic inflows into the first cell $i=1$ does not exceed neither the maximum traffic flow $\underline{Q}_{max}$ which is a model parameter (rather than pre-specified to be the maximum measured traffic flow) nor the receiving capacity of cell $i=1$. Another advantage is ensuring flexibility in traffic conditions implied by the traffic demand i.e. it would not be necessary to impose traffic demand with free-flow condition at all times. Such advantages, however, can cause the boundary condition at the origin to not necessarily match the actual traffic inflows into cell $i=1$ as a result of uncertainty in the input traffic demand (i.e. measurements), the parameter vector used, and/or CTM model structure. In the case of $q_{N+1/2}(t)$ (i.e. traffic outflow from cell N to cell $N+1$), it is the sending capacity of the upstream cell N under free-flow conditions

[1]All figures in this research were plotted using R (R Core Team, 2015) and the Lattice package (Sarkar, 2008), unless otherwise stated.

limited by the receiving capacity of $N+1$ under congested conditions. The latter is determined by the estimated (from traffic flow and speed measurements) boundary condition $\tilde{\underline{\rho}}_{N+1}(t)$. If traffic density at the destination cell $N+1$ is imposed, it is also not necessary to ensure that it is in free-flow condition at all times. However, the main disadvantage behind this approach is the accumulation of counting errors of measurements as suggested by Papageorgiou et al. (1989).

The five conditions of Equation 3.2 represent four properties of the trapezium fundamental diagram plus the *C*ourant-*F*riedrichs-*L*ewy (CFL) condition that the minimum cell length needs to be longer than the free-flow distance to avoid instances when vehicles cross multiple cells in a single time step leading to model instabilities (Daganzo, 1995a). Since, in this study, the minimum of cell lengths is pre-defined and u_f is the unknown parameter, the condition is set on u_f rather than the minimum cell length.

The above formulation describes the dynamics of traffic density per lane for a simple road network, i.e. without merge or diverge nodes. However, an on-ramp (i.e. merge node) can be located at the upstream start of a cell and an off-ramp (i.e. diverge node) can be located at the downstream end of a cell, as shown earlier in Figure 3.2. At merge or diverge nodes, CTM formulation needs to be modified.

In the presence of an on-ramp at the start of cell i (i.e. at upstream cell interface $i-1/2$) with traffic demand $\tilde{d}_{on,i-1/2}(t)$, there exist two scenarios: the first is when cell i can accommodate the entire mainline cell $i-1$ outflow and on-ramp demand; and the second is when the mainline cell $i-1$ outflow and on-ramp demand exceed the receiving capacity of cell i, in which case the priority is assumed to be given for the on-ramp traffic demand and the remaining receiving capacity for the mainline outflow. This assumption can however be changed to reflect the percentage priority of the different entering links. In either scenarios, it is assumed that the on-ramp has unlimited capacity or at least that on-ramp traffic demand is lower than its capacity. The mainline outflow $q_{i-1/2}(t)$ and the on-ramp outflow $q_{on,i-1/2}(t)$ are modified based on Equation 3.3, accordingly.

$$q_{i-1/2}(t) = \begin{cases} S_{i-1}(t), & \text{if } S_{i-1}(t) \leq R_i(t) - \tilde{d}_{on,i-1/2}(t) \\ \max\left\{0, R_i(t) - \tilde{d}_{on,i-1/2}(t)\right\}, & \text{otherwise} \end{cases} \tag{3.3a}$$

$$q_{on,i-1/2}(t) = \begin{cases} \tilde{d}_{on,i-1/2}(t), & \text{if } S_{i-1}(t) \leq R_i(t) - \tilde{d}_{on,i-1/2}(t) \\ R_i(t) - q_{i-1/2}(t), & \text{otherwise} \end{cases} \tag{3.3b}$$

In the presence of an off-ramp at the end of cell i (i.e. at downstream cell interface $i+1/2$), it is assumed that the off-ramp has unlimited capacity (Muñoz et al., 2004). The traffic outflow is thus modified to Equation 3.4a where $\tilde{\theta}_{i+1/2}(t)$ represents the split ratio of the off-ramp. Also, the traffic outflow to the off-ramp is calculated based

on Equation 3.4b.

$$q_{i+1/2}(t) = \min\left\{\left[1 - \tilde{\theta}_{i+1/2}(t)\right] S_i(t), R_{i+1}(t)\right\} \tag{3.4a}$$

$$q_{off,i+1/2}(t) = \frac{\tilde{\theta}_{i+1/2}(t) q_{i+1/2}(t)}{1 - \tilde{\theta}_{i+1/2}(t)} \tag{3.4b}$$

In the presence of both an on-ramp and off-ramp on cell i, Equation 3.2a needs to be updated to take into account on-ramp inflows and off-ramp outflows at each time step t, as shown in Equation 3.5.

$$\underline{\rho}_i(t+1) = \underline{\rho}_i(t) + \frac{T_s}{l_i \lambda_i}\left[q_{i-1/2}(t) + q_{on,i-1/2}(t) - q_{i+1/2}(t) - q_{off,i+1/2}(t)\right] \tag{3.5}$$

By calculating the traffic density per lane at each cell i and time step t, traffic speed can be defined for each cell interface using traffic inflows and outflows (i.e. using traffic flows at cell interfaces). However, such definition does not guarantee equilibrium traffic speed i.e. the main assumption of first-order models, such as CTM. Since the CTM model parameter vector is to be optimised using traffic measurements, equilibrium traffic flow and speed are considered the best approximation of traffic measurements by the CTM. This study is thus focussed on traffic speed at each cell rather than at cell interfaces. It is worth noting here that some of the studies reviewed in § 2.2.2 and § 2.4.1 did use traffic speeds at cell interfaces for the CTM calibration (Spiliopoulou et al., 2014) or for emission predictions (J. Lin & Ge, 2006; Aziz & Ukkusuri, 2012; Rehimi & Landolsi, 2013). Equilibrium traffic flow and space-mean traffic speed at each cell i and time step t are defined based on Equation 3.6, respectively.

$$q_i(t) = \min\left\{u_f \underline{\rho}_i(t), \underline{Q}_{max}, w\left[\underline{\rho}_{max} - \underline{\rho}_i(t)\right]\right\}\lambda_i \tag{3.6a}$$

$$u_i(t) = \frac{q_i(t)}{\underline{\rho}_i(t)\lambda_i} \tag{3.6b}$$

Liu et al. (2014) have developed a multi-class LWR-type model extension. In their extension, they have proposed using a queue model at the mainline and on-ramp origins in order to ensure all the specified traffic demand at the origin(s) is entering the network with time while not exceeding the receiving capacity constraints of the cells. The same simple queue model is used here for similar purposes. The queue model for an on-ramp origin located at cell interface $i - 1/2$ is given by Equation 3.7 where $h_{on,i-1/2}(t)$ represents the queue length at time step t in vehicles. It is modified and added to the traffic demand if it is higher than 0 at each time step of the modelling time period. The updated traffic demand $\acute{\tilde{d}}_{on,i-1/2}(t)$ needs to be used in Equation 3.3 instead of $\tilde{d}_{on,i-1/2}(t)$. Equation 3.7 is similarly applied for the mainline origin with traffic demand, $\tilde{d}_0(t)$, traffic outflow, $q_{1-1/2}(t)$, and queue length, $h_{1-1/2}(t)$, instead.

The updated traffic demand at the mainline origin needs to be used for the calculation of $S_0(t)$ shown in Equation 3.2c.

$$h_{on,i-1/2}(t+1) = h_{on,i-1/2}(t) + T_s\left[\acute{d}_{on,i-1/2}(t) - q_{on,i-1/2}(t)\right] \tag{3.7a}$$

$$\acute{d}_{on,i-1/2}(t) = \tilde{d}_{on,i-1/2}(t) + \frac{h_{on,i-1/2}(t)}{T_s}; h_{on,i-1/2}(0) = 0 \tag{3.7b}$$

Few studies have extended the basic CTM model structure described above to take into account the impact of VSL operation. For instance, Chiou et al. (2012) and Z. Li et al. (2015) have proposed different methods to determining the optimal VSL control strategy through CTM which can improve safety or reduce collision risks on road networks. Hadiuzzaman and Qiu (2013) have also modified the CTM in order to analyse when VSL control can be effective. The main modification is to assume that VSL operation influences the demand part of the fundamental diagram, i.e. the free-flow traffic speed (u_f) and the maximum traffic flow ($\underline{Q}_{max}$), as shown in Figure 3.3B above. In accordance to that, the sending and receiving capacities of cells i and $i+1$ are calculated using Equation 3.8 instead. Such modification assumes 100% compliance of road traffic to imposed speed limits; ensuring 100% compliance typically requires strict enforcement of VSL (Chiou et al., 2012). The expected consequence of such modification is a reduction of traffic outflows to downstream cells, for the congestion to dissipate with time. The only additional input to CTM under VSL operation is a 2-D matrix of speed limits for each cell i and time step t, denoted as $\tilde{u}_{VSL_i}(t)$. In the absence of a VSL, the free-flow traffic speed parameter is assumed. The maximum traffic flow per lane $\underline{Q}_{VSL_i}(t)$, however, can be calculated from the fundamental diagram. Since the trapezium fundamental diagram is used here rather than the triangular one which is the case for the above-mentioned studies, the minimum of the original $\underline{Q}_{max}$ and the new $\underline{Q}_{VSL_i}(t)$ is used to limit the sending or receiving capacities.

$$S_i(t) = \min\left\{\min\left[u_f, \tilde{u}_{VSL_i}(t)\right]\underline{\rho}_i(t), \min\left[\underline{Q}_{max}, \underline{Q}_{VSL_i}(t)\right]\right\}\lambda_i \tag{3.8a}$$

$$R_{i+1}(t) = \min\left\{w\left[\underline{\rho}_{max} - \underline{\rho}_{i+1}(t)\right], \min\left[\underline{Q}_{max}, \underline{Q}_{VSL_{i+1}}(t)\right]\right\}\lambda_{i+1} \tag{3.8b}$$

$$\text{such that } \underline{Q}_{VSL_i}(t) = \frac{\underline{\rho}_{max}}{1/\tilde{u}_{VSL_i}(t) + 1/w}.$$

Studies which have developed a VSL-modification of CTM have often focussed on devising optimal VSL-control strategies. They focussed less on calibrating and validating the modified CTM on a VSL operated network against real traffic and speed limit data. This study takes advantage of a VSL-operated motorway in order to do so. This is expanded upon in later sections.

This section have defined the structure, input, and parameter vector of CTM on a road network with and without the operation of VSL. From now on, these are referred

to as 'CTM-basic' and 'CTM-vsl', respectively. The main model output is the 2-D matrix of traffic density per lane, from which the 2-D matrices of (equilibrium) traffic flow and speed can be estimated. The predicted traffic density, $\rho_i(t)$, and flow per lane, $\underline{q}_i(t)$, matrices can easily be calculated using the number of lanes of each cell i. The general traffic flow modelling form described in Equation 3.1b is adjusted to Equation 3.9 for f = CTM-basic and Equation 3.10 for f = CTM-vsl.

$$\begin{aligned} Z_{r,m} &= (q, \underline{q}, u, \rho, \underline{\rho})_{r,m} = f\Big[\tilde{I}_{r,m}, \beta(\tilde{Z}_{r,m})\Big] \\ &= f\Big\{\Big[\underline{\tilde{\rho}}_i(0), \tilde{d}_0(t), \underline{\tilde{\rho}}_{N+1}(t), \tilde{d}_{on,i-1/2}(t), \tilde{\theta}_{i+1/2}(t)\Big]_{r,m}, \\ &\quad \Big[(u_f, \underline{Q}_{max}, w, \underline{\rho}_{max})(\tilde{Z}_{r,m})\Big]\Big\} \end{aligned} \tag{3.9}$$

$$\begin{aligned} Z_{r,m} &= (q, \underline{q}, u, \rho, \underline{\rho})_{r,m} = f\Big[\tilde{I}_{r,m}, \beta(\tilde{Z}_{r,m})\Big] \\ &= f\Big\{\Big[\underline{\tilde{\rho}}_i(0), \tilde{d}_0(t), \underline{\tilde{\rho}}_{N+1}(t), \tilde{d}_{on,i-1/2}(t), \tilde{\theta}_{i+1/2}(t), \tilde{u}_{VSL_i}(t)\Big]_{r,m}, \\ &\quad \Big[(u_f, \underline{Q}_{max}, w, \underline{\rho}_{max})(\tilde{Z}_{r,m})\Big]\Big\} \end{aligned} \tag{3.10}$$

3.2.2 Second-order METANET

Second-order Payne-type models have also been briefly described mathematically in their continuous form. Here, a detailed description of the discretised second-order METANET is provided. Under the METANET model, the road network can be described as that for the CTM, illustrated earlier in Figure 3.2. METANET allows the calculation of traffic density per lane and speed at each cell i and time step t using three main recursive equations, rather than two as is the case of the CTM. These are the equation of conservation law of vehicles, the equation of traffic speed, and the equation of fundamental diagram. The equation of traffic speed, of course, adds to the complexity of the modelling process in comparison to the CTM. Especially that while the conservation law of vehicles equation is based on physical principles, the traffic speed equation is rather heuristic (Hegyi, 2004). The METANET model structure for cells

without any merge or diverge nodes is described in Equation 3.11 and Equation 3.12.

$$\underline{\rho}_i(t+1) = \underline{\rho}_i(t) + \frac{T_s}{l_i \lambda_i}\Big[q_{i-1}(t) - q_i(t)\Big]; \underline{\rho}_i(0) = \tilde{\underline{\rho}}_i(0) \tag{3.11a}$$

$$\underline{\rho}_i(t+1) = \min\Big\{\underline{\rho}_i(t+1), \underline{\rho}_{max}\Big\} \tag{3.11b}$$

$$q_i(t) = \underline{\rho}_i(t) u_i(t) \lambda_i \tag{3.11c}$$

$$\begin{aligned} u_i(t+1) = u_i(t) + \frac{T_s}{\tau}\Big[U[\underline{\rho}_i(t)] - u_i(t)\Big] + \frac{T_s}{l_i} u_i(t)\Big[u_{i-1}(t) - u_i(t)\Big] \\ - \frac{T_s \eta}{\tau l_i} \frac{\underline{\rho}_{i+1}(t) - \underline{\rho}_i(t)}{\underline{\rho}_i(t) + \kappa}; u_0(t) = \tilde{u}_0(t); \underline{\rho}_{N+1}(t) = \tilde{\underline{\rho}}_{N+1}(t) \end{aligned} \tag{3.11d}$$

$$u_i(t+1) = \max\Big\{u_{min}, u_i(t+1)\Big\} \tag{3.11e}$$

such that $(\underline{\rho}_{max}, \tau, \eta, \kappa, u_{min}) > 0$.

To initialise the recursive Equations of 3.11, model input data in the form of initial conditions and boundary conditions is required and is also assumed here to be based on traffic measurements. In addition to those required as input to CTM, METANET requires traffic speed at mainline origin cell 0 at each time step t, denoted as $\tilde{u}_0(t)$. Thus, for a simple road network (i.e. without any on-ramps and off-ramps), the METANET model input data is modified to $\tilde{I} = \Big[\tilde{\underline{\rho}}_i(0), \tilde{d}_0(t), \tilde{u}_0(t), \tilde{\underline{\rho}}_{N+1}(t)\Big]$. Given these, Equation 3.11a shows that the traffic density per lane can be updated according to the conservation law of vehicles which involves the use of traffic inflows and outflows. As described in Equation 3.11c, traffic inflows and outflows depend on the dynamics of traffic speed, as opposed to the CTM. Hence, it is not necessary to distinguish between traffic flow at cell interfaces and those within cells. Traffic speed equation depends mainly on four terms, as shown in Equation 3.11d, and are described in order of appearance as follows:

- the average traffic speed at the previous time step;
- a relaxation term to take into account the tendency of traffic to relax (accelerate or decelerate) to an equilibrium traffic speed (i.e. desired traffic speed);
- a convection term to take into account changes (increase or decrease) in traffic speed due to traffic inflows and outflows; and
- an anticipation term to take into account changes (increase or decrease) in traffic speed due downstream traffic conditions.

The equilibrium traffic speed, denoted as $U[\underline{\rho}_i(t)]$, of the relaxation term is defined based on the original METANET fundamental diagram of Figure 3.4A below and takes the exponential form of Equation 3.12a. This is considered the original METANET traffic speed-density equation. Frejo et al. (2012) have suggested a simple modification, since the exponential form can be replaced (Messner & Papageorgiou, 1990), if necessary. Frejo et al. (2012) have fitted both (the original and their modified fundamental diagram) using traffic data measured on a US motorway and suggested that

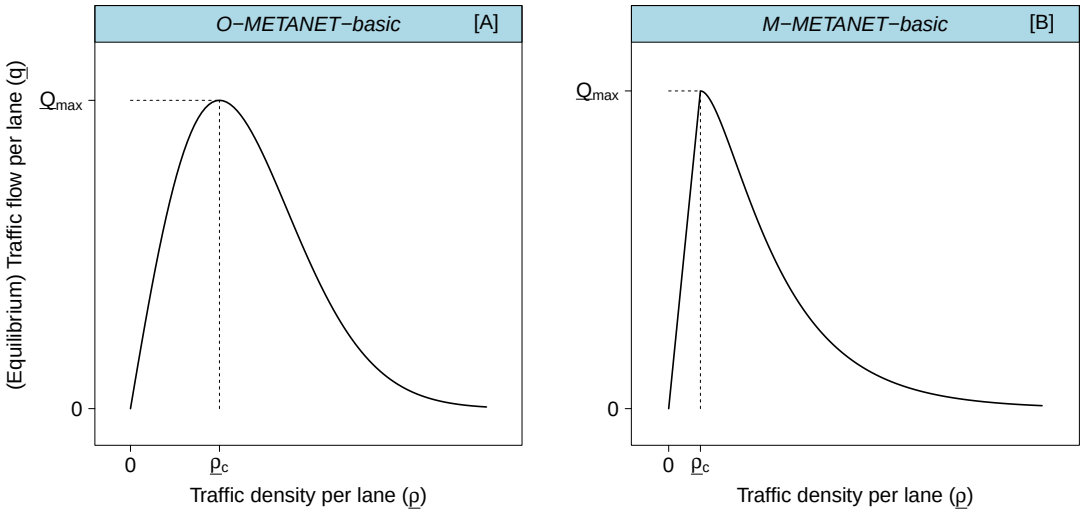


FIGURE 3.4: Traffic flow per lane [veh h^{-1}] as a function of traffic density per lane [veh km^{-1}] considered as [A] the O-METANET-basic fundamental diagram where $Q_{max} = \rho_c U[\rho_c] = \rho_c u_f e^{-1/\alpha}$; and [B] the M-METANET-basic fundamental diagram where $Q_{max} = \rho_c u_f$.

their modified METANET traffic speed-density equation can reduce the fit error without an increase in computational power; their conclusion, however, might as well be case (or data)-specific. Equation 3.12b shows their modified METANET equilibrium traffic speed-density relationship. Figure 3.4B above shows its corresponding fundamental diagram. The main modification is that under congested traffic conditions, the fundamental diagram takes the exponential form whereas under free-flow traffic conditions, the traffic flow is simply the product of traffic density and the free-flow traffic speed parameter.

$$U[\rho_i(t)] = u_f \exp\left[\frac{-1}{\alpha}\left(\frac{\rho_i(t)}{\rho_c}\right)^{\alpha}\right] \tag{3.12a}$$

$$U[\rho_i(t)] = \min\left\{u_f, \frac{u_f}{e^{-1/\alpha}} \exp\left[\frac{-1}{\alpha}\left(\frac{\rho_i(t)}{\rho_c}\right)^{\alpha}\right]\right\} \tag{3.12b}$$

such that $(u_f, \alpha, \rho_c) > 0$;

$$u_f \leq \frac{\min(l_i)}{T_s}.$$

From now on, METANET with the original fundamental diagram is referred to as O-METANET, and METANET with the modified one is referred to as M-METANET. In this study, the methodological framework is applied for the M-METANET. The reason behind that is not because it might better fit traffic data but rather because of:

- the dependence of the free-flow traffic speed parameter of the fundamental diagram on its congested parameters is reduced which implies desirable property to the parameter optimisation problem - this has been particularly highlighted by

Papageorgiou et al. (1990) who have mentioned that, for the original fundamental diagram, the deterioration of one parameter may be partially compensated by the modification of other parameters;

- below the critical density, the first derivative of the fundamental diagram is the free-flow traffic speed parameter, u_f (i.e. independent on traffic density) whereas that of the original fundamental diagram, it is: $u_f\left[1-\left(\underline{\rho}/\underline{\rho}_c\right)^{\alpha}\right]\exp\left[\frac{-1}{\alpha}\left(\frac{\underline{\rho}}{\underline{\rho}_c}\right)^{\alpha}\right]$ (i.e. dependent on traffic density). The definition and value of the free-flow traffic speed parameter is thus comparable to that of the trapezium fundamental diagram assumed in CTM.

In order to show an example of the above and a critical review of this assumption, a comparison of the parameter optimisation results of O-METANET and M-METANET is provided in § 5.2.

The equilibrium traffic speed-density equations (either original or modified) give rise to three model parameters. These are:

- the free-flow traffic speed parameters, u_f, in [km h^{-1}] which is subject to the CFL condition;
- the critical traffic density per lane parameter, $\underline{\rho}_c$, in [veh km^{-1}]; and
- additional unitless parameter α.

Papageorgiou et al. (1990) and Messner and Papageorgiou (1990) have also modified the original Payne formulation in order to account for both the influence of lane-drop(s) (or increase) and on-ramp(s) on average traffic speed by adding two terms presented in Equation 3.13, respectively. $\Delta\lambda_i$ in Equation 3.13a represents the difference between the number of lanes at cell i and cell $i+1$, i.e. $\lambda_i - \lambda_{i+1}$. If it is a lane drop (i.e. positive difference), this is assumed to cause a reduction in traffic speed. In the case of a lane increase (i.e. negative difference), this is assumed to cause an increase in traffic speed. If an on-ramp (i.e. a merge node) is located at the start of cell i, Equation 3.13b suggests a reduction in traffic speed as a result of the merging phenomena depending on the traffic flows from the on-ramp.

$$u_i(t+1) = u_i(t+1) - \frac{\phi T_s \Delta\lambda_i}{l_i\lambda_i}\frac{\underline{\rho}_i(t)u_i^2(t)}{\underline{\rho}_c} \tag{3.13a}$$

$$u_i(t+1) = u_i(t+1) - \frac{\delta T_s}{l_i\lambda_i}\frac{q_{on,i}(t)u_i(t)}{\underline{\rho}_i(t)+\kappa} \tag{3.13b}$$

such that $(\phi,\delta) > 0$.

Traffic flow at an on-ramp, $q_{on,i}(t)$, is typically assumed to depend on its traffic demand, on-ramp capacity, and the maximum traffic flow that can enter the mainline road network, as described in Equation 3.14a. For simplicity, on-ramp capacity is assumed the same as the mainline. To ensure the entire traffic demand enters the road network, a simple queue model is proposed, similar to that of the CTM (i.e.

Equation 3.7a). This is similarly applied for the mainline origin traffic outflow, $q_0(t)$, as shown in Equation 3.14b.

$$q_{on,i}(t) = min\left\{ \acute{\tilde{d}}_{on,i}(t), C_{on,i}, C_{on,i}\frac{\underline{\rho}_{max} - \underline{\rho}_i(t)}{\underline{\rho}_{max} - \underline{\rho}_c} \right\} \tag{3.14a}$$

$$q_0(t) = min\left\{ \acute{\tilde{d}}_0(t), C_0, C_0\frac{\underline{\rho}_{max} - \underline{\rho}_1(t)}{\underline{\rho}_{max} - \underline{\rho}_c} \right\} \tag{3.14b}$$

where $C_{on,i} = \underline{Q}_{max}\lambda_{on,i}$;
$C_0 = \underline{Q}_{max}\lambda_0$;
$\underline{Q}_{max} = \underline{\rho}_c U[\underline{\rho}_c]$.

Messner and Papageorgiou (1990) have also modified the original Payne formulation to take into account the influence of upstream traffic speed in the convection term when there is more than one upstream link; this is in the presence of an on-ramp where there is two upstream links (the mainline and the on-ramp itself) rather than one. Accordingly, $u_{i-1}(t)$ of the convection term of Equation 3.11d is modified to Equation 3.15 and is referred to as the 'virtual traffic speed'. Traffic speed at on-ramp(s) is assumed to be input data from traffic measurements.

$$u_{i-1}(t) = \frac{u_{i-1}(t)q_{i-1}(t) + u_{on,i}(t)q_{on,i}(t)}{q_{i-1}(t) + q_{on,i}(t)}; u_{on,i}(t) = \tilde{u}_{on,i}(t) \tag{3.15}$$

Similar to the CTM, in the presence of off-ramps, unlimited capacity is assumed. Equation 3.16a describes the influence of an off-ramp (i.e. a diverge node) at the end of cell i on the mainline traffic outflow to cell $i+1$ which is reduced at each time step t based on the estimated (from traffic flow measurements) split ratios $\tilde{\theta}_i(t)$. Messner and Papageorgiou (1990) have also modified the original Payne formulation to take into account the influence of downstream traffic density in the anticipation term when there is more than one downstream cell; this is in the presence of an off-ramp where there is two downstream cells (the mainline and the off-ramp itself) rather than one. Accordingly, $\underline{\rho}_{i+1}(t)$ of the anticipation term of Equation 3.11d is modified to Equation 3.16c and is referred to as 'virtual traffic density'. Traffic density per lane at off-ramp(s) is considered as input data from traffic measurements.

$$q_i(t) = \left[1 - \tilde{\theta}_i(t)\right]\underline{\rho}_i(t)u_i(t)\lambda_i \tag{3.16a}$$

$$q_{off,i}(t) = \tilde{\theta}_i(t)\underline{\rho}_i(t)u_i(t)\lambda_i \tag{3.16b}$$

$$\underline{\rho}_{i+1}(t) = \frac{\underline{\rho}^2_{i+1}(t) + \underline{\rho}^2_{off,i}(t)}{\underline{\rho}_{i+1}(t) + \underline{\rho}_{off,i}(t)}; \underline{\rho}_{off,i}(t) = \tilde{\underline{\rho}}_{off,i}(t) \tag{3.16c}$$

In the presence of both an on-ramp and off-ramp on cell i, Equation 3.11a needs to be updated to take into account on-ramp inflows and off-ramp outflows at each time

step t, as shown in Equation 3.17.

$$\underline{\rho}_i(t+1) = \underline{\rho}_i(t) + \frac{T_s}{l_i \lambda_i}\Big[q_{i-1}(t) + q_{on,i}(t) - q_i(t) - q_{off,i}(t)\Big] \tag{3.17}$$

With that, the METANET model parameter vector β consists of: three parameters related to the traffic speed-density equation (or the fundamental diagram) as described earlier; and seven parameters related to the traffic speed and density equations. All of which are positive parameters. These are described as follows:

- $[\tau, \eta, \kappa, \phi, \delta, u_{min}]$ are the traffic speed equation parameters with τ being a relaxation time parameter in [h] (in case it is in seconds, then τ in the model equations needs to be replaced by $\tau/3600$); η being an anticipation parameter in [km^2 h^{-1}]; κ being a parameter added to avoid infinite value of the anticipation and on-ramp merging terms when the traffic density per lane is very low, in [veh km^{-1}]; ϕ and δ being lane-drop and on-ramp merging unitless parameters; and u_{min} being a constraint parameter to avoid traffic speed less than the minimum value. In the absence of lane-drops and/or on-ramps on the entire road network, ϕ and/or δ are not included as parameters (and their corresponding terms are set to 0 for all cells and time steps) in order to avoid unnecessary parameterisation.
- $\underline{\rho}_{max}$ is used both as a constraint parameter for the traffic density equation to avoid values higher than the maximum value (Poole & Kotsialos, 2016) and as a constraint in the queue model equations.

Rather than considering u_{min} and $\underline{\rho}_{max}$ as model parameters, this study sets these to specific values mainly to avoid unnecessary uncertainty around their values since they are added to the METANET as constraints to traffic speed and density. Papageorgiou et al. (1990) have suggested constraining traffic speed by a small enough value, for e.g. $u_{min} = 1$, rather than having it as parameter to be optimised as has been later suggested in Poole and Kotsialos (2016). Poole and Kotsialos (2016) have suggested $\underline{\rho}_{max}$ to be optimised but within a minimum and maximum bound (required by certain optimisation algorithms). This study sets u_{min} to 1 [km h^{-1}], and $\underline{\rho}_{max}$ to 190 [veh km^{-1}] which is the maximum bound suggested in Poole and Kotsialos (2016).

There exist a debate on the impact of VSL on the METANET model formulation, particularly on its fundamental diagram equation. One of which is the use of VSL-specific fundamental diagram equations which assumes that speed limits: reduce traffic speed and thus traffic flow during under-critical traffic density; and increase traffic flow at over-critical traffic density (Alessandri et al., 1999; Papamichail et al., 2008; Papageorgiou et al., 2008; Carlson et al., 2010, for example). The latter implies that the speed limit can influence traffic states that have a lower traffic speed than the imposed speed limit. According to Hegyi et al. (2005b), this is rather unrealistic and exaggerates the influence of VSL. The alternative proposition is that the speed limits only influence the desired traffic speed which are higher than the imposed ones. This can

be expressed as the minimum of: equilibrium traffic speed based on experienced traffic density; and traffic speed imposed by the speed limit. Nevertheless, Hegyi et al. (2005b) have emphasised the importance of a more quantitative validation of the proposed VSL model. Studies often focus on the influence of VSL on the fundamental diagram outside the scope of a dynamic traffic flow model such as CTM or METANET. In this study, the second approach is used which is similar to the assumption used for the CTM and is also less parameterised in comparison to the first approach. It is important to note, however, that here none of the alternatives are favoured. However, this study focusses on the influence of the chosen approach on the model performance in comparison to using the non-VSL model structure. This is made in an attempt to answer whether the VSL-related assumptions can actually improve the model performance or not on a VSL-operated road network. Such an assumption leads to Equation 3.18 which is a modification of Equation 3.12. Their corresponding fundamental diagrams are shown in Figure 3.5.

$$U[\underline{\rho}_i(t)] = min\left\{\tilde{u}_{VSL_i}(t), u_f \exp\left[\frac{-1}{\alpha}\left(\frac{\underline{\rho}_i(t)}{\underline{\rho}_c}\right)^{\alpha}\right]\right\} \tag{3.18a}$$

$$U[\underline{\rho}_i(t)] = \min\left\{\tilde{u}_{VSL_i}(t), u_f, \frac{u_f}{e^{-1/\alpha}} \exp\left[\frac{-1}{\alpha}\left(\frac{\underline{\rho}_i(t)}{\underline{\rho}_c}\right)^{\alpha}\right]\right\} \tag{3.18b}$$

The VSL assumption also influences the maximum traffic flow, $\underline{Q}_{max}$, required for the calculation of traffic outflows from the mainline origin. As depicted in Figure 3.5, in the presence of speed limit at cell i and time step t, $\underline{Q}_{VSL_i}(t)$ can be calculated for

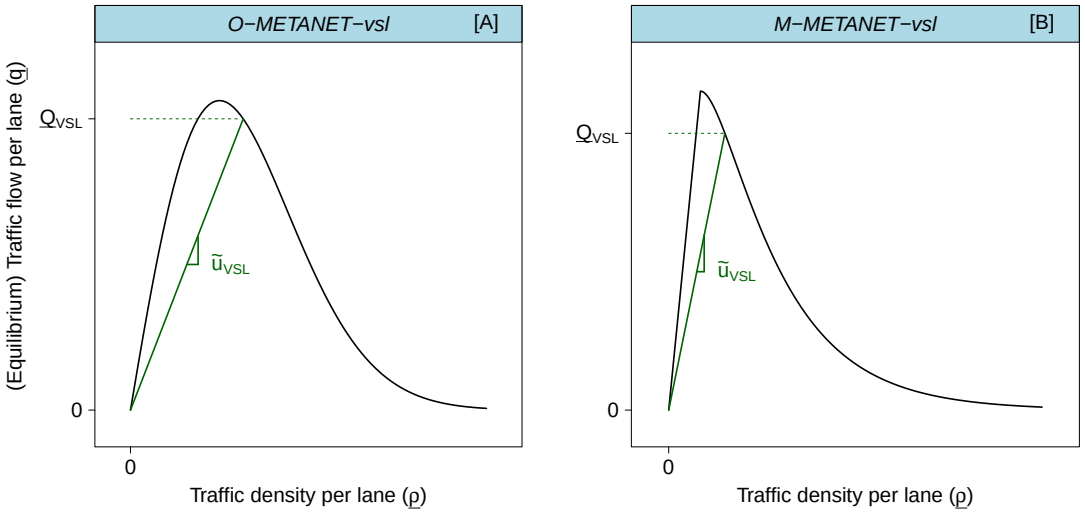


FIGURE 3.5: Traffic flow per lane [veh h^{-1}] as a function of traffic density per lane [veh km^{-1}] considered as [A] the O-METANET-vsl fundamental diagram; and [B] the M-METANET-vsl fundamental diagram.

the O-METANET and M-METANET as follows, respectively:

$$\underline{Q}_{VSL_i}(t) = \underline{\rho}_c \tilde{u}_{VSL_i}(t) \left\{ -\alpha \log \left[\frac{\tilde{u}_{VSL_i}(t)}{u_f} \right] \right\}^{1/\alpha} \tag{3.19a}$$

$$\underline{Q}_{VSL_i}(t) = \underline{\rho}_c \tilde{u}_{VSL_i}(t) \left\{ -\alpha \log \left[\frac{\tilde{u}_{VSL_i}(t) e^{-1/\alpha}}{u_f} \right] \right\}^{1/\alpha} \tag{3.19b}$$

This section has defined the structure, input, and parameter vector of METANET on a road network: using the original or modified traffic speed-density equation; and with and without the operation of VSL. From now on, these are referred to as: 'O-METANET-basic'; 'M-METANET-basic'; 'O-METANET-vsl'; and 'M-METANET-vsl'. The main model outputs are the 2-D matrices of traffic density per lane, speed, and flow per lane. The predicted traffic density, $\rho_i(t)$, and traffic flow per lane, $\underline{q}_i(t)$, matrices can then be calculated using the number of lanes of each cell i. The general traffic flow modelling form described in Equation 3.1b is modified to Equation 3.20 for f = O-METANET-basic or M-METANET-basic, and Equation 3.21 for f = O-METANET-vsl or M-METANET-vsl.

$$\begin{aligned} Z_{r,m} &= (q, \underline{q}, u, \rho, \underline{\rho})_{r,m} = f\left[\tilde{I}_{r,m}, \beta(\tilde{Z}_{r,m})\right] \\ &= f\Bigg\{ \left[\underline{\tilde{\rho}}_i(0), \tilde{d}_0(t), \underline{\tilde{\rho}}_{N+1}(t), \tilde{u}_0(t), \tilde{d}_{on,i}(t), \tilde{u}_{on,i}(t), \tilde{\theta}_i(t), \underline{\tilde{\rho}}_{off,i}(t)\right]_{r,m}, \\ &\quad \left[(u_f, \alpha, \underline{\rho}_c, \tau, \eta, \kappa, \phi, \delta)(\tilde{Z}_{r,m})\right] \Bigg\} \end{aligned} \tag{3.20}$$

$$\begin{aligned} Z_{r,m} &= (q, \underline{q}, u, \rho, \underline{\rho})_{r,m} = f\left[\tilde{I}_{r,m}, \beta(\tilde{Z}_{r,m})\right] \\ &= f\Bigg\{ \left[\underline{\tilde{\rho}}_i(0), \tilde{d}_0(t), \underline{\tilde{\rho}}_{N+1}(t), \tilde{u}_0(t), \tilde{d}_{on,i}(t), \tilde{u}_{on,i}(t), \tilde{\theta}_i(t), \underline{\tilde{\rho}}_{off,i}(t), \tilde{u}_{VSL_i}(t)\right]_{r,m}, \\ &\quad \left[(u_f, \alpha, \underline{\rho}_c, \tau, \eta, \kappa, \phi, \delta)(\tilde{Z}_{r,m})\right] \Bigg\} \end{aligned} \tag{3.21}$$

3.2.3 Average speed-based emission predictions

The concept behind average speed-based emission models has been introduced briefly in § 2.3.2. Here, such models are described in details in the context of their integration with both CTM and METANET. The potential pitfalls of integrating high-resolution cell-based traffic flow model outputs with average speed-based emission models which give rise to additional uncertainty components are discussed in more details in the second step of § 3.3. Those related to the emission model are quite general and are not confined to a specific choice of an average speed-based emission model (i.e. its emission factors). This section describes the integration approach followed by the studies reviewed in § 2.4. The most comprehensive is that of Csikós et al. (2015).

Given the traffic density, flow, and speed model outputs of CTM and METANET, average speed-based emissions of an air pollutant a at each cell i of length l_i (and number of lanes λ_i) and time step t of length T_s can be calculated, in principle, using either Equations of 3.22. The total emissions over all cells and/or all time steps can then be easily calculated using Equation 3.23. Equation 3.22a utilises as input: the proportion of each vehicle category, γ_c; and both traffic flow model outputs of traffic density, $\underline{\rho}_i(t)$, and speed, $u_i(t)$, to calculate emissions in grams produced in a particular cell during a particular time step, $E_i^a(t)$.

$$E_i^a(t) = \sum\nolimits_c \boxed{\underline{\rho}_i(t)\lambda_i l_i \gamma_c u_i(t) T_s}\, ef^{a,c}\left[u_i(t), \omega^{a,c}\right] \tag{3.22a}$$

$$E_i^a(t) = \sum\nolimits_c \boxed{\underline{q}_i(t)\lambda_i l_i \gamma_c T_s}\, ef^{a,c}\left[u_i(t), \omega^{a,c}\right] \tag{3.22b}$$

$$E_i^a = \sum\nolimits_t E_i^a(t) \tag{3.23a}$$

$$E^a(t) = \sum\nolimits_i E_i^a(t) \tag{3.23b}$$

$$E^a = \sum\nolimits_i \sum\nolimits_t E_i^a(t) \tag{3.23c}$$

Traffic density and speed of Equation 3.22a can be replaced by traffic flow to calculate the same. When emissions are calculated with the same spatiotemporal resolution of the model outputs, Equation 3.22a and Equation 3.22b must give the same result. When the model outputs are aggregated, results will change to a certain degree; since the multiplication of the separately aggregated traffic density and speed might not be equal to the aggregated traffic flow. In such case, there is the question of which one to use for predicting emissions. Since the developed methodological framework involves the aggregation of the model outputs to match the spatiotemporal resolution of measurements (in order to estimate errors), it is worth mentioning the two calculation methods. One way to determine if there is a difference between the two is to compare aggregated traffic density outputs with the aggregated flow divided by aggregated speed outputs. Because the case studies of this research does not show any substantial differences between the two, only traffic density, rather than flow, will be used from now on i.e. Equation 3.22a. In § 5.3, an example comparison is provided to illustrate the similarity.

The first part of Equation 3.22 (in red box) allows the calculation of total distance (e.g. kilometres) travelled by vehicles of certain category in a certain cell during a certain time step. Since neither CTM nor METANET are inherently multi-class traffic flow models, their outputs do not take into consideration the different driving conditions of each vehicle class. Given that only *P*assenger *C*ars (PCs) and *L*ight *D*uty *V*ehicles (LDVs) are subject to the *N*ational *S*peed *L*imits (NSL), traffic is assumed to be PCs and LDVs only, and their corresponding emission factors are used. In addition, any

traffic flow model does not distinguish between vehicles of different fuel types and emission standards since it is not expected that they influence traffic conditions. All of which (vehicle class, fuel type, and emission standard), make up a particular vehicle category. Typically the vehicle class, fuel type, and emission standard of any vehicle can be set either randomly or deterministically to reflect the percentage of each. The latter are normally estimated using external vehicle data bases. Here, traffic density (or the total vehicle distance travelled) at each cell and time step is divided in a deterministic manner depending on the proportion of different fuel types and emission standards of PCs and LDVs only. The proportion of other vehicle classes (mainly heavy duty vehicles on motorway networks) are assumed to be either PCs or LDVs. This type of data, referred to as the fleet mix data, are road network and time period-specific. So, such input is to be explored in the data preparation Chapter 4.

The parameterisation of average speed-based emission models is embedded in the second part: $ef^{a,c}\left[u_i(t), \omega^{a,c}\right]$ or the emission factors used. This is the predicted amount of a particular air pollutant a emitted by a vehicle of category c in grams per distance travelled. Thus, total vehicular distance travelled is needed to calculate the total emissions produced. Emission factors use average traffic speed $u_i(t)$ as input and $\omega^{a,c}$ as an air pollutant and vehicle category-specific parameter vector.

While average speed-based emission modelling structure is unique (i.e. as shown in Equation 3.22), the underlying emission factors' functions vary depending on the exact model used, as described in § 2.3. Because the road networks in this study are UK-based, COPERT emission factors (Ntziachristos et al., 2009) are used. COPERT is developed based on European vehicles and is widely used in European countries (22 out of 27 countries) (Ntziachristos et al., 2009), both in research and practice. COPERT is mainly developed using laboratory-based driving cycles such as the Common Artemis Driving Cycles (CADC). An example of COPERT(4v10) emission factors of NO_x air pollutant ($a = NO_x$) for four different vehicle categories is provided in Figure 3.6. The emission factors are functions of average speed; here this is assumed to be cell-based average traffic speed. Minimum emission factors lie within an average speed between 60 and 80 [km h^{-1}] for both Euro 4 and Euro 5 vehicles. The non-linearity of the functions poses an interesting question regarding uncertainty propagation, particularly the fact that the example functions have similar values at low and high average speeds. This will be expanded upon in the results. Their corresponding emission factors' functions are provided as follows: Euro 4 and Euro 5 Diesel PCs; and Euro 4 and Euro 5 Diesel LDVs (for illustration of what a function can represent). ω_i $\{i = 1, 2, ..., 9\}$ are parameters set out by COPERT.

$$ef(u, \omega) = (\omega_1 + \omega_2 u + \omega_3 u^2)$$
$$ef(u, \omega) = (\omega_1 + \omega_2 u + \omega_3 u^2)(1 - \omega_4)$$
$$ef(u, \omega) = (\omega_5 u^2 + \omega_6 u + \omega_7)(1 - \omega_8)$$
$$ef(u, \omega) = (\omega_5 u^2 + \omega_6 u + \omega_7)(1 - \omega_9)$$

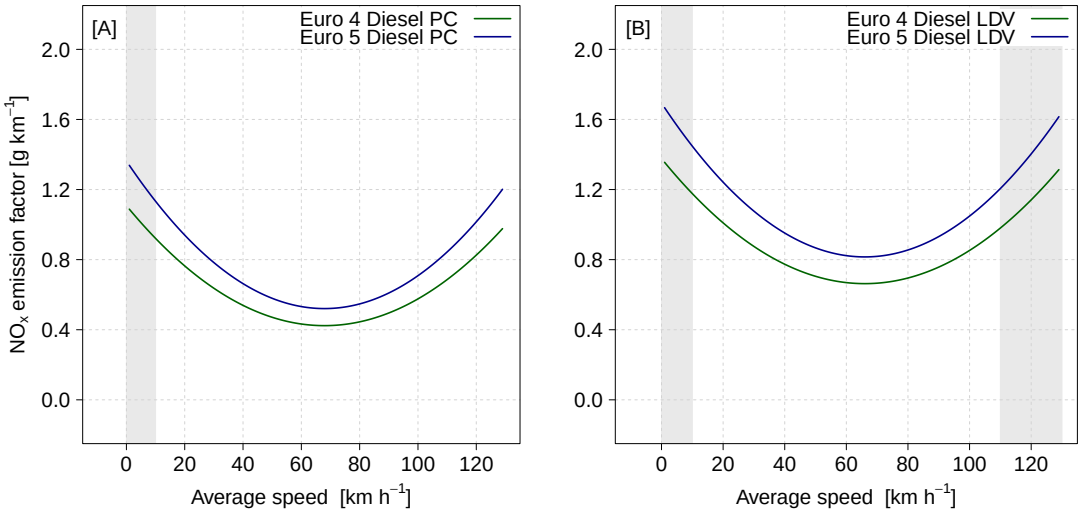


FIGURE 3.6: NO_x COPERT(4v10) emission factors [g km^{-1}] as a function of average speed [km h^{-1}] for [A] Euro 4 and Euro 5 Diesel PCs; and [B] Euro 4 and Euro 5 Diesel LDVs. The range of average speed at which the emission factors are developed do not include the grey shaded areas which is [0, 10] [km h^{-1}] for [A] and [0, 10] and [110, 130] [km h^{-1}] for [B].

The average speed-based emission model can in itself be seen as a nested model where its structure, input, and parameter vector depend on another model: emission factors. Let $E^a_{r,m}$ denote the 2-D matrix capturing the space-time dynamics of emission predictions for air pollutant a, road network r, and time period m, such that:

$$E^a_{r,m} = g^a\Big[(\gamma, \underline{\rho}, u)_{r,m}, ef^a(u_{r,m}, \omega^a)\Big] \tag{3.24}$$

where g refers to the general average speed-based emission model structure of Equation 3.22a. The first term represents the model input, and the second term represents the parameter vector embedded in the emission factors' functions of all vehicle categories. If the traffic inputs are replaced by the traffic flow model structure, the final 'traffic flow - emission' modelling chain can be summarised as follows:

$$E^a_{r,m} = g^a\Bigg\{\Big[\gamma_{r,m}, f[\tilde{I}_{r,m}, \beta(\tilde{Z}_{r,m})]\Big], ef^a\Big[f[\tilde{I}_{r,m}, \beta(\tilde{Z}_{r,m})], \omega^a\Big]\Bigg\} \tag{3.25}$$

To summarise, the emission model structure g takes two forms. The form which depends on traffic density will be used here. This shall be empirically justified in § 5.3. The traffic flow model structure f can take four forms: CTM-basic; CTM-vsl; M-METANET-basic; or M-METANET-vsl.

It is perhaps important to note here that COPERT is only one approach to predicting emissions based on either CTM or METANET. This approach has been selected due to the dominant use of average speed-based emission models in the literature reviewed in § 2.4. Its dominant use is a result of its direct integration with CTM or

METANET i.e. it does not require any major intermediary steps to predicting emissions. This facilitates, as a first step, the development of a methodology to uncertainty propagation. In addition to that, the wide availability of average traffic speed data facilitates the development of a methodology that uses real data to quantifying uncertainty. This is however at the expense of average speed-based model's inability to take into account variations in traffic speeds or individual driving behaviour that could otherwise be considered in more detailed emission models. The use of the latter may help reveal differences in uncertainty arising from the outputs of CTM in comparison to METANET that might not be captured by using COPERT or average speed-based emission models. By defining the final generic (average speed-based) modelling chain here, the next section describes in details the sources of uncertainty in the final 2-D emission prediction matrix, $E^{a}_{r,m}$ and expands on those arising from the use of average speed-based emission models, in general, and COPERT, in particular. Potential use of other emission model types and the implication of their use on the developed methodology is further highlighted in Chapter 6.

3.3 Sources of uncertainty

By examining the modelling chain described earlier, sources of uncertainty in emission predictions at each cell and time step (and ultimately in total emission predictions) of a road network r and time period m can be categorised into three, the levels of which depend on whether the modelling chain is to be applied for current scenarios or future scenarios.

1. Those related to the average speed-based emission model, in general, and COPERT, in specific.
2. Those related to the selected macroscopic traffic flow model.
3. Those related to the integration compatibility of selected macroscopic traffic flow models and COPERT.

Figure 3.7 shows the potential sources arising from the three components of the first two categories. These are the model structure, input, and parameter vectors of the average speed-based emission model and macroscopic traffic flow model. Those related to the emission model are described first below, noting that both the structure and parameter vector are dependent on the selected air pollutant.

The emission model structure relies on the assumption that macroscopic traffic behaviour in terms of cell-based traffic speed can predict emissions without consideration of disaggregate vehicle dynamics. The limitations of such assumption which gave rise to more detailed (instantaneous) emission models, are described in § 2.3.2. Uncertainty in such structure is generic to average speed-based emission models, and not only to COPERT.

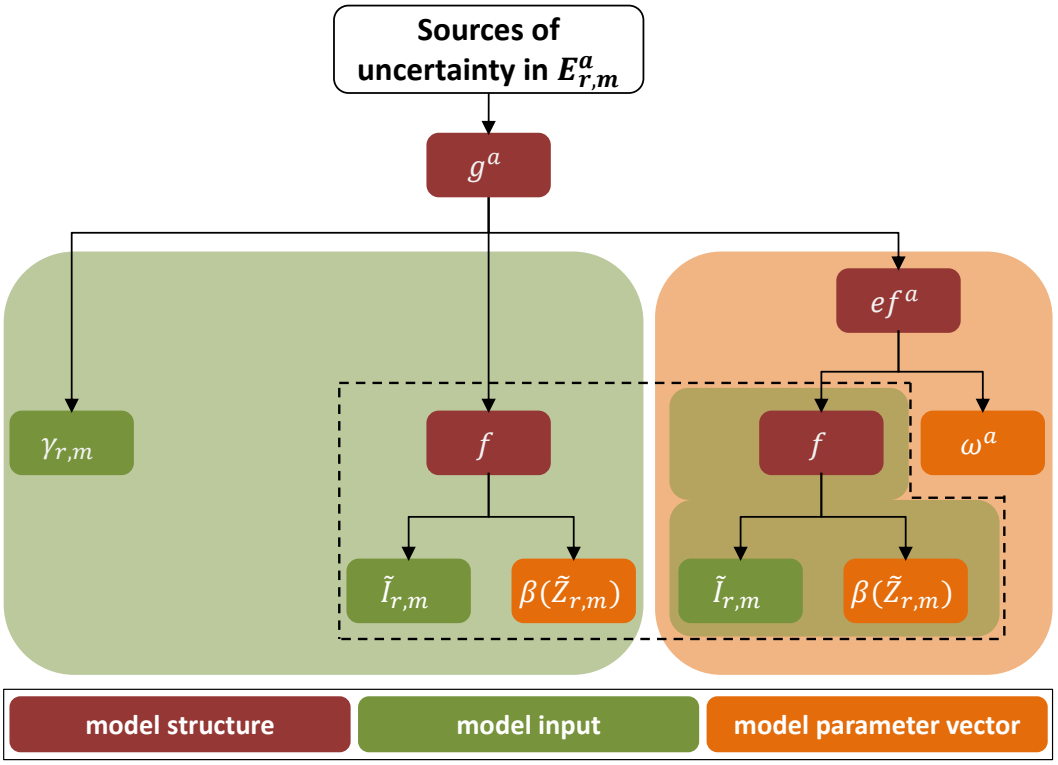


FIGURE 3.7: Flow chart describing the sources of uncertainty in average speed-based emission predictions $E^a_{r,m}$ of air pollutant a, road network r, and time period m, based on Equation 3.25.

The emission model input corresponds to both traffic data input and vehicle fleet mix. Uncertainty in traffic data input is directly related to those from the selected traffic flow models, described next separately. Vehicle fleet mix can be a source of uncertainty mainly for two reasons, both of which are generic to average speed-based emission models, and not only to COPERT:

- The first is the limited spatial and temporal resolution. For instance, fleet mix data used in the UK with categorisation level matching that of COPERT are aggregated spatially based on urban, rural, or motorway roads in and out of the City of London, and temporally based on each year. These numbers thus neglect motorway-to-motorway and link-to-link spatial differences as well as seasonal, daily, and within day temporal differences.
- The second is a result of assumptions being made on the proportions of vehicles other than PCs and LDVs.

The emission model parameter vector are nested within the emission factors' functions. They are thus highly dependent on the uncertainty of the functions themselves, which are highly parameterised vehicle category-specific non-linear functions. Uncertainty arising from these functions and their parameters are due to four main reasons, which at least apply for the case of COPERT.

1. They are built based on laboratory measurement data rather than real-world-based measurement data, the limitations of which are expanded upon in § 2.3.1.

This induces measurement uncertainty in terms of possible inaccuracies with respect to the real-world.

2. They are built based on limited measurement data. This also induces measurement uncertainty in terms of potential sampling error.
3. They ignore inherent variability of emission levels of different vehicles (even under the same category and driving cycle) and different trips (even for the same vehicle and driving cycle) (Ntziachristos et al., 2009). Such inherent variability exists even under more complex emission models.
4. The form of average speed-based functions and ultimately their parameters might not necessarily be the best fit to available data.

In addition to the above, sources related to the macroscopic traffic flow model depend on the selected model (i.e. CTM-basic, CTM-vsl, M-METANET-basic, or M-METANET-basic). As reflected in Figure 3.7, uncertainty in traffic density and speed predictions influences uncertainty in total distance travelled by each vehicle category on each cell and time step. Uncertainty in traffic speed predictions also influences uncertainty in emission factors, as they act as input. The three sources are described below.

The traffic flow model structure describes traffic dynamics at an aggregate level without taking into account vehicle dynamics, but also models traffic deterministically. Deterministic modelling ignores inherent variability in traffic dynamics due to external factors such as weather or infrastructure. In addition to that, given a particular model structure, uncertainty arises due to:

- limitations of CTM-basic and CTM-vsl in describing traffic dynamics such as multi-class, lane-changing phenomenon, heterogeneous driving behaviour, and lane-drops and merging phenomena (J. Li et al., 2012). It also arises from the assumptions of on-ramp merging, off-ramp capacity, and VSL impact on traffic dynamics, described earlier; or
- limitations of M-METANET-basic and M-METANET-vsl in the equation of traffic speed dynamics which is not entirely based on physical principles. Both are also limited in describing multi-class dynamics and lane-changing phenomenon.

The macroscopic traffic flow model input is dependent on the selected model i.e. if it is CTM-basic, CTM-vsl, M-METANET-basic, or M-METANET-basic, with the M-METANET requiring more input than CTM. Model input uncertainty is a result of potential inaccuracies and imprecisions in traffic measurement data. J. Li et al. (2012) suggested that while these can contribute to uncertainty in traffic flow model predictions, they are generally controllable and less dominant in comparison to uncertainty in the model parameter vector. However, this can only apply to current scenarios where real measurement data exist as input. Modelling of future scenarios often involves using modelled input (such as traffic demand) rather than direct measurement data. Also, in real-world applications, measurement data does not entirely match the

requirements of macroscopic traffic flow models, particularly in terms of the spatiotemporal resolution. For instance, measurement data are often of 30 seconds or one minute resolution, whereas input data required is at a resolution of modelling time step, which can in principle be one second. Accordingly, the resulting assumption creates additional uncertainty, even when the measurement data themselves are accurate and precise (i.e. error-free). One can argue, however, that these are model implementation uncertainty, rather than model input.

The traffic flow model parameter vector is also dependent on the selected model. The additional complexity of METANET implies additional parameters. As such, while the complexity might reduce model structure uncertainty, the latter arises in the additional parameters. This is basically due to difficulties in finding the 'global' solution to the optimisation problem. As has been described in § 2.2.2, such difficulties gave rise to multiple optimisation algorithms, multiple objective functions, and multiple performance assessment indicators. Of course, this is related to current scenarios, where an optimisation problem can be formulated. In the case of future scenarios, an optimal solution (even if it is not the 'global' one) cannot be determined due to lack of data, resulting in using previously calibrated parameter vectors on the same road network but for earlier time periods.

The above-mentioned sources are described separately for each model type. However, in this modelling chain, 'compatibility' uncertainty also arises. This is due to two reasons. First, macroscopic traffic flow models produce high spatiotemporal outputs yet aggregate for each cell. This induces uncertainty since COPERT is built based on trip-based average speeds, which implies longer distances than a cell and longer trip time than the time step of macroscopic traffic flow models. Second, in an ideal scenario, traffic density and speed predictions for each class of vehicles is obtained and is compatible with its corresponding emission factor. When using a uni-class model, predictions average out the contribution of each vehicle class. This means, even when they are assumed to correspond to either PCs or LDVs, there exist uncertainty in traffic speed assigned to them due to their aggregation in the first place.

Depending on interactions between the different sources described above, accumulated uncertainty from the traffic flow model structure, input, and parameter vector reflects the uncertainty in traffic flow model predictions (i.e. model outcome uncertainty). The accumulated uncertainty from the emission model structure, input, and parameter vector, including traffic flow model predictions, reflects the uncertainty in emission predictions. The latter cannot be transformed to prediction errors due to unavailability of the 'true' value of emissions, neither for current nor for future scenarios. However, the availability of real traffic data which allows the calculation of prediction errors of traffic flow models implies that the uncertainty in its outputs can be used to better understand those in emission predictions. Since traffic flow model predictions are the most recurring sources of uncertainty, as shown in Figure 3.7, the research

methodology of the next section focusses on quantifying the prediction errors of traffic flow models and propagating these to average speed-based emission predictions.

3.4 Five-phase methodological framework

In a typical scenario where uncertainty is not taken into account at all, traffic density and speed are predicted for a given road network and time period, and then fed into the emission model to obtain the 2-D emission prediction matrix. The main purpose here is to quantify uncertainty of each emission prediction value of the 2-D matrix in terms of a probability distribution. Based on the probability distribution obtained for each emission prediction value, the final most useful measure to uncertainty is its confidence interval at a specific confidence level. In order to do so for CTM and METANET-based emission predictions, this research has developed a data-driven methodological framework comprising one preliminary phase: data preparation, and four main phases: ensemble-based optimisation; prediction error estimation; grid-based Monte Carlo sampling; and uncertainty propagation.

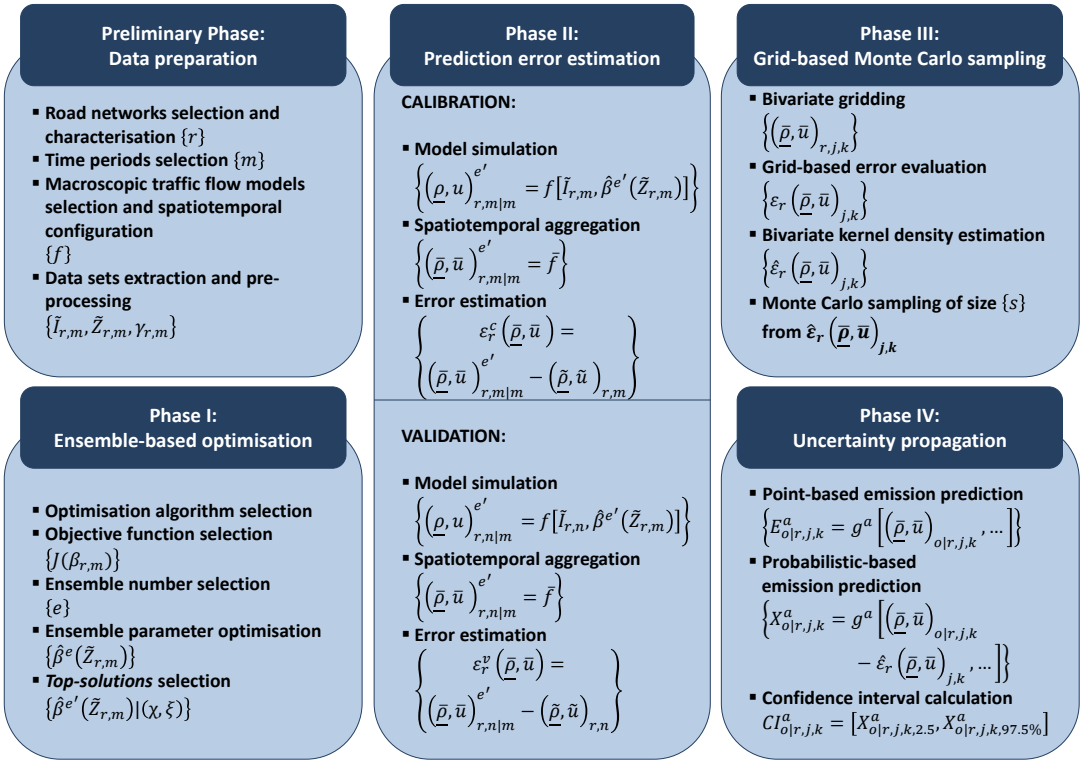


FIGURE 3.8: Five-phase data-driven methodological framework to uncertainty propagation from macroscopic road traffic flow models to average speed-based emission predictions. [c and v] in Phases III and IV are dropped for convenience.

Figure 3.8 provides a summary of key steps and notations arising in each of the five phases, and thus used as an icon in the following sections. The preliminary phase along with the first two phases are developed in order to quantify uncertainty in traffic flow model outputs. The last two phases are developed in order to propagate these to emission predictions. The following sections provide the rationale behind each phase and a detailed description of each step. Here, it is important to note that each of these phases are independent of which macroscopic traffic flow model structure is selected (i.e. f).

3.4.1 Preliminary Phase: Data preparation

The 'data preparation' preliminary phase follows directly from the description of CTM and METANET. Data preparation involves as a first step the selection and characterisation of a road network(s) and the selection of time periods to be studied for each road network. While the application of the methodological framework is not dependent on the number of road networks, multiple road networks are selected due to several reasons. These are as follows:

- to ensure transferability of results in the proposed methodological framework to other road networks - this is unlike most of the calibration and validation reviewed studies which have often focussed on a single road network;
- to identify the strengths and weaknesses of the proposed methodological framework which can either be road network-dependent or road network-independent;
- to ensure that both VSL-related traffic flow models as well as non-VSL related traffic flow models are tested;
- to test whether CTM-vsl and M-METANET-vsl model structures improve the performance of the basic ones given a VSL-operated road network.

This is directly followed by the selection of time periods. For the purposes of this study, multiple time periods for each road network is required for two main reasons.

1. It allows both calibrating each selected traffic flow model and validating them against other time periods. This eventually means the ability to capturing calibration and validation prediction error distributions, separately, and studying their influence on emission predictions. This will be expanded upon in Phase II, § 3.4.3.
2. It allows capturing prediction error distributions based on different traffic and ambient conditions (i.e. not restricted to a single day) of a road network. This implies the possibility of developing probabilistic-based emission predictions for each road network rather than each time period of a road network. This will be expanded upon in Phase II, § 3.4.3.

The selection and spatiotemporal configuration of macroscopic traffic flow models f follows directly, particularly, after identifying whether VSL is under operation or not

on selected road networks and time periods. The last step is data sets' extraction and any necessary pre-processing for each road network and time period. Data sets to be extracted need to include:

- **traffic data** required for the simulation (i.e. input data) of either CTM-basic and M-METANET-basic, or CTM-vsl and M-METANET-vsl, depending on VSL operation;
- **traffic data** required for the calibration of the selected traffic flow model parameter vectors; and
- **fleet mix data** required for the prediction of emissions.

Let R denote the total number of selected road networks and M_r denote the total number of selected time periods of road network r such that the pre-processed extracted data sets can be summarised by:

$$\left\{\tilde{I}_{r,m}, \tilde{Z}_{r,m} = (\tilde{q}, \underline{\tilde{q}}, \tilde{u}, \tilde{\rho}, \underline{\tilde{\rho}})_{r,m}, \gamma_{r,m}; r = 1, 2, ..., R; R \geq 1; m = 1, 2, ..., M_r; M_r > 1\right\}$$

where $\tilde{I}_{r,m}$ is measured input data dependent on the selected traffic flow model, with reference to Equation 3.9, 3.10, 3.20, or 3.21; $\tilde{Z}_{r,m}$ is measured traffic flow, flow per lane, speed, and estimated traffic density and density per lane data at all measurement sites located within the mainline boundaries required for parameter calibration; and $\gamma_{r,m}$ is the fleet mix data of road network r and time period m. As will be explained in Chapter 4, traffic density is not directly measured by loop detectors but rather estimated based on measurements of traffic flow and speed.

3.4.2 Phase I: Ensemble-based optimisation

Once traffic data sets required are prepared, an optimisation problem can be formulated and the parameter vector β for each road network r and time period m can be calibrated. However, there exist two issues that need to be tackled. Both are a result of the complexity and non-linearity of the optimisation problem, demonstrated earlier in § 2.2.2, which implies that a single 'global' solution is hard to obtain, if it exists. The first is to develop a systematic approach to the optimisation problem where the effort to reaching a 'good' solution are clearly set out. The second is to develop a rigorous approach to selecting a 'good' solution. These two issues are tackled through the ensemble-based optimisation phase, where a pre-defined number of optimisation runs (an ensemble) is performed and a statistical approach is used to selecting 'good' solutions. This phase allows to possibly get an ensemble of parameter vectors (rather than a single 'good' solution), each of which satisfies certain statistical criteria. However, one can treat this ensemble of parameter vectors in two ways. The first is to arbitrary select one and use it to estimate prediction errors. In selecting only one parameter vector from the ensemble one loses a lot of information on:

- the range and distribution of possible individual parameters;
- the joint variability of each two parameters (i.e. their covariance);
- the uncertainty around the particular choice of a parameter vector; and
- the possible relationship between prediction errors and predicted traffic variables themselves when data points are reduced to those from a single solution.

To avoid this, this phase makes use of each parameter vector in the ensemble. Perhaps, it is important to refer interested readers here to a study found to have developed an ensemble-based optimisation framework by Golaz et al. (2007) in the context of climate models, particularly, the modelling of clouds. Although developed in a different context and for different purpose, their framework was also motivated by the complexity of their optimisation problem.

In the context of this work, the ensemble-based optimisation phase includes five main steps, described next. These are determining a derivative-free optimisation algorithm, an objective function, and a number of optimisation runs forming an ensemble. This is followed by an ensemble parameter optimisation and *top-solutions* selection.

There exist a number of optimisation algorithms to select from. Algorithms range in complexity, computational cost, and data requirements. Here, despite the recognised deficiencies of the unconstrained N-M algorithm (i.e. Nelder-Mead, also referred to as downhill simplex) in terms of its possibility of falling to a 'false' minimum and its dependence on the initial parameter set, N-M is selected for four main reasons.

1. It is an off-the-shelf derivative-free method which implies it is simple and straightforward to use.
2. It is a relatively fast algorithm.
3. It does not require parameter bounds which are otherwise required, especially when using evolutionary methods such as GA. This is advantageous for the following reasons.
    - It allows the parameter vector space to be explored and any potential 'good' solution to be obtained. Although not using parameter bounds might allow the optimisation algorithm to explore 'unexpected' values, it is argued here that these are either disregarded after the application of the statistical approach or otherwise, if selected, can reveal information on the model structure.
    - It avoids making misleading assumptions. Although bounds on physical values are often based on previous observations, bounds on parameters such as those in the METANET model (other than the fundamental diagram ones) are often based on judgments/assumptions, which is disadvantageous/restrictive especially within an ensemble-based optimisation framework.
4. It has been shown by Spiliopoulou et al. (2015) that its results are comparable with those from more complex methods such as GA and CEM.

Detailed description of the method can be found in Nelder and Mead (1965), Powell (1998), or Ngoduy et al. (2004). Briefly, N-M searches for an unconstrained minimum of an objective function without the need to calculate its derivative (this is referred to as a direct search method). The algorithm starts with an initial simplex of size $k \geq n+1$, where n is the number of parameters. A simplex is defined as a $(k-1)$-dimensional geometrical shape with k vertices; each vertex corresponds to an objective function value of a parameter vector. The algorithm then iteratively performs a number of transformations (referred to as reflection, contraction, expansion, and shrinking) in order to find a better simplex which reduces the objective function values at its vertices. This iterative process terminates if two convergence criteria are met. The criteria are defined based on the specific implementation, which in this study is that of Bihorel and Baudin (2015). These are set as follows.

1. The size of the simplex is lower than a certain tolerance level. This is defined as the maximum length of the vector from each vertex of the simplex to the first vertex of the simplex. The tolerance level is set here to 0.1.
2. The absolute difference of the highest and lowest objective function values in the simplex is lower than a certain tolerance level. The tolerance level is set here to 10^{-4}.

Different studies have used different objective functions in terms of what traffic variable they use or the statistic to be minimised. In this research, the mean of the sum of absolute errors of measured versus predicted traffic speed, shown in Equation 3.26, is used where $\tilde{N}$ is the number of dual loop detectors (excluding the origin and destination ones); $\tilde{T}$ is the number of measurement time steps; $\tilde{u}_{\tilde{i}}(\tilde{t})$ is the measured traffic speed at loop detector $\tilde{i}$ and measurement time step $\tilde{t}$; and $\bar{u}_{\tilde{i}}(\tilde{t})$ is the space-time aggregated predicted traffic speed at loop detector $\tilde{i}$ and measurement time step $\tilde{t}$. Notice that from now on, any *aggregated* predicted traffic variable X is denoted as $\bar{X}$.

$$J(\beta_{r,m}) = \frac{1}{\tilde{N}} \frac{1}{\tilde{T}} \sum_{\tilde{i}=1}^{\tilde{N}} \sum_{\tilde{t}=1}^{\tilde{T}} \left[|\bar{u}_{\tilde{i}}(\tilde{t}) - \tilde{u}_{\tilde{i}}(\tilde{t})|_{r,m} \right] \tag{3.26}$$

Traffic speed is chosen in the objective function because it is directly measured by dual loop detectors unlike density which is estimated; and because traffic flow is relatively easier to predict in comparison to traffic speed as a result of the conservation equation (Spiliopoulou et al., 2014). $\tilde{N}$ is used since those at origin and destination are not predicted by the model. Similarly, those at time $\tilde{t} = 0$ are excluded as they are only predicted by CTM but not METANET. Perhaps the main assumption here, and, in fact, in the entire methodological framework is the spatiotemporal aggregation of the macroscopic traffic flow model outputs to match the resolution of the measurements, which in most cases are more aggregate. Temporally, traffic speed predictions are time-averaged to match the resolution of measurement data. Spatially, only cells

where measurement sites (i.e. dual loop detectors) exists are included in the objective function.

The last requirement for the optimisation is satisfying the CFL and additional parameter vector conditions of the CTM (four fundamental diagram properties) and METANET (positive parameters) set out earlier. Since parameter vector bounds are not set for N-M, penalty terms are added to the objective function which produces a high value if any of the conditions are violated; this ensures that all conditions are satisfied.

In defining the optimisation algorithm and objective function, the three ensemble-based-related steps follow. Given a traffic flow model, the approach entails running the algorithm multiple times, resulting in a set of optimal parameter vectors. Let E denote the number of optimisation runs, with $\hat{\beta}^e(\tilde{Z}_{r,m})$ being the optimal parameter vector obtained from run e, such that

$$\left\{\hat{\beta}^e(\tilde{Z}_{r,m}) = argmin\, J(\beta^e_{r,m}); e = 1, 2, ..., E; E \gg 1\right\}$$

Each run is initialised with a randomly sampled vector β^e, as required by the optimisation algorithm to initialise the simplex, within which each parameter is drawn from a pre-defined distribution. These distributions as well as the number of optimisation runs E are considered as phase inputs to be specified. Generally, the distributions can be based either on previous studies/subjective judgment or preferably on real traffic data, when parameters are physical/observable. As will be described empirically in § 5.2, this study makes use of real traffic data in order to infer the initial distribution of each parameter related to the fundamental diagram. In the case where the parameters are unobservable (five parameters in METANET), distributions are assumed similar to those suggested in previous studies.

Some of the optimisation runs might result in poor solutions. In an attempt to improve the model assessment process and disregard poor solutions automatically, this study proposes the following. First, objective function values obtained from all optimisation runs are stored and ordered. Those below a threshold percentile χ of all objective function values are kept, and their parameter vectors are extracted. Second, for each remaining parameter vector (i.e. a total of $\chi \times E$), the model is simulated and the time-averaged predicted traffic density per lane and speed at measurement locations are obtained. Two statistical hypothesis tests are proposed here and applied to predicted versus measured traffic variables. Both of which are non-parametric in order to avoid making assumptions of underlying distributions obtained from each optimisation run. These are as follows.

1. The non-parametric Mann-Whitney (Bauer, 1972) to test whether or not the distribution of a predicted traffic variable differs from that of a measured traffic variable by a location shift at a pre-defined significance level ξ (significance level

is normally denoted as α but this has been used earlier for the METANET fundamental diagram). The null hypothesis here is that the true location shift is 0. So, if the resulting p-value is lower than ξ, then the null hypothesis is rejected and the true location shift is not equal to 0. If the resulting p-value is higher than ξ, then the null hypothesis, that the true location shift is 0, is not rejected.

2. The non-parametric Fligner-Killeen (Conover et al., 1981) to test whether or not the variances of a predicted and measured traffic variable differ at a pre-defined significance level ξ. The null hypothesis here is that the variances are the same. So, if the resulting p-value is higher than ξ, then the null hypothesis, that the variances are the same, is not rejected.

The corresponding parameter vector $\hat{\beta}^e(\tilde{Z}_{r,m})$ is selected as a *top-solution* if one or both of the following two conditions is satisfied:

1. Both null hypotheses are not rejected based on the predicted and measured *traffic density per lane*. This is referred to as Condition 1 or C1 from now on.
2. Both null hypotheses are not rejected based on the predicted and measured *traffic speed*. This is referred to as Condition 2 or C2 from now on.

The above ensures that the difference in distribution and variance in at least one traffic variable (traffic density per lane or speed) is not statistically significant. Requiring the predictions to satisfy both conditions simultaneously turned out to be too demanding, especially for the case of CTM which only captures the dynamics of traffic density and assumes steady-state traffic speed-density conditions. This will be illustrated in § 5.2. Hence, satisfying one condition is assumed here to be enough.

Let E' denote the number of *top-solutions* such that,

$$\left\{\hat{\beta}^{e'}(\tilde{Z}_{r,m})|(\chi,\xi); e' = 1, 2, ..., E'; E' < E; 0 < \chi < 100\%; 0 < \xi < 1\right\}$$

By obtaining the *top-solutions* for each road network r, time period m, and traffic flow model f, the second 'prediction error estimation' phase can be undertaken.

To summarise, this 'ensemble-based optimisation' phase requires six inputs: optimisation algorithm, objective function, distribution of initial parameters required for the optimisation algorithm, ensemble number E, threshold percentile χ, and significance level ξ. While the first two are determined and explained in detail here, the last four inputs are to be set out in the results § 5.2. Keeping in mind that this phase is not restricted to the choice of the first two inputs. The outputs of this phase are the *top-solutions* and are the main input to the next phase.

3.4.3 Phase II: Prediction error estimation

The outputs (or predictions) of simulating the *top-solutions* of road network r, time period m, and traffic flow model f using the input data of road network r and time period m are referred to as the *calibration traffic flow model outputs*. However, the outputs of simulating those using the input data of road network r and time period n $\{n = 1, 2, ..., M_r - 1; n \neq m\}$ are referred to as the *validation traffic flow model outputs*. For instance, if $r = 1$, then

- $f\left[\tilde{I}_{1,1}, \hat{\beta}^{e'}(\tilde{Z}_{1,1})\right]$ gives the calibration outputs for time period $m = 1$; and
- $f\left[\tilde{I}_{1,2}, \hat{\beta}^{e'}(\tilde{Z}_{1,1})\right]$ gives the validation outputs of time period $n = 2$ given a *top-solution* e' from time period $m = 1$.

As such, the 'prediction error estimation' phase is divided into two sub-phases: 'calibration' and 'validation' prediction error estimation. The concept behind doing so here and in subsequent phases is to distinguish between:

1. uncertainty in emission predictions for a current scenario, when one can optimise the traffic flow model given the availability of traffic data; and
2. uncertainty in emission predictions for a future scenario, when one cannot optimise the traffic flow model given the absence of traffic data.

As a result, it is assumed here that validation outputs and their corresponding prediction errors are best approximation to uncertainty in emission predictions for future scenarios. Although it is acknowledged that additional uncertainty in traffic flow model structure and input in future scenarios can arise. This additional uncertainty can be a result of known changes yet unknown impacts, or even a result of both unknown changes and unknown impacts. For instance, a future scenario can incorporate a potential change in infrastructure (e.g. additional on-ramp). This is a known change incorporated in the modelling scenario, but there is additional uncertainty in for example the on-ramp input data and in expected impact on traffic dynamics. Alternatively, a future scenario might not incorporate a change in infrastructure, yet there is a possibility that it does happen but neither the planner nor the modeller can anticipate its happening. Validation errors here do not take these additional uncertainties into account.

Estimating prediction errors involves three main steps: model simulation, spatiotemporal aggregation, and error estimation. Given the *top-solutions* $\hat{\beta}^{e'}(\tilde{Z}_{r,m})$, let $(\underline{\rho}, u)^{e'}_{r,m|m}$ and $(\underline{\rho}, u)^{e'}_{r,n|m}$ denote the calibration and validation traffic density and speed outputs at all cells i and all time steps t respectively, such that,

$$\left\{(\underline{\rho}, u)^{e'}_{r,m|m} = f\left[\tilde{I}_{r,m}, \hat{\beta}^{e'}(\tilde{Z}_{r,m})\right]\right\}$$

$$\left\{(\underline{\rho}, u)^{e'}_{r,n|m} = f\left[\tilde{I}_{r,n}, \hat{\beta}^{e'}(\tilde{Z}_{r,m})\right]; n = 1, 2, ..., M_r - 1; n \neq m\right\}$$

As a result of typically higher spatiotemporal resolution in comparison to real measurements explained earlier, predicted outputs on cells where measurement sites exist

are extracted and time-averaged to match the temporal resolution of real measurements. Let $(\underline{\bar{\rho}}, \bar{u})^{e'}_{r,m|m}$ and $(\underline{\bar{\rho}}, \bar{u})^{e'}_{r,n|m}$ denote the aggregated predicted traffic density and speed outputs required for the estimation of calibration and validation errors respectively, such that,

$$\left\{(\underline{\bar{\rho}}, \bar{u})^{e'}_{r,m|m} = \bar{f}\left[\tilde{I}_{r,m}, \hat{\beta}^{e'}(\tilde{Z}_{r,m})\right]\right\}$$

$$\left\{(\underline{\bar{\rho}}, \bar{u})^{e'}_{r,n|m} = \bar{f}\left[\tilde{I}_{r,n}, \hat{\beta}^{e'}(\tilde{Z}_{r,m})\right]\right\}$$

In principle, prediction errors of each traffic variable can be estimated based on the calibration and validation outputs, separately, given a *top-solution* for time period m and road network r. It is argued here that while such errors account for the total accumulated uncertainty in the simulated traffic flow model, it assumes a particular choice of *top-solution* and time period. However, this choice amongst other potential *top-solutions* and other time periods, is typically unknown, and the different *top-solutions* and time periods might lead (but not necessarily) to different prediction errors in calibration and/or validation. Hence, it is argued here that using a single prediction error distribution for all *top-solutions* and time periods on a road network, takes into account this additional uncertainty in the parameter and time period choice to modelling traffic flow on a typical day. The influence of this argument on the uncertainty in emission predictions is discussed in Phase III, § 3.4.4.

Given the measurements of traffic variables, let the 2-D matrices of the two calibration error distributions and two validation error distributions of road network r be denoted as $\varepsilon^{c}_{r}(\underline{\bar{\rho}}, \bar{u})$ and $\varepsilon^{v}_{r}(\underline{\bar{\rho}}, \bar{u})$ respectively, such that,

$$\left\{\varepsilon^{c}_{r}(\underline{\bar{\rho}}, \bar{u}) = (\underline{\bar{\rho}}, \bar{u})^{e'}_{r,m|m} - (\underline{\tilde{\rho}}, \tilde{u})_{r,m}; \forall e', m\right\}$$

$$\left\{\varepsilon^{v}_{r}(\underline{\bar{\rho}}, \bar{u}) = (\underline{\bar{\rho}}, \bar{u})^{e'}_{r,n|m} - (\underline{\tilde{\rho}}, \tilde{u})_{r,n}; \forall e', m, n\right\}$$

As a result of the dependence of traffic speed on density and the consequent dependence of traffic speed errors on density errors, as will be illustrated empirically in § 5.3, the above are treated as two bivariate traffic speed-density error distributions corresponding to two bivariate traffic speed-density predictions (calibration and validation). These are the outputs of this phase and the main input to the next 'grid-based Monte Carlo sampling' phase.

3.4.4 Phase III: Grid-based Monte Carlo sampling

Given the traffic flow model outputs and their prediction errors, uncertainty in emission predictions can be determined for every road network r, time period m, and *top-solution* e'. However, in considering a single bivariate error distribution for all *top-solutions* and time periods, uncertainty in emission predictions does not have to be restricted to the specific traffic speed and density

outputs of an (unknown) choice of a *top-solution* and time period, but rather on where these points lie in the feasible prediction region. Doing so takes advantage of the population of prediction points and their errors obtained earlier across all its time periods (one main reason why $M_r \geq 1$ is recommended) and its *top-solutions* (one main reason behind the ensemble-based optimisation approach). The following describes four steps of the first phase to achieving this. These are bivariate gridding, grid-based error evaluation, bivariate kernel density estimation, and Monte Carlo sampling.

The resulting error distributions vary across the feasible region of traffic speed-density predictions; errors occurring at high speed free-flow conditions are systematically different from those associated with highly congested, low speed conditions. This will be empirically demonstrated in § 5.4. Therefore, it is not appropriate to associate the entire bivariate error distribution with the entire feasible prediction region. Alternatively, a bivariate error distribution of traffic speed-density could be associated with each traffic speed-density point, but estimating this would require incredibly high data density throughout the feasible prediction region. Hence, a grid-based approach is instead proposed here, which allows to efficiently take into account such variation across the feasible prediction region. A grid-based approach entails dividing the feasible prediction region into small enough grids in each direction (traffic speed and density).

Let $(\underline{\bar{\rho}}, \bar{u})_{r,j,k}$ define the $[j,k]$ grid square of traffic speed-density prediction region of road network r. The bivariate error distribution associated with the $[j,k]$ grid square can easily be evaluated from the outputs of Phase II. These are considered the first two steps of this phase. Let this be denoted as,

$$\left\{\varepsilon_r(\underline{\bar{\rho}}, \bar{u})_{j,k}; j = 1, 2, ..., J; k = 1, 2, ..., K\right\}$$

where J is the number of grids in traffic density and K is number of grids in speed. Both can, but not necessarily, be equal. Also, both are phase input to be specified during implementation.

While in principle, bivariate sampling can directly be undertaken from the grid-based error distributions, a bivariate kernel density estimation for each grid square is proposed in order to smooth the original data and calculate its bivariate probability density function. At the same time, an analytical approximation of each distribution (for example, a bivariate normal) would not be needed and in fact, is not recommended since it makes huge assumptions on the underlying distributions of bivariate errors in each $[j,k]$ grid square, the total of which can be large depending on the feasible region and the choice of J and K. Given that, let the estimated bivariate kernel density be denoted as,

$$\left\{\hat{\varepsilon}_r(\underline{\bar{\rho}}, \bar{u})_{j,k}\right\}$$

The most useful method to propagating the bivariate empirical error distributions to emission predictions in the final phase is through numerical simulation. This is

because of the complexity of the emission model in terms of its non-linear vehicle category-specific emission factors' functions which prevents the use of analytical or approximation methods. Several numerical simulation methods are available (Cullen & Frey, 1999). The well-established Monte Carlo simulation is proposed here. The first step to Monte Carlo simulation is sampling. The second step is the simulation and is described in the next phase. Using the estimated bivariate kernel density of each grid square, a Monte Carlo random sample of the bivariate error distribution with size s can be generated and associated with the corresponding gridded bivariate predicted speed-density region. The sample size s depends on the confidence level required and is determined in the next 'uncertainty propagation' phase.

This is considered the final step of Phase III. Its only input is the gridding variables J and K, which will be set out in § 5.4. By applying the above step on the bivariate calibration errors and validation errors of each road network, separately, the main outputs of this phase are s-sample points from each of,

$$\left\{\hat{\varepsilon}_r^c(\underline{\rho},\bar{u})_{j,k}\right\}$$

$$\left\{\hat{\varepsilon}_r^v(\underline{\rho},\bar{u})_{j,k}\right\}$$

3.4.5 Phase IV: Uncertainty propagation

The final phase of the methodological framework is the 'uncertainty propagation' phase. It involves three steps. The first one is point-based prediction of emissions, given a particular air pollutant, over the entire traffic speed-density region of road network r. Given that predictions over the entire feasible region is of interest rather than only the exact prediction points, point-based emissions need to be calculated at each point in the region. This can be specified depending on the resolution (or smoothness) of predictions required across the region. It can be for example, every 1 km h^{-1} in traffic speed and 1 veh km^{-1} in traffic density per lane.

Let $E^a_{o|r,j,k}$ denote the predicted emissions for air pollutant a evaluated at a traffic speed-density point o within the $[j,k]$ grid square based on Equation 3.24 such that,

$$\left\{E^a_{o|r,j,k} = g^a\left[\gamma_r,(\underline{\rho},\bar{u})_{o|r,j,k},ef^a\left(\bar{u}_{o|r,j,k},\omega^a\right)\right];o\in[r,j,k]\right\}$$

The uncertainty at each traffic speed-density point can be quantified by adding the corresponding sampled errors of the same grid square. Depending on the sample size s, each point o is therefore replaced by s points which represent the captured uncertainty of point o. The Monte Carlo simulation follows directly. The emission model is run s times given the s traffic speed-density points. This is considered the probabilistic-based emission prediction at point o in comparison to the point-based one obtained earlier.

Let $X^a_{o|r,j,k}$ denote the s-sample distribution of emission predictions at the traffic speed-density point o in $[j,k]$ grid square such that,

$$\left\{ X^a_{o|r,j,k} = g^a \left[\gamma_r, (\underline{\rho}, \bar{u})_{o|r,j,k} - \hat{\varepsilon}_r(\underline{\rho}, \bar{u})_{j,k}, ef^a\left(\bar{u}_{o|r,j,k} - \hat{\varepsilon}_r(\bar{u})_{j,k}, \omega^a\right) \right] \right\}$$

The entire distribution at each point o is not exactly informative, given the large number of points o in any feasible region. Hence, the 2.5^{th} and 97.5^{th} percentiles of the distribution $X^a_{o|r,j,k,2.5\%}, X^a_{o|r,j,k,97.5\%}$ can be evaluated and considered the lower and upper bounds of the emission prediction at 95% confidence level; this is the last uncertainty propagation step. Let this Confidence Interval (CI) be denoted as,

$$CI^a_{o|r,j,k} = [X^a_{o|r,j,k,2.5}, X^a_{o|r,j,k,97.5\%}]$$

Based on this CI, the additive upper and additive lower bounds can also be calculated by subtracting it from the point-based emission prediction at the same point i.e. $E^a_{o|r,j,k}$.

The selected sample size s determines the precision of calculated percentiles (Morgan et al., 1992); to be 95% confident that the actual 97.5^{th} (or 2.5^{th}) percentile is between the estimates of 96.5^{th} and 98.5^{th} (or 1.5^{th} and 3.5^{th}) percentiles, a sample size $s = 1000$ was found to be sufficient.

Evaluating the lower and upper bounds at equally spaced traffic speed-density prediction points in each grid square provides a surface map of the confidence interval of the point-based emission predictions of air pollutant a throughout the entire feasible prediction region.

By applying the described three steps using the calibration and validation sampled grid-based error distributions separately, two surface maps of confidence intervals can be deduced for any road network. The only input to doing so is the desired resolution of equally spaced points o in any grid-square. This will be set out in § 5.5. The surface maps can be used to predict emissions at the spatiotemporal resolution of measurements alongside their lower and upper bounds, given a specific time period and *top-solution*. They can also be used to predict total emissions alongside their lower and upper bounds. An example showing the different application of the developed surface maps is provided separately in the final section of the results, § 5.6.

3.5 Summary

This chapter serves the first objective of this research: developing a data-driven methodological framework that enables quantifying and propagating uncertainty in the outputs of macroscopic traffic flow models to emission predictions. The chapter starts by defining the 'traffic flow - emission' modelling chain to be used as the basis to developing and applying the methodological framework. The final described modelling chain can take four forms: CTM-basic, CTM-VSL, M-METANET-basic, and M-METANET-vsl

average speed-based emission modelling. In describing the modelling chain, two main points were argued but will be illustrated empirically in § 5.2.2 and § 5.3.1, respectively.

1. The use of a modified fundamental diagram in METANET.
2. The use of aggregated traffic density rather than flow-based density to estimate emissions.

The chapter follows by describing in details all the sources of uncertainty arising in the emission and traffic flow model structure, input, and parameter vectors as well as those arising from their integration, referred to as 'compatibility' uncertainty. Describing the modelling chain and it sources of uncertainty have shown that the outputs of macroscopic traffic flow models are the most recurring input to predicting emissions. Given the availability of real traffic data (assumed to the 'true' values), the total accumulated uncertainty from macroscopic traffic flow models can be estimated and propagated. As such, the focus on uncertainty in traffic data input is justified and a methodological framework to doing so is established in a five data-driven phases. Each phase was described in detail. The final developed framework is quite generic and is not dependent on the specific choice of the four above-mentioned forms of traffic flow models or the specific choice of COPERT as an average speed-based emission model. Hence, the framework is equally valid given a variation of models under the category of macroscopic traffic flow or average speed-based emission models. Also, as noticed there is not any dependence on the locations of road networks or the source of traffic data. The framework is applicable regardless of the location or data source.

In describing the phases of the methodological framework, two main arguments were also made but will be illustrated empirically in § 5.3.1 and § 5.4.1, respectively.

1. The need for bivariate analysis of traffic flow model outputs.
2. The need for a grid-based approach to sampling prediction errors.

One of the main limitations of the developed framework lies in the ensemble-based optimisation of phase I, where it is often necessary to aggregate traffic flow model outputs to the measurement resolution for parameter optimisation. As a result, an understanding of the errors at the resolution of traffic flow modelling outputs are not entirely understood. Nevertheless, the developed methodological framework does not depend on any particular measurement resolution i.e. obtaining higher resolution measurement data improves the implementation of the framework, but not the framework itself.

To conclude, the results of applying the methodological framework to real road networks are provided in Chapter 4 for the first preliminary phase, and in Chapter 5 for the four main phases.

Chapter 4

Data Preparation Preliminary Phase

4.1 Introduction

Chapter 3 provided a detailed description of the methodological framework, which consists of five connected phases. This chapter focusses only on the data preparation preliminary phase of the methodological framework, as it underpins the entire methodological framework. The data preparation phase involves four main steps: selection and characterisation of road networks; selection of time periods; selection and space-time configuration of macroscopic traffic flow models; and extraction and pre-processing of corresponding data sets. These four steps are not entirely disconnected; the selection of a road network is largely determined by the existence of long-term historical data, necessary for consequent phases. Hence, this chapter starts with § 4.2 by providing a generic description of historical traffic data that the selection of road networks and time periods can then be based on. This is followed by describing additional data of fleet mix. The generic description includes the data sources, the data extraction process, and all the necessary data pre-processing, such as its manipulation and transformation. These are common for any time period of any road network. Section § 4.3 follows on to provide a detailed description of:

- the selected road networks;
- the selected time periods; and
- the selected macroscopic traffic flow models.

With the data sets from the different identified sources being selected, extracted, and pre-processed, a summary of the data is provided in § 4.4 and § 4.5. The data summary mainly helps in setting a background for the analysis of the results, for Phases I to IV, presented in Chapter 5.

4.2 Data sources, extraction, and pre-processing

The developed methodological framework requires two types of data: traffic data and fleet mix data. Traffic data are needed to calibrate and validate the traffic flow models (and/or simulate the traffic flow models). Fleet mix data are needed for the prediction of emissions. The present section looks into the sources of such data and the generic data handling approach adopted.

4.2.1 Traffic data

The traffic data required mainly includes average traffic flow, average traffic speed, average traffic density, and speed limits, if VSL is under operation. The *M*otorway *I*ncident *D*etection and *A*utomatic *S*ignalling (MIDAS) network is a primary source of traffic data for England's motorway road network, which is operated by Mott MacDonald on behalf of Highways England (Mott MacDonald, 2016). MIDAS is a network of more than 8,000 temporary and permanent (and/or long-term) dual loop detector measurement sites (hereafter referred to as measurement sites) installed and distributed along England's motorway road network, as shown in Figure 4.1 below. These measurement sites are divided into seven different regions, reflecting *R*egional *C*ontrol *C*entres (RCCs) across different geographical locations.

The MIDAS data base contains four years of historical traffic data, which is accessible to users upon request. A single data file within the data base includes traffic data from all measurement sites located within a selected region, for a selected day. A measurement site is distinguished by a unique identifier (referred to as "Geographic.Address") reflecting the motorway name (for example, "M6"), distance from a motorway-specific nominal starting point (for example, "6615"), direction of traffic (for example, "A" indicating clockwise direction), and whether it is located on a motorway's mainline, on-ramp, or off-ramp (for example, "A" indicating a motorway mainline). The traffic data of each measurement site during an entire day includes the following:

- Number of vehicles of four different vehicle categories passing the measurement site each minute irrespective of the lane. The vehicle categories are based on the estimated vehicle length. These are vehicles with lengths less than 5.2 metres, between 5.2 and 6.6 metres, between 6.6 and 11.6 metres, and more than 11.6 metres.
- Average speed of all vehicles passing each lane of the measurement site per minute. Average speed is calculated arithmetically and is reported in [km h^{-1}].
- Number of vehicles passing each lane of the measurement site per minute. The number of vehicles across all lanes is equal to the number of vehicles of all four categories.

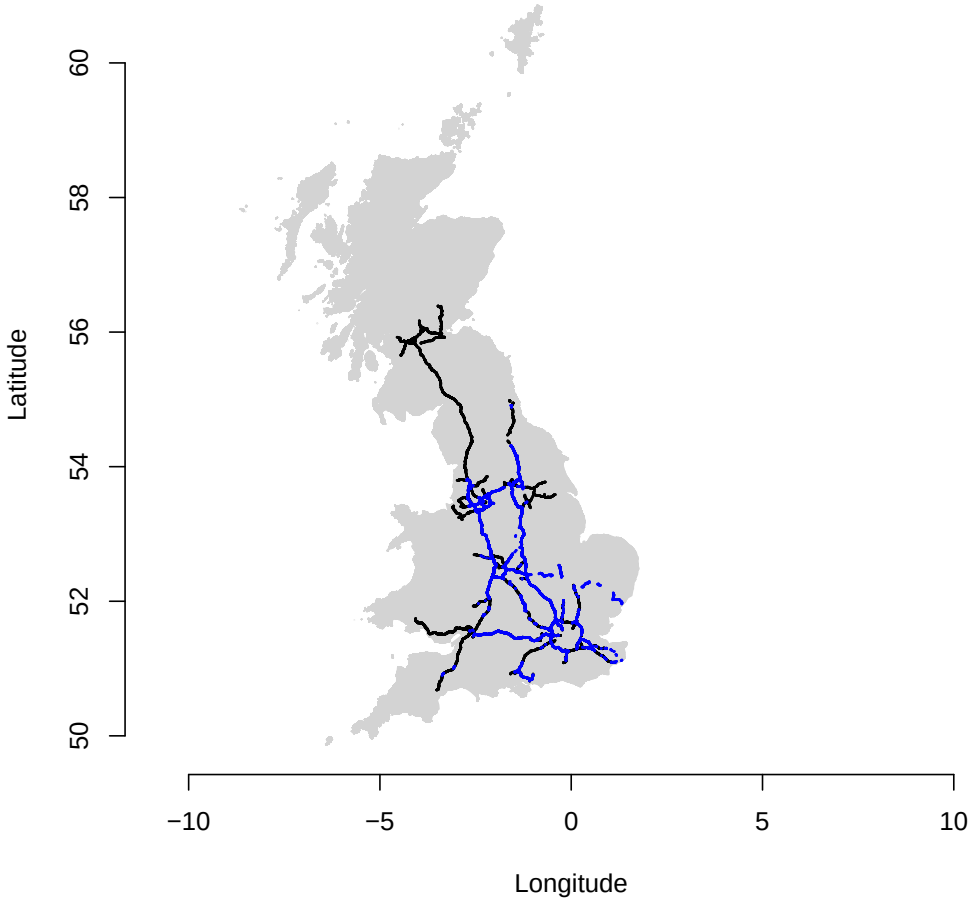


FIGURE 4.1: Great Britain map (Strategi®, 2015) overlaid by England's motorway road network managed by Highways England in black (Boundary-Line™, 2016) and the MIDAS network of dual loop detector measurement sites (8,853 in total as of 2014) in dark blue which is operated by Mott MacDonald on behalf of Highways England (Mott MacDonald, 2016).

- Average occupancy for each lane of the measurement site per minute. This is reported in percentage, which signifies the percentage of time a lane is occupied by vehicles.
- Average headway for each lane of the measurement site per minute. This is reported in metres, which signifies the average distance between passing vehicles.

Here, only the first three are to be extracted and pre-processed. An example of what a data file looks like for a particular region and a particular day is provided here. This shows the first five rows of the first three items explained above. The average occupancy and average headway columns alongside additional equipment-related columns are not shown here. The example shows that measurement site "M6/6615A" has three lanes. Hence, data columns for lanes four to seven are set to 'NA', with seven lanes being the maximum number of lanes a measurement site can be located on.

```
## example of raw traffic data - top 5 rows
```

```
##   Geographic.Address     Date  Time Number.of.Lanes Flow.Category.1. Flow.Category.2. Flow.Category.3. Flow.Category.4.
## 1          M6/6615A 16/04/13 00:00               3                2                1                0                4
## 2          M6/6615A 16/04/13 00:01               3                3                0                1                6
## 3          M6/6615A 16/04/13 00:02               3                2                0                1                2
## 4          M6/6615A 16/04/13 00:03               3                2                0                2                2
## 5          M6/6615A 16/04/13 00:04               3                4                0                1                2
##   Speed.Lane.1.  Speed.Lane.2. Speed.Lane.3.  Speed.Lane.4. Speed.Lane.5. Speed.Lane.6. Speed.Lane.7.
## 1            90            131           NA            NA            NA            NA            NA
## 2            90            120           NA            NA            NA            NA            NA
## 3           106            129           NA            NA            NA            NA            NA
## 4            95            125           NA            NA            NA            NA            NA
## 5            99            144           NA            NA            NA            NA            NA
##   Flow.Lane.1. Flow.Lane.2. Flow.Lane.3. Flow.Lane.4. Flow.Lane.5. Flow.Lane.6. Flow.Lane.7.
## 1            5            2            0           NA           NA           NA           NA
## 2            7            3            0           NA           NA           NA           NA
## 3            3            2            0           NA           NA           NA           NA
## 4            4            2            0           NA           NA           NA           NA
## 5            5            2            0           NA           NA           NA           NA
```

For each measurement site, few transformations on each row of data have to be made in order to calculate the following:

- average traffic flow across all lanes ($\tilde{q}$) in [veh h^{-1}];
- average traffic flow per lane ($\underline{\tilde{q}}$) in [veh h^{-1}];
- average (space-mean) traffic speed across all lanes ($\tilde{u}$) in [km h^{-1}];
- average traffic density across all lanes ($\tilde{\rho}$) in [veh km^{-1}]; and
- average traffic density per lane ($\underline{\tilde{\rho}}$) in [veh km^{-1}].

Average traffic flow is easily calculated using Equation 4.1; summing the number of vehicles for each lane ($\tilde{n}_{\lambda_i}$) and multiplying by 60 minutes per hour since the time resolution of the measurements is one minute. Dividing by the total number of lanes allows the calculation of the average traffic flow per lane.

$$\tilde{q} = 60 \sum\nolimits_i (\tilde{n}_{\lambda_i}) \qquad \forall i \tag{4.1a}$$

$$\underline{\tilde{q}} = \tilde{q}/\lambda \tag{4.1b}$$

The space-mean traffic speed across all lanes is required as an input for the macroscopic traffic flow models. It can be calculated by averaging speeds of individual vehicles over a stretch of road during a particular time interval, or otherwise, by taking the harmonic mean of speeds of individual vehicles passing a particular point in space during a particular time interval. The latter is under the assumption of stationary and homogeneity state (Van Lint, 2004). The harmonic mean of any quantity x with n number is defined as:

$$n / \sum\nolimits_i (1/x_i) \qquad \forall i$$

If the harmonic mean traffic speed for each lane is reported by the raw data, then space-mean traffic speed across all lanes can be easily calculated (i.e. by calculating their harmonic mean, again). However, raw data provides the arithmetic mean traffic speed for each lane. This is also referred to as the time-mean traffic speed. The arithmetic

mean of any quantity x with n number is defined as:

$$\sum_i (x_i)/n \qquad \forall i$$

Hence, the calculation of the space-mean traffic speed across all lanes using the time-mean traffic speed for each lane is not straightforward and an approximation is required. Perhaps, the simplest but most inadequate approach is to calculate the arithmetic mean of raw speed data for each lane ($\tilde{u}_{\lambda_i}$), as represented in Equation 4.2a, and to assume that this is equivalent to the space-mean traffic speed across all lanes, required for the models.

$$\tilde{u} = \sum_i \tilde{u}_{\lambda_i}/\lambda \qquad \forall i \tag{4.2a}$$

$$\tilde{u} = \sum_i (\tilde{n}_{\lambda_i}\tilde{u}_{\lambda_i})/\sum_{\lambda_i} \tilde{n}_{\lambda_i} \qquad \forall i \tag{4.2b}$$

$$\tilde{u} = \sum_i \tilde{n}_{\lambda_i}/\sum_i (\tilde{n}_{\lambda_i}/\tilde{u}_{\lambda_i}) \qquad \forall i \tag{4.2c}$$

Given that the number of vehicles passing each lane is provided, Equation 4.2a can be improved by calculating the weighted arithmetic mean of raw traffic speed data for each lane, as suggested by Treiber and Kesting (2012). This is represented by Equation 4.2b. The assumption here is that the weighted arithmetic mean is equivalent to the space-mean traffic speed required, and is considered as a better approximation in comparison to the non-weighted arithmetic mean traffic speed.

Poole and Kotsialos (2016), have instead suggested the approximation of space-mean traffic speed across all lanes by assuming that all vehicles passing a particular lane have the same speed. For instance, if 50 [km h^{-1}] represents the arithmetic mean of 3 vehicles passing a particular point in space during 7:00 AM and 7:01 AM, then it is assumed that each of the 3 vehicles have passed with a speed of 50 [km h^{-1}]. With this assumption, the reported arithmetic mean traffic speed for each lane is equivalent to the harmonic mean for each lane. Based on that, the space-mean traffic speed across all lanes is then calculated using Equation 4.2c. A number of studies (Van Lint & Van der Zijpp, 2003; Van Lint, 2004; Rakha & Zhang, 2005; Seetharaman et al., 2011, for example) have focussed on developing statistical models which allow estimating space-mean traffic speed, given the time-mean that is typically reported by dual loop detectors. Such conversion methods rely heavily on the variance about the time-mean (or space-mean) traffic speed, which is not directly measured in the case of the MIDAS network. Here, the focus is on the first three approaches.

A comparison between the non-weighted arithmetic mean traffic speed of Equation 4.2a, the weighted arithmetic mean of Equation 4.2b, and the approximate harmonic mean of Equation 4.2c is shown in Figure 4.2 below. The comparison is based on six hours traffic speed data extracted for a 3.1 kilometres stretch on the M60 motorway consisting of eight measurement sites.

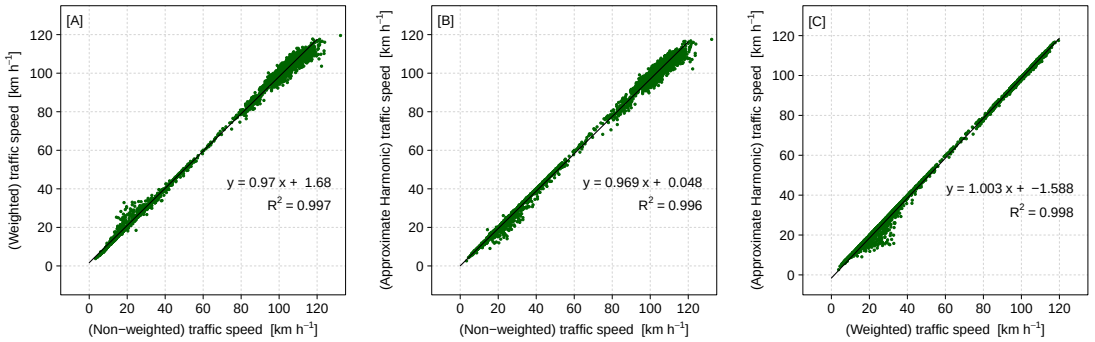


FIGURE 4.2: Example of the relation between average traffic speed [km h^{-1}] calculated in three different ways (Equation 4.2): [A] non-weighted versus weighted arithmetic mean; [B] non-weighted arithmetic versus approximate harmonic mean; and [C] weighted arithmetic versus approximate harmonic mean. The linear regression (black line), its equation ($y = ax + b$), and the coefficient of determination (R^2) are reported on each. 1 minute data are extracted for a 3.1 kilometres stretch on the M60 motorway consisting of 8 measurement sites for a period of 6 hours on 13 April 2016.

Figure 4.2 does not compare the above calculations with the ground-truth space-mean traffic speed as this is unavailable; it does, however, provide a comparison of the magnitude of differences from available methods. Figure 4.2A and Figure 4.2B compare the weighted arithmetic mean and the approximate harmonic mean with the non-weighted arithmetic mean respectively. They show high distances to their linear fits during both congested time periods (traffic speed between 10 [km h^{-1}] and 40 [km h^{-1}]) and high traffic speeds. While, Figure 4.2C compares the approximate harmonic mean with the weighted arithmetic mean traffic speed. The latter shows relatively high distances to the linear fit only during congested time periods. Figure 4.2C also shows that the approximate harmonic mean traffic speed is always lower than the weighted mean, by approximately 1.588 [km h^{-1}] (given that the linear fit slope is approximately one).

In order to avoid making the rather strong assumption that all vehicles passing through a lane have the same speed, the weighted arithmetic mean traffic speed is used. The latter is considered a better estimate in comparison to the non-weighted arithmetic mean which is sometimes used in calibration and validation studies. Yet, in the presence of space-mean traffic speed or a better approximation, the methodological framework remains equally valid.

Given the average traffic speed, the average traffic density across all lanes is then calculated using the speed-density relationship of Equation 4.3 (Ngoduy & Maher, 2012). Dividing the total number of lanes at a measurement site allows the calculation of the average traffic density per lane.

$$\tilde{\rho} = \tilde{q}/\tilde{u} \tag{4.3a}$$

$$\tilde{\underline{\rho}} = \tilde{\rho}/\lambda \tag{4.3b}$$

The above constitutes the major raw data transformations applied to traffic data used in this study. While the data extraction and pre-processing scripts are compiled in an open source R package[1], a flow-diagram and a description of the MIDAS-based developed functions and their integration with the traffic flow modelling functions are provided in Appendix A.

In addition to the MIDAS network being the major source of traffic data, geographical coordinates (based on the World Geodetic System 1984 or WGS84) for each measurement site have also been separately provided by Highways England, using the Network and Asset Model of the *N*ational *T*raffic *I*nformation *S*ervice (NTIS) of 2013 (Highways England, 2016). These are merged with the traffic data in order to identify the locations of measurements sites on the selected road network as well as the distance between consecutive sites. In examining the geographical coordinates on geographical maps, two issues arose:

- the geographical coordinates do not always match the location of the dual loop detectors identified on pavement; and
- the distance between dual loop detectors is not necessarily 500 metres (0.5 kilometre) as is often assumed to be the case.

The exact geographical coordinates of measurement sites on pavement has been identified using the street view on Google Earth (Google Inc., 2016) and used in the analysis rather than the geographical coordinates provided by NTIS. Also, the actual distance between sites are calculated rather than using the assumption of 500 metres distance. These two issues and the extent of these discrepancies is expanded upon in § 4.5.

An example extract of a final data set is provided here. This shows the same five rows (presented earlier) after data transformations have led to the calculation of average traffic flow, flow per lane, speed, density, and density per lane.

```
## example of traffic data after pre-processing - top 5 rows
```

```
##   Geographic.Address Number.of.Lanes            datetime    lon   lat
## 1           M6/6615A               3 2013-04-16 00:00:00 -2.335 53.07
## 2           M6/6615A               3 2013-04-16 00:01:00 -2.335 53.07
## 3           M6/6615A               3 2013-04-16 00:02:00 -2.335 53.07
## 4           M6/6615A               3 2013-04-16 00:03:00 -2.335 53.07
## 5           M6/6615A               3 2013-04-16 00:04:00 -2.335 53.07
##   avg.flow avg.flow.per.lane avg.spd avg.den avg.den.per.lane
## 1      420              140   101.7   4.129           1.3764
## 2      600              200    99.0   6.061           2.0202
## 3      300              100   115.2   2.604           0.8681
## 4      360              120   105.0   3.429           1.1429
## 5      420              140   111.9   3.755           1.2516
```

Based on that, a 2-D matrix that captures the space-time dynamics of each of these variables can be obtained. The space-time resolution of the 2-D matrices is based on

[1] A compilation of R (R Core Team, 2015) scripts necessary for data extraction and pre-processing in the form of R package is made available on `https://github.com/arwasayegh/mdep`. "mdep" package refers to "MIDAS Data Extraction and Pre-processing". The package does not provide any traffic data set. It is thus only useful for users who have access to the MIDAS data base.

the number of measurement sites and measurement time steps available for a selected road network and time period. Hence, the transformed data sets are subsetted based on space (the origin and destination measurement sites) and time (the start and end of a time period) to obtain a 2-D matrix of each of the traffic variables. An example extract of the 2-D matrices is provided here. The example is based on six hours traffic data extracted for a 3.1 kilometres stretch on the M60 motorway consisting of eight measurement sites. This example only shows the average traffic flow per lane and density per lane matrices as these are similar to the average traffic flow and density matrices, respectively.

```
## example of 2-D matrix of average traffic flow per lane - top 5 rows
```

```
##              datetime M60/9140A M60/9145A M60/9148A M60/9151A M60/9157A M60/9163A M60/9167A M60/9172A
## 1 2013-04-16 14:00:00       920      1040       980       840      1240      1060      1240      1540
## 2 2013-04-16 14:01:00      1020       960       600       740       900      1060      1040      1540
## 3 2013-04-16 14:02:00      1180      1120       840       840      1000       740       800      1320
## 4 2013-04-16 14:03:00      1360      1320       880       780      1120      1040       940      1280
## 5 2013-04-16 14:04:00      1180      1180       840       960      1240      1040       900      1500
```

```
## example of 2-D matrix of average traffic speed - top 5 rows
```

```
##              datetime M60/9140A M60/9145A M60/9148A M60/9151A M60/9157A M60/9163A M60/9167A M60/9172A
## 1 2013-04-16 14:00:00    105.57    103.73    106.61     108.5     99.84    101.57     99.39     87.69
## 2 2013-04-16 14:01:00    102.35    102.44    104.90     107.6    104.64     99.58     95.56     81.25
## 3 2013-04-16 14:02:00    102.10    102.11    105.48     109.4    103.32    102.78     98.85     84.21
## 4 2013-04-16 14:03:00     96.26     95.86    102.55     107.0    102.39    101.17    101.72     91.20
## 5 2013-04-16 14:04:00     99.44    100.71     99.76     100.9     95.13     98.31     99.04     87.92
```

```
## example of 2-D matrix of average density per lane - top 5 rows
```

```
##              datetime M60/9140A M60/9145A M60/9148A M60/9151A M60/9157A M60/9163A M60/9167A M60/9172A
## 1 2013-04-16 14:00:00     8.715    10.026     9.192     7.745    12.420     10.44    12.476     17.56
## 2 2013-04-16 14:01:00     9.966     9.372     5.720     6.879     8.601     10.64    10.883     18.95
## 3 2013-04-16 14:02:00    11.557    10.969     7.964     7.676     9.679      7.20     8.093     15.67
## 4 2013-04-16 14:03:00    14.128    13.770     8.582     7.290    10.938     10.28     9.241     14.03
## 5 2013-04-16 14:04:00    11.866    11.717     8.420     9.517    13.035     10.58     9.087     17.06
```

While the average traffic flow and speed at an on-ramp are directly used in the traffic flow models, calculated average traffic flow at an off-ramp needs to be transformed to off-ramp split ratios. These are typically calculated as the ratio of measured off-ramp average traffic flow to the total measured average traffic flow. The latter is defined as the mainline average traffic flow on the first measurement site downstream an off-ramp, plus the off-ramp average traffic flow (Muñoz et al., 2004).

The last component of the traffic data is speed limits. While these are not part of the MIDAS data base, they are separately requested from Highways England on road networks where VSL is under operation. Speed limit data is linked with the MIDAS network data base by the unique identifier of any measurement site defined earlier. Speed limit data is provided for each lane of a measurement site in miles per hour and is reported whenever speed limits change, rather than at specific time intervals. An indicator of whether speed limits are enforced (vehicles need to comply) or advisory (vehicles do not need to comply) is also provided for each posted speed limit. It is assumed that the speed limit on a measurement site is the arithmetic average of speed

limits across all lanes. This is considered the major data transformation of speed limit data. In order to match the one minute resolution of traffic data, the speed limit at any particular minute remains constant until it changes. Also, in order to match the units of other traffic data, speed limits are converted to [km h^{-1}].

4.2.2 Fleet mix data

The second type of data required for predicting average-speed based emissions on a particular road network is the fleet mix. COPERT emission factors (selected in this study) have also been associated with UK fleet composition data developed by the NAEI. Thus the proportion of vehicles by class, fuel type, Euro standard (and engine size for certain vehicle classes) are provided at the same level of aggregation of the developed COPERT emission factors of a particular air pollutant. Fleet mix data since 2011, which are of interest in this study as will be explained next, are based on traffic projections published by the Department for Transport (DfT) in 2013 (Department for Transport, 2013; Department for Environment Food and Rural Affairs, 2016). Road network specific assumptions and data summary are provided in § 4.4.2.

4.3 Road networks, time periods, and models

Having provided a background on the sources of traffic data and the generic data handling approach, this section provides a description of selected road networks and time periods. Depending on the operation of VSL, the traffic flow models for each road network are then specified and their space-time configurations are defined.

Three motorway road networks have been selected, i.e. $R = 3$. Each road network has different characteristics. These are a 3.1 kilometres road network on the non-VSL M60 motorway, a 3.8 kilometres road network on the non-VSL M1 motorway, and a 4.8 kilometres road network on the VSL M25 motorway. Each is equipped with a number of measurement sites as part of the MIDAS network. Figure 4.3 below shows their location along with the geographical locations of measurement sites on each one of them. These are based on the actual geographical locations of measurement sites on the pavement.

The first M60 road network is located on the western side of the Manchester Ring Motorway between Junction 8 (J8) and J11. Encircling Greater Manchester and with two on-ramps and two off-ramps located along the mainline route, this section is characterised by its recurrent congestion specifically during the AM and PM peak periods with a two-way annual average daily traffic on J11 reported to reach 138,000 vehicles per day during 2013 (Department for Transport, 2016b). Eight measurement sites are

FIGURE 4.3: Map[2]location of the three selected road networks: [A] non-VSL M60, [B] non-VSL M1, and [C] VSL M25 represented by the dark blue coloured lines. Light blue coloured circles indicate the location of measurement sites available along the motorway mainline, excluding on-ramps and off-ramps measurement sites.

located on the M60 mainline route, and additional four measurement sites located on each on-ramp and off-ramp.

The second road network is located on the northern end of the M1 motorway between J31 and J33 near Sheffield City. The M1 motorway is considered to be a major connection between London and northern UK. It is thus also characterised by recurrent congestion during the AM and PM peak periods with a two-way annual average daily traffic on J31, which is reported to reach 126,000 vehicles per day during 2013 (Department for Transport, 2016b). Eight measurement sites are located on the M1 mainline route with additional measurement site on an on-ramp, the only ramp located along the route.

The third road network is located on the western side of the M25 London Orbital Motorway between J14 and J12. Encircling greater London, the M25 motorway is one of the busiest motorways in the UK with a two-way annual average daily traffic on J14, which is reported to reach 228,000 vehicles per day during 2013 (Department for Transport, 2016b). Unlike the first two, this is considered a controlled motorway section and thus operates VSL. Speed limits on this M25 mainline route are most often enforced rather than advisory. They can vary by measurement site, motorway lane, and time period typically between 40 miles per hour, i.e. 65 [km h^{-1}], and 60 miles per hour, i.e. 96 [km h^{-1}]. In the absence of VSL signs, the motorway speed limit is assumed to be the NSL, which is 70 miles per hour, i.e. 113 [km h^{-1}]. 11 measurement sites are located on this mainline route with additional two measurement sites located on an off-ramp and on-ramp. Between the off-ramp and on-ramp, the number of lanes drops from five to four lanes, unlike the first two which are three lanes sections throughout. Table 4.1 below provides a summary of characteristics of each selected road network.

[2]Maps were plotted using R (R Core Team, 2015) and RgoogleMaps package (Loecher & Ropkins, 2015).

TABLE 4.1: summary of selected non-VSL M60, non-VSL M1, and VSL M25 road network characteristics

Road Network	**M60**	**M1**	**M25**
Location	Manchester Orbital	Sheffield City	London Orbital
Direction	Northbound	Northbound	Anti-clockwise
Length in [km]	3.1	3.8	4.8
Start point OS[1]grid reference	[377526, 394773]	[448017, 388049]	[503129, 174537]
End point OS grid reference	[375778, 396773]	[444800, 389325]	[501650, 170116]
VSL operation	No	No	Yes
Number of lanes	3	3	5 and 4
Number of on-ramps	2	1	1
Number of off-ramps	2	-	1
Location of on-ramps [km]	1.34, 2.69	1.97	2.71
Location of off-ramps [km]	0.66, 2.11	-	1.55
Number of measurement sites[2]	8	8	11

[1] Ordnance Survey.
[2] The number includes 1 mainline origin and 1 mainline destination dual loop detector measurement sites but excludes on-ramp(s) and off-ramp(s) measurement sites. There exist a measurement site on each on-ramp and off-ramp of each road network.

For each of the selected road networks, 6 time periods, i.e. $M_r = 6$ $\{r = 1, 2, 3\}$, from the MIDAS data base have been selected. Time periods of each data set range between three and six hours and typically includes the AM peak period or the PM peak period of weekdays only. Notably, congested traffic conditions have been observed to occur more often during the PM time periods for the M60 road network, but during the AM time periods for the M1 road network. In 2014, VSL was on operation on the M25 road network. For each selected time period, accident data released by the Department for Transport (2016a) has been cross checked in order to ensure that major accidents have not occurred. Selected time periods are summarised in Table 4.2 below for each road network. From now on, time periods will be referred to by their numbers [1-6] as indicated in the table.

TABLE 4.2: summary of time periods for each of the selected non-VSL M60, non-VSL M1, and VSL M25 road networks

Road Network	**M60**	**M1**	**M25**[1]
Time resolution	1 minute	1 minute	1 minute
Number of time periods (M_r)	6	6	6
Length of time periods [h]	[6, 6, 6, 6, 3, 3]	[3, 3, 3, 3, 5, 4]	[3, 3, 3, 3, 5, 5]
Number of AM time periods	0	6	3
Number of PM time periods	6	0	3
Date of time periods [dd-mm-yyyy]	16-04-2013 [1]	01-07-2013 [1]	06-05-2014* [1]
	17-04-2013 [2]	02-07-2013 [2]	06-05-2014 [2]
	18-04-2013 [3]	03-06-2013 [3]	12-05-2014* [3]
	22-04-2013 [4]	04-06-2013 [4]	12-05-2014 [4]
	23-04-2013 [5]	25-06-2013 [5]	20-05-2014* [5]
	24-04-2013 [6]	26-06-2013 [6]	20-05-2014 [6]

[1] Asterisked time periods [1, 3, 5] are the AM time periods.

In selecting the road networks and time periods and identifying their characteristics, particularly their VSL operation, traffic flow models used for each road network is selected and the cell (or space) configuration of each road network can be specified. Both CTM and M-METANET are used for each road network in order to assess and compare their performance in each phase. CTM-vsl and M-METANET-vsl are used for calibration and validation using the time periods of the M25 road network only. However, CTM-basic and M-METANET-basic are also used for the same road network in order to compare the performance of both models on a VSL operated road network, and to check whether the VSL model structure improves upon the basic CTM and M-METANET model structures or not and provide possible justifications. Based on that, selected models for each road network are:

- CTM-basic and M-METANET-basic for the M60 road network, hereafter referred to as M60/CTM-basic and M60/M-METANET-basic traffic flow modelling scenarios, respectively;
- CTM-basic and M-METANET-basic for the M1 road network, hereafter referred to as M1/CTM-basic and M1/M-METANET-basic traffic flow modelling scenarios, respectively;
- CTM-vsl and M-METANET-vsl for the M25 road network, hereafter referred to as M25/CTM-vsl and M25/M-METANET-vsl traffic flow modelling scenarios, respectively; and
- CTM-basic and M-METANET-basic also for the M25 road network, hereafter referred to as M25/CTM-basic and M25/M-METANET-basic traffic flow modelling scenarios, respectively.

For all three road networks, general guidelines have been followed for the cell configuration. These are:

- the location of an on-ramp or off-ramp always indicates the start of a new cell;
- the identified measurement sites are always located at the end of cells;
- cells are not necessarily of equal lengths; and
- cells might not necessarily have a measurement site since the distance between these can be more than 600 metres. These cells are referred to here as non-equipped cells.

Based on these guidelines, cells for the three road networks have been configured as follows and are summarised in Table 4.3 below:

- The M60 road network has been divided into 15 cells (excluding the mainline origin and destination cells) with lengths varying between 0.1 kilometres to 0.285 kilometres; nine of which are non-equipped cells.
- The M1 road network has been divided into 13 cells (excluding the mainline origin and destination cells) with lengths varying between 0.22 kilometres to 0.33 kilometres; seven of which are non-equipped cells.

- The M25 road network has been divided into 18 cells (excluding the mainline origin and destination cells) with lengths varying between 0.14 kilometres to 0.38 kilometres; nine of which are non-equipped cells.

TABLE 4.3: spatiotemporal configuration of selected macroscopic traffic flow models for each of the selected non-VSL M60, non-VSL M1, and VSL M25 road networks

Road Network	**M60**	**M1**	**M25**
Selected models (f)	CTM-basic; M-METANET-basic	CTM-basic; M-METANET-basic	CTM-vsl; M-METANET-vsl; CTM-basic; M-METANET-basic
Number of cells (N)[1]	15	13	18
Number of non-equipped cells	9	7	9
Number of equipped cells[2]	6	6	9
Cell lengths [min,max] in [km]	[0.1, 0.285]	[0.22, 0.33]	[0.14, 0.38]
Model time step in [seconds]	2	3	3
Model time step (T_s) in [h][3]	2/3600	3/3600	3/3600
u_f CFL constraint [km h^{-1}]	≤ 180	≤ 264	≤ 168

[1] The number excludes the 1 mainline origin cell 0 and 1 mainline destination cell $N+1$.
[2] The number excludes the 1 mainline origin and 1 mainline destination measurement sites. Origin cell 0 and destination cell $N+1$ are equipped and their measurements are used as input data.
[3] Unit of T_s as input to the models is in hours rather than seconds.

Data sets are then extracted and pre-processed based on the description of § 4.2. The resulting 2-D matrices of traffic variables provide the input data and the parameterisation data needed for any selected traffic flow model. However, it is clear that the spatial resolution of measurement data (i.e. the number of measurement sites along a mainline route) does not match the space (cell) configuration of each road network (i.e. the total number of cells). Also, it is not feasible to simulate the traffic flow models at a time resolution as low as one minute, which is the time resolution reported by measurement sites. Typically, the choice of the time step at which the traffic flow models (investigated in this research) depends on the purpose of the study and the computational feasibility, since higher time resolution leads to higher computational time. Also, the selection of time step depends on the minimum cell length on a road network. Shorter cell lengths require the time step to be reduced as well in order to ensure that the constraint on the free-flow speed parameter (set to satisfy CFL condition) is plausible. Most often, a time step less than or equal to ten seconds is used. This study uses a time step of three seconds for all road networks except for the M60, in which a time step of two seconds is used. The use of two seconds time step for the M60 is a result of having a minimum cell length of 0.1 kilometres, the shortest amongst all other road networks. In order for the spatiotemporal resolution of traffic data to match the resolution of input data demanded by the traffic flow models, constant interpolation of the data has been applied where necessary, spatially and temporally (similar assumption has been made in Zhong et al. (2015), as an example).

4.4 Data summary

Having pre-processed the traffic data sets corresponding to each road network and time period, a summary of traffic data is provided in this section. There is a particular focus on those required for the traffic flow models, as this aids the analysis of the methodological framework results. Also, given the location of road networks and the years of selected traffic data sets, fleet mix data is extracted accordingly and summarised here.

4.4.1 Traffic data

Given a total number of 18 data sets for all road networks, only an example of input data extracted and pre-processed on a particular road network and time period is provided here. Figure 4.4 and Figure 4.5 below show time series data of mainline boundary conditions and ramps boundary conditions required for all CTM and METANET modelling scenarios. The input data corresponds to the M60 road network and traffic conditions of time period [1] (i.e. the PM period of the 16th of April 2013). Figure 4.6, however, corresponds to the additional mainline origin and off-ramps boundary conditions required for the METANET ones only. The only additional transformation on these input data is the constant interpolation of each time series to match the spatiotemporal resolution of the traffic flow models.

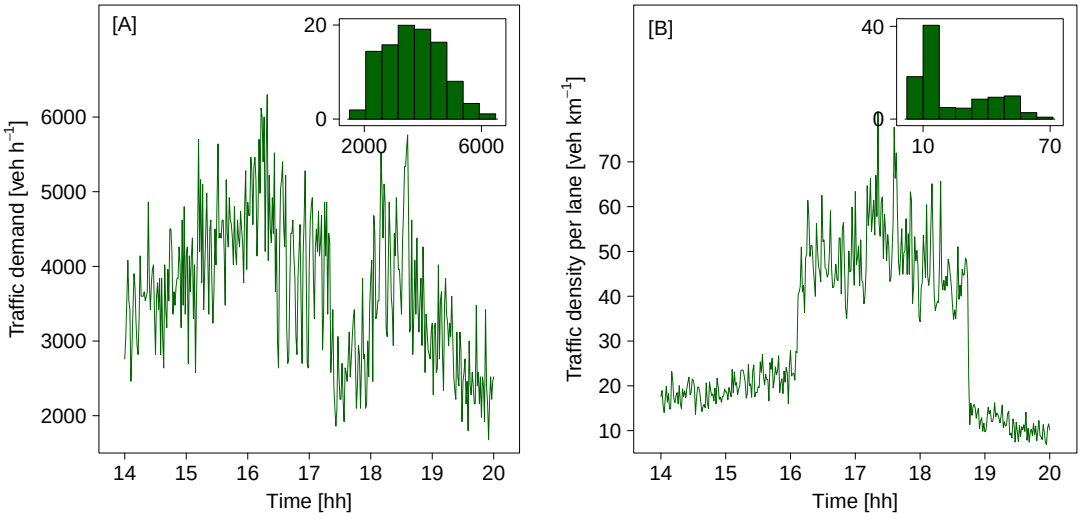


FIGURE 4.4: Summary of input data for time period [1] (14:00 till 20:00 on 16th April 2013) of the selected M60 road network: [A] time series of traffic demand [veh h^{-1}] at the mainline origin overlaid by the time series histogram; and [B] time series of traffic density per lane [veh km^{-1}] at the mainline destination overlaid by the time series histogram.

The model calibration and validation comprises:

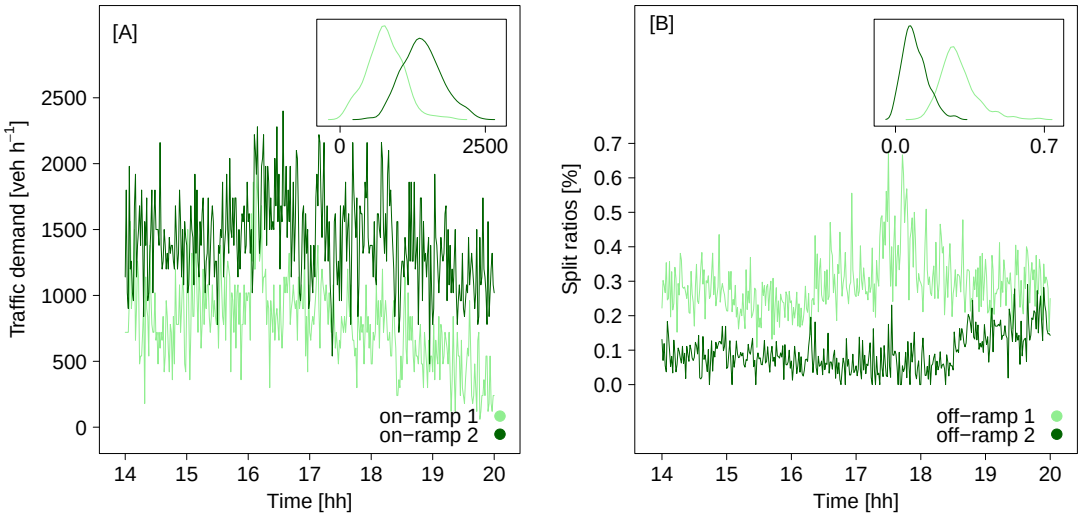


FIGURE 4.5: Summary of input data for time period [1] (14:00 till 20:00 on 16th April 2013) of the selected M60 road network: [A] time series of traffic demand [veh h^{-1}] at the two on-ramps overlaid by the time series (Gaussian) kernel density; and [B] time series of off-ramp split ratios [%] overlaid by the time series (Gaussian) kernel density.

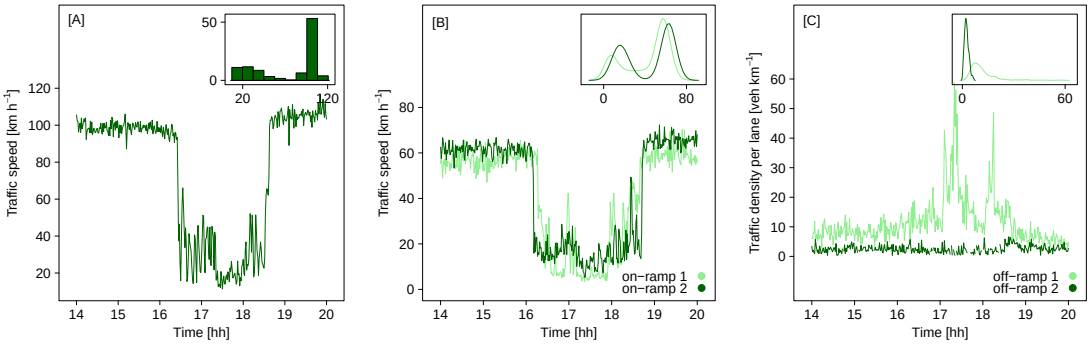


FIGURE 4.6: Time series of [A] traffic speed [km h^{-1}] at the mainline origin overlaid by the time series histogram, [B] traffic speed [km h^{-1}] at two on-ramps ovarlaid by the time series (Gaussian) kernel density, and [C] traffic density per lane [veh km^{-1}] at two off-ramps ovarlaid by the time series (Gaussian) kernel density for time period [1] (14:00 till 20:00 on 16th April 2013) of the selected M60 road network. These are considered additional input data for METANET traffic flow models only.

- data required for the ensemble-based optimisation of parameters which are mainly the 2-D matrices of average traffic speed with the origin, destination, and initial conditions data removed as these are used as input;
- data required for the selection of *top-solutions* which are mainly the 2-D matrices of average traffic density per lane and speed with the origin, destination, and initial conditions data removed as these are used as input; and
- data required for the estimation of calibration and validation prediction errors which are mainly the 2-D matrices of average traffic density per lane and speed with the origin, destination, and initial conditions data removed as these are used

as input.

Figure 4.7 and Figure 4.8 below show the distribution of these data for each of the 6 selected time periods on each of the three road networks. The average traffic speed and density per lane distributions of each data set across all networks show two peaks, as expected, representing the free-flow traffic conditions (i.e. high traffic speed/low density per lane) and congested traffic conditions (i.e. low traffic speed/high density per lane). Data for both free-flow and congested traffic conditions are essential in order to determine the model parameters relating to both flow regimes. Amongst all road networks, a minimum average traffic speed and maximum traffic density per lane have been observed on the M60 with values of 3.8 [km h^{-1}] and 95.6 [veh km^{-1}], respectively, while a maximum average traffic speed and minimum traffic density per lane have been observed on the M1 with values of 140.5 [km h^{-1}] and 0.28 [veh km^{-1}], respectively.

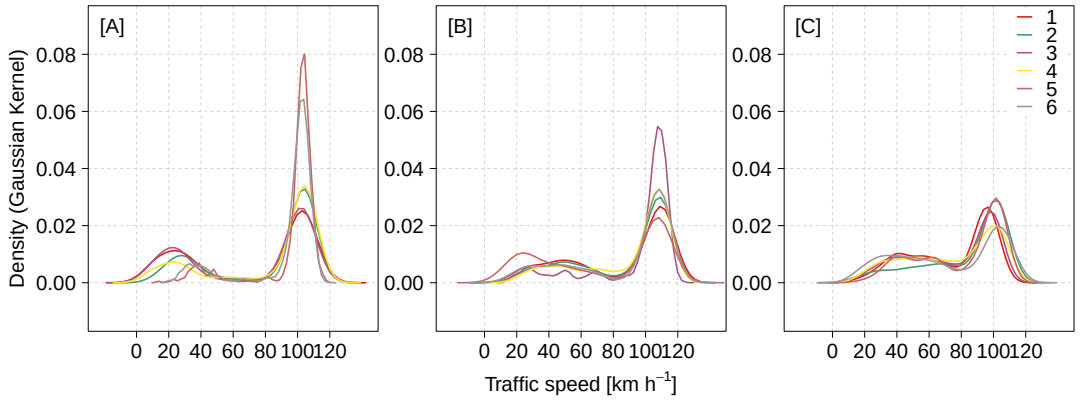


FIGURE 4.7: Kernel density plots (with the kernel assumed to be Gaussian) for the average traffic speed data [km h^{-1}] at each of the three road networks: [A] M60, [B] M1, and [C] M25. Each plot includes 6 curves representing the kernel density of the 6 time periods selected (with data used as model input removed) for each route (with number of data points: 2160, 2160, 2160, 2160, 1080, and 1080 for the [A] M60 data sets; 1080, 1080, 1080, 1080, 1800, and 1400 for the [B] M1 data sets; and 1620, 1620, 1620, 1620, 2700, and 2700 for the [C] M25 data sets.

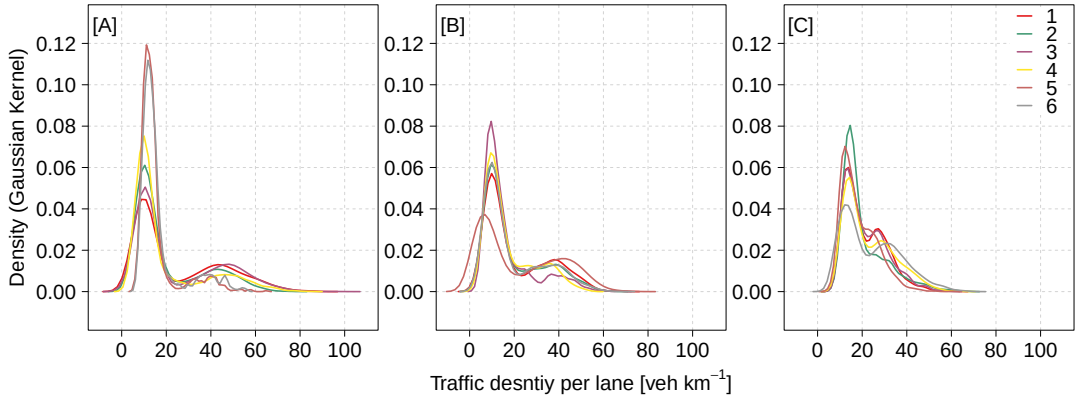


FIGURE 4.8: Kernel density plots for the average traffic density per lane data [veh km^{-1}] at each of the three road networks: [A] M60, [B] M1, and [C] M25.

4.4.2 Fleet mix data

The UK fleet mix data described in § 4.2.2 provides the proportion of vehicle categories depending on the location of the road network, particularly if it is in or out of London. It also distinguishes between the different road types, as follows:

- urban, rural, and motorway if the road network is located outside London; and
- central, inner, outer, and motorway if the road network is located inside London.

Accordingly, the proportions of vehicles for the M60 and M1 are considered to be different than those of the M25 which is located inside London. Obviously, the proportion of vehicles for motorways is considered for all road networks. Also, given that the traffic data sets are extracted from year 2013 for the M60 and M1 and from year 2014 for the M25, the proportion of vehicles for these years is extracted accordingly for each road network.

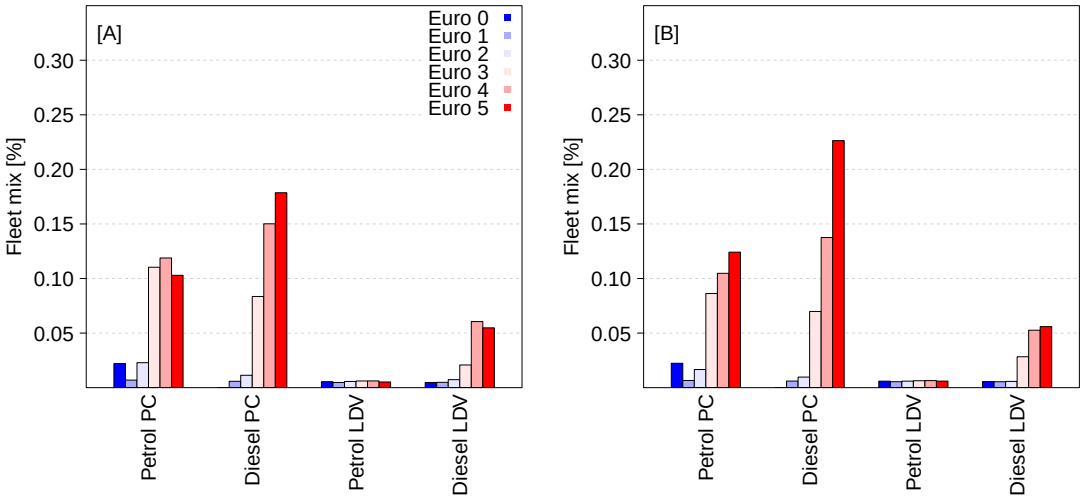


FIGURE 4.9: Fleet mix proportions used for [A] the M60 and M1 road networks during 2013 and [B] the M25 road network during 2014 (Department for Transport, 2013; Department for Environment Food and Rural Affairs, 2016). Proportions are split based on vehicle type (PCs and LDVs), fuel type (Petrol and Diesel), and Euro standard (Euro 0 to Euro 5)

Figure 4.9 above shows the proportion of PCs and LDVs per fuel type (petrol or diesel) and per Euro standard (Euro 0 to Euro 5 only as Euro 6 does not appear until 2015) used for each of the M60 and M1 (Figure 4.9A) and for the M25 (Figure 4.9B). The proportion of heavy good vehicles and coaches are assumed to be distributed equally amongst the proportions of PCs and LDVs. Figure 4.9 distinguishes fleet mix data based on the corresponding COPERT emission factors; since this the main use of fleet mix data. For instance, it does not attempt to distinguish the proportion of vehicles by engine size since the emission factor (the function and its parameters) for a particular vehicle class and fuel type is the same regardless of the engine size. This is always the case except for the Euro 0 PCs where its COPERT emission factor depends

on the engine size (i.e. less than 1.4 Litres, between 1.4 and 2.0 Litres, and higher than 2.0 Litres). Also, COPERT emission factor for London taxis is considered to be the same as the Diesel LDVs. Hence, proportions of London taxis are added to those of Diesel LDVs, depending on their Euro standard. Lastly, COPERT emission factor for vehicles with failed catalyst is considered to be the same as Euro 0 vehicles. This is reflected in the Euro 0 fleet mix proportion of Figure 4.9. Based on that, Figure 4.9 shows the following:

- on the M25 road network, 81% of vehicles are PCs, 36% of which are petrol vehicles and 45% are diesel vehicles. The remaining 19% are LDVs, 3.6% of which are petrol vehicles and 15.4% are diesel vehicles; and
- on the M60 and M1 road networks, 81.4% of vehicles are PCs, 38.4% of which are petrol vehicles and 43% are diesel vehicles. The remaining 18.6% are LDVs, 3.3% of which are petrol vehicles and 15.3% are diesel vehicles.

4.5 A note on measurement sites' locations

It was highlighted in § 4.2 that the geographical locations of measurement sites obtained from the NTIS (referred to here as 'reported' locations) can be different than the location of measurement sites on pavement (referred to here as 'actual' locations). In addition to that, distances between measurement sites are not necessarily 500 metres, as is typically assumed for the MIDAS network. In fact, standards on spacing of MIDAS detector loops specified in the Design Manual for Roads and Bridges (DMRB) (Highways England, 2005) highlights that distances between measurement sites can be 500 metres (±20%). While the 'actual' locations and the 'actual' distances have been used in this study, this section highlights the extent of the discrepancies that can occur in the case where the 'reported' locations and 500 metres distances are used instead. For example, if the route under study is long enough to prevent the modeller from extracting the 'actual' locations, it can - most often - be time-consuming to double check geographical coordinates with street level locations unless the process is automated.

The 'actual' locations of all measurement sites used in this study (a total of 27 locations) have been compared with the 'reported' locations. The absolute difference in distance between the 'reported' and 'actual' locations on the three road networks is shown in Figure 4.10A. The absolute difference ranges between 0 to 170 metres with a mean (μ) of 47 metres. Figure 4.10A shows that differences in the data from the M25 are higher in comparison to those from the M60 and M1. In order to calculate the confidence interval around the mean, non-parametric bootstrapping is undertaken to avoid the normality assumption imposed by the margin-of-error method. The confidence interval is estimated by taking a random sample with replacement from the data to generate 2000 bootstrap replicates. The mean of each bootstrap replicate (μ^*) is

calculated and the approximate 95% confidence interval is calculated. The distribution of μ^* is shown in Figure 4.10B and the approximate 95% confidence interval of the mean value of difference between 'actual' and 'reported' locations is [29, 64] metres. Figure 4.10C shows how the calculated confidence interval changes by changing the number of bootstrap replicates (from 100 to 10000).

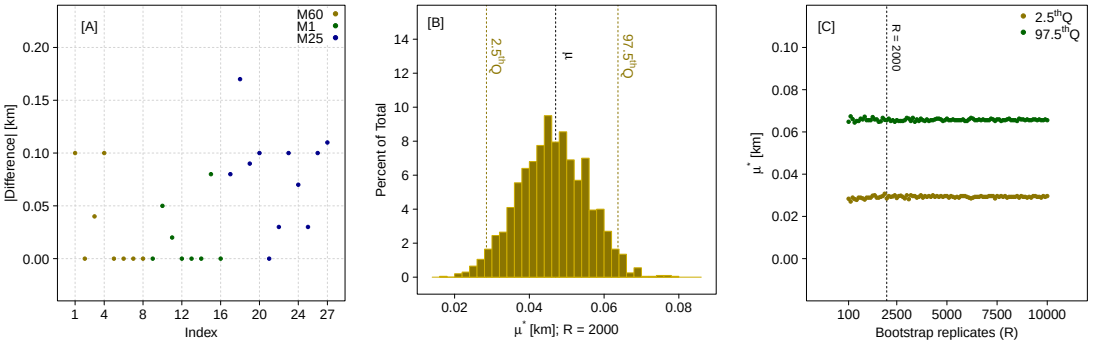


FIGURE 4.10: [A] Absolute difference in kilometres between 'reported' and 'actual' locations of measurement sites; [B] distribution of μ^* in kilometres, the mean values of the non-parametric bootstrap replicates (for R = 2000) with the dotted black line showing the mean, μ, based on the measurements and dotted golden lines showing the 2.5^{th} and 97.5^{th} quantiles of the mean values of bootstrap replicates; and [C] the variation of 95% confidence interval with varying bootstrap replicates, R, between 100 and 10000.

On the other hand, it has also been observed that the mean distance between measurement sites is 480 metres with minimum distances reaching 210 metres and maximum distance reaching 930 metres. If a 500 metres distance between measurement sites is assumed, then the absolute differences from the 'actual' distances have been observed to range between 0 to 430 metres, as shown in Figure 4.11A, with a mean (μ) absolute difference of 102 metres. Similar bootstrapping exercise has been undertaken to calculate the confidence interval of the mean absolute difference. Figure 4.11B shows the distribution of μ^* (means of the bootstrap replicates) while Figure 4.11C shows how the confidence interval changes by changing the number of bootstrap replicates (from 100 to 10000). Using 2000 bootstrap replicates, the approximate 95% confidence interval around the mean is [67, 145] metres.

The first limitation of such analysis is the sample size (number of points used to undertake bootstrapping). The second limitation is related to the errors that can arise from the fact that dual loop detectors in themselves are typically 6.5 metres, which makes it hard to calculate the exact differences using a single 'actual' coordinate and a single 'reported' coordinate. The last limitation is the possibility that discrepancies at certain motorways can be higher than other motorways. Despite that, this brief section attempts to highlight the issue of measurement sites' locations since potential discrepancies can lead to:

- measured data being compared with modelled data of adjacent cells, rather than cells on which the measurement site is actually located. This becomes more

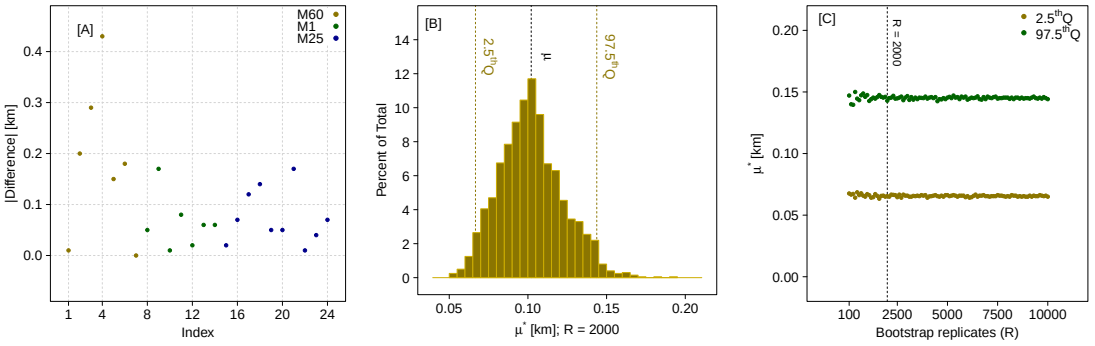


FIGURE 4.11: [A] Absolute difference in kilometres between the typical assumption of 500 metres and 'actual' distances of two consecutive measurement sites; [B] distribution of μ^* in kilometres, the mean values of the non-parametric bootstrap replicates (for R = 2000) with the dotted black line showing the mean, μ, based on the measurements and dotted golden lines showing the 2.5^{th} and 97.5^{th} quantile of the mean values of bootstrap replicates; and [C] the variation of 95% confidence interval with varying bootstrap replicates, R, between 100 and 10000.

important when measurement sites are located near on-ramps and off-ramps and hence ensuring that the correct location of the measurement site (before or after a ramp) is essential. This of course can be double checked by looking at the measurements, particularly at the average traffic flow which typically changes before and after a ramp.

- Changes to the road network layout configuration since the location of measurement sites (in this study) are set to be at the end of each cell. Cell lengths can easily be altered if the location of a measurement site is not accurate enough.

4.6 Summary

This chapter is considered the results of the preliminary phase of the methodological framework. It is also considered the main input to the results of subsequent phases of Chapter 5. The chapter has provided a detailed description of road networks ($R = 3$) and time periods ($M_r = 6$) selected for this study. It has also given a background on the sources of data and how it is handled, including all the assumptions being made with reference to scripts which can be used if access to the MIDAS network data base is acquired. While particular road networks and time periods were selected and described in earlier sections, such generic background allows the transferability of this phase to other road networks and time periods of the MIDAS network data base.

Key assumptions made in this chapter are the estimation of space-mean traffic speed as the weighted arithmetic mean; the existence of non-equipped cells on each road network (i.e. measurement sites are not available on all cells); the constant spatiotemporal interpolation of traffic data to match the input and parameter estimation requirements

of traffic flow models; and the assumption of heavy duty vehicles' proportions being considered either PCs or LDVs as a result of uni-class traffic flow models.

The selection and description of road networks and time periods provided the necessary background to select the macroscopic traffic flow models to be used in consequent phases; particularly whether the VSL formulation of the CTM and METANET traffic flow models is to be used or not. A total of eight traffic flow modelling scenarios have been set. Although three road networks are selected, both the non-VSL and VSL formulation of CTM and METANET are applied to the VSL-operated road network for comparison purposes. This gave rise to four CTM-related and four M-METANET-related modelling scenarios. The extracted and pre-processed data sets have also been analysed and compared in order to help better understand the results of subsequent phases.

The last section of this chapter highlighted the extent of the issue regarding the geographical locations of measurement sites. The analysis has shown that the 95% confidence interval of mean absolute difference in the 'reported' versus 'actual' locations is [29, 64] metres, while that of the mean absolute difference in the typical assumption of 500 metres and the 'actual' distance of two consecutive measurement sites is [67, 145] metres.

Chapter 5

Results and Discussion

5.1 Introduction

Having prepared all the necessary data, Chapter 5 presents and discusses the results of the four main phases, described in § 3.4. The following four main sections focus entirely on the results of each phase. Each of these sections are divided into two:

- The first presents and discusses results related to the four CTM modelling scenarios; and
- The second presents and discusses the results related to the four M-METANET modelling scenarios, with a comparison to those obtained from the CTM where appropriate.

A detailed description of each phase is only provided when presenting CTM-related results and unnecessary repetition is avoided when presenting those of the M-METANET. Also, in some cases where additional analysis is provided to illustrate a point, the case study of M60 road network is considered, while those of other road networks are provided in an appendix, when necessary.

Both Phases I and Phase III involve making phase-input assumptions. These are mainly the three ensemble-based input assumptions and the grid-number input assumption. As a result, two sensitivity analysis exercises are performed alongside the results of original assumptions. The concept behind these is not to find the 'optimal' assumption but rather to generally comment on the possible influences of varying assumptions on the ensemble-based optimisation results (given phase I assumptions) and on the confidence intervals of emission predictions (given phase III assumptions).

Having studied the application of the methodological framework on real road networks, § 5.6 ends with a final example of how one can use the final surface maps obtained for each modelling scenario in Phase IV to predict emissions alongside their confidence intervals given a choice of a time period and *top-solution*.

5.2 Phase I: Ensemble-based optimisation

Both the optimisation algorithm and optimisation objective function have been set earlier. The first step in Phase I is to determine an initial distribution of parameter vectors from which values can be randomly drawn to initialise the optimisation algorithm. This is followed by determining the other three inputs of Phase I: the number of optimisation runs E; the threshold percentile for objective function values χ; and the significance level for the two selected statistical tests ξ.

Since parameter vectors are model-specific, the distributions of their initial values are described for the CTM and METANET, separately. In addition to that, the choice of optimisation runs' number depends on both the computational power one can afford but at the same time on the ability to explore all possible *top-solutions*. The choice of threshold percentile and significance levels is a trade-off between not rejecting solutions that should be rejected, and rejecting solutions that should not be rejected. These inputs are set the same for all modelling scenarios as follows: $E = 500$, $\chi = 50\%$, and $\xi = 0.01$ (1%). The influence of varying these inputs on the results are discussed in § 5.2.3.

The following sections discuss the results of Phase I obtained for each of the CTM and M-METANET modelling scenarios. Because prediction errors are to be estimated for calibration outputs and validation outputs in Phase II, initial results of calibration performance and validation performance are presented here separately for the CTM in § 5.2.1.1 and § 5.2.1.2, and METANET in § 5.2.2.1 and § 5.2.2.2.

5.2.1 First-order CTM

5.2.1.1 Model calibration

The parameters of the CTM are those of the fundamental diagram. The distributions of their initial values have been deduced using an initial optimisation process on three-months of normal weekdays (a total of 55 days) traffic data on the M25 road network described earlier. Similar method has also been proposed in (Zhong et al., 2015). Here, an SQP optimisation algorithm (Ghalanos & Theussl, 2015; Ye, 1987) is used, as proposed by Frejo et al. (2012), to minimise the mean absolute errors of predicted and measured traffic flow per lane for each daily data set. The distribution of each of the four parameters based on the 55 days are shown in Figure 5.1. Using the mean and variance of the values of each parameter, an approximate normal distribution $\sim \mathcal{N}(\mu, \sigma^2)$ is used to draw the initial parameters of each optimisation run e of the CTM modelling scenarios. The normal approximation results along with the p-values

obtained from a commonly used test of normality, Shapiro-Wilk test (Royston, 1982a, 1982b), are:

- $u_f \sim \mathcal{N}(104, 1.8)$, p-value $= 0.002$;
- $\underline{Q}_{max} \sim \mathcal{N}(1558, 4623)$, p-value $= 0.01$;
- $w \sim \mathcal{N}(18, 27)$, p-value $= 1 \times 10^{-4}$ (p-value $= 0.1$ in case the outlier value of 42 km h^{-1} is disregarded); and
- $\underline{\rho}_{max} \sim \mathcal{N}(111, 234)$, p-value $= 0.06$.

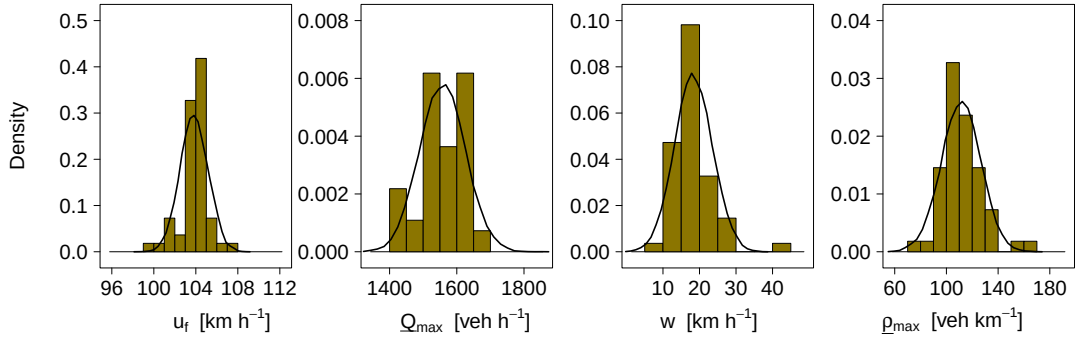


FIGURE 5.1: Histograms of the fitted CTM parameter vector on each of the 55 normal weekdays (from April 2014 to June 2014) traffic flow-density data on the M25 road network overlaid by their normal distribution approximations $\sim \mathcal{N}(\mu, \sigma^2)$.

Parameters are randomly drawn E times to initialise the ensemble of optimisation runs for each time period of each modelling scenario: M60/CTM-basic, M1/CTM-basic, M25/CTM-vsl, and M25/CTM-basic. From now on, the results of the four CTM modelling scenarios follow this order. Figure 5.2 shows the objective function values of each ensemble and highlights the *top-solutions*, based on criteria set out earlier, in dark blue coloured circles. Figure 5.3 shows only those for the *top-solutions* in order to better understand the distribution of their objective function values. The range of values is dependent on the time period, road network, and modelling scenarios. Relatively higher values are obtained for the M25 optimisation runs which possibly reflects the complexity of modelling a VSL-operated road network.

Figure 5.2 and Figure 5.3 show that *top-solutions* are not identified for all time periods. Given the population of initial conditions tested in each case and the fact that N-M optimisation algorithm is a proven and reliable way to find a 'good' solution if one exists, this can be attributed to the CTM model structure itself being unable to reproduce all features of measured traffic data or at least features that allow to meet the statistical testing criteria. This can be a result of deficiencies such as the use of simplified fundamental diagram, use of steady-state speed-density relationship, and ignored influence of different vehicle classes, lane drops, and lane changing.

The number of *top-solutions* obtained should always be less than or equal to 250 (since the threshold percentile is 50%). Table 5.1 below shows the total number of *top-solutions* which have an objective function value less than the threshold and have

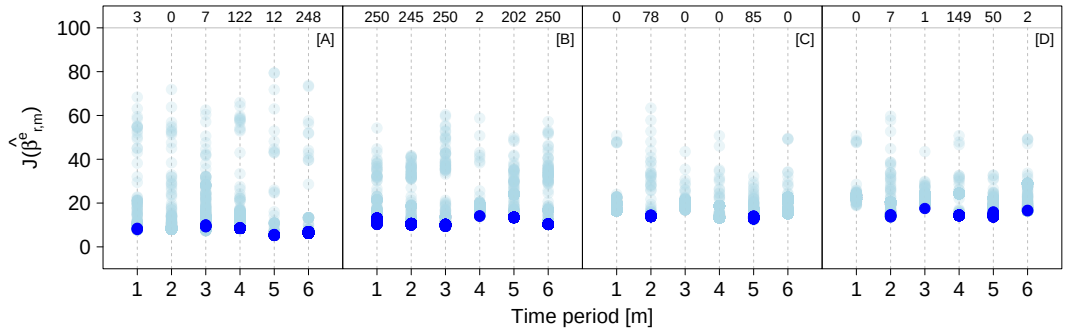


FIGURE 5.2: Minimised objective function values obtained from the calibration runs of [A] M60/CTM-basic, [B] M1/CTM-basic, [C] M25/CTM-vsl, and [D] M25/CTM-basic. From now on, the results in figures of the four CTM modelling scenarios follow this order. Dark blue coloured circles highlight the objective function values of *top-solutions* (the number of *top-solutions* for each time period is highlighted on top of each figure). One calibration run for [C] M25/CTM-vsl and one for [D] M25/CTM-basic did not converge based on the pre-defined convergence criteria and thus are not shown here.

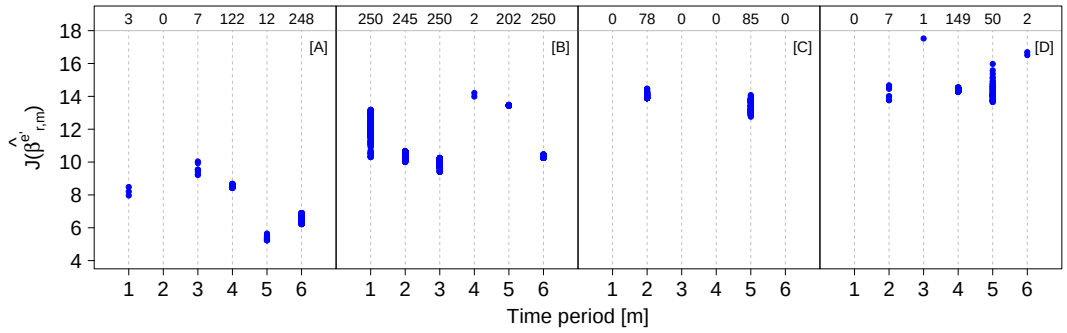


FIGURE 5.3: Minimised objective function values obtained from *top-solutions* only of each CTM modelling scenario.

satisfied the two statistical tests using either the predicted and measured traffic density per lane data (C1) or speed data (C2). It shows that both conditions are rarely satisfied at once for a single optimisation run. A total of 392 and 1199 are obtained for the M60/CTM-basic and M1/CTM-basic, respectively. This is relatively higher to those obtained for the two M25 modelling scenarios. While this can be seen a positive outcome (i.e. ability to find parameters which reproduce certain features of measured data), it can also imply that uncertainty around the parameters for these two modelling scenarios is larger in comparison. While uncertainty in the modelling structure dominates for the VSL-operated road network which makes it difficult to find a *top-solution*, uncertainty in the parameter vector dominates for non-VSL road networks where structural deficiencies are relatively less. For the M25, the CTM-basic resulted in 209 *top-solutions* compared to only 163 under the CTM-vsl; *top-solutions* are obtained for 5 time periods (out of 6) under the CTM-basic model structure but for 2 time periods only under CTM-vsl. This suggests that the CTM-vsl may sometimes be problematic and may not necessarily be appropriate to describe traffic behaviour for

all time periods especially when VSL operation changes by time, loop detector, and by lane as it does for this road network. CTM-vsl assumes a uniform impact of VSL on traffic behaviour, when this might not be necessarily true. For instance, compliance to different speed limits on different lanes and different sections during different time periods might not be necessarily the same. That said, the number of *top-solutions* for M25/CTM-basic drops to 11 (for 1 time period out of 6) when the significance level of statistical test is set to 0.1 instead of 0.01 (i.e. under more demanding criteria). Since this high drop does not happen for the M60 and M1 road networks, this indicates that CTM-basic may also be problematic for VSL-operated road networks. For the CTM-vsl, however, there exist two options to *possibly* overcome this issue: the first is adding a single VSL compliance parameter; and the second is adding a space-time varying compliance parameter. One issue with the first option is the possibility of obtaining both positive and negative compliance from different optimisation runs which can both be counter-intuitive (in terms of loosing its physical meaning) and result in over-fitting. The second option however leads to an over-parameterised model and to the problematic issue of fitting parameters based on very few data points.

TABLE 5.1: number of *top-solutions* of each CTM modelling scenario satisfying the threshold percentile and statistical tests criteria

Statistical Condition	**M60 CTM-basic**	**M1 CTM-basic**	**M25 CTM-vsl**	**M25 CTM-basic**
C1	382	1078	1	12
C2	10	202	163	199
C1&C2	0	81	1	2
Total	392	1199	163	209

Few *top-solutions* obtained based on the above-mentioned criteria involved values for the maximum traffic density parameter which is lower than traffic density measurements at the destination boundary of other time periods. This was observed during the validation process where they led to negative outputs. These *top-solutions* were removed from the analysis which were only 28 for the M60/CTM-basic and 6 for the M25/CTM-vsl. This issue can be critical when using an optimised parameter vector on real-time or future time periods where the destination boundary condition is not known beforehand.

Now that *top-solutions* are determined, their parameter vectors are examined in details. As a start, an example of how the values of the CTM parameter vector translate to the trapezoidal fundamental diagram is provided. Figure 5.4 shows the fundamental diagrams obtained based on the *top-solutions* of each time period of the M60 road network overlaid on the flow-density measurements of each time period. Time period [2] does not have any *top-solution* and thus is not shown. These fundamental diagrams do not reflect the flow-density prediction points corresponding to each measured point but rather only the expected relationship since predicted density from CTM is not the

same as those from measurements. The same plots are generated for the other three modelling scenarios and are shown in Appendix B.

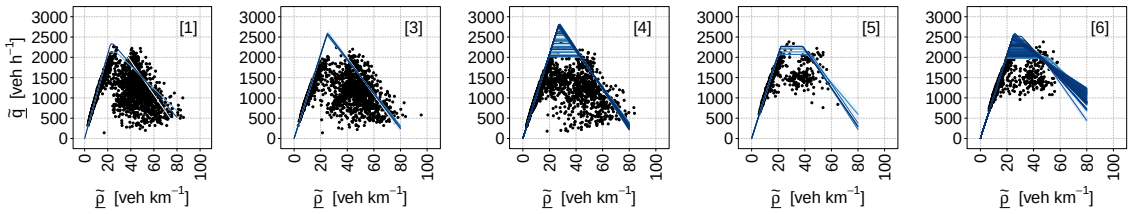


FIGURE 5.4: Measured traffic flow per lane [veh h^{-1}] and density per lane [veh km^{-1}] scatter plots (black dots) overlaid by fundamental diagrams (shades of blue lines) based on *top-solutions* of M60/CTM-basic: 3 for time period [1], 7 for time period [3], 122 for time period [4], 12 for time period [5], and 248 for time period [6]. Time period [2] does not have any *top-solution* and thus is not shown.

Table 5.2 shows the range of values of each parameters for each modelling scenario, irrespective of time period. Very similar ranges across different modelling scenarios are obtained for the free-flow speed parameter, u_f, in comparison to other parameters.

TABLE 5.2: [minimum, 99th percentile, maximum] of parameter vector values obtained from the *top-solutions* of each CTM modelling scenario

$\boldsymbol{\beta}$	**M60 CTM-basic**	**M1 CTM-basic**	**M25 CTM-vsl**	**M25 CTM-basic**
u_f	[102, 104, 105]	[104, 110, 111]	[103, 105, 105]	[100, 105, 106]
Q_{max}	[1981, 2832, 2853]	[1719, 2045, 2048]	[1635, 1823, 1893]	[1601, 1788, 1797]
w	[21, 50, 51]	[4, 26, 29]	[14, 40, 41]	[10, 37, 43]
ρ_{max}	[84, 137, 139]	[87, 293, 504]	[60, 138, 142]	[60, 174, 178]

To better understand and compare the variation across the different parameters, the *C*oefficient of *V*ariation (CV), (i.e. relative standard deviation) of each one obtained for each time period in each modelling scenario is calculated given the below generic equation:

$$CV = \frac{\sigma}{\mu} \times 100$$

where σ is the calculated standard deviation of the parameter values and μ is the calculated mean parameter value. Only time periods which have a minimum of 10 *top-solutions* were selected, and CV was calculated for each time period. Table 5.3 shows the minimum, mean, and maximum CV of each parameter obtained across these time periods. Lowest variability is observed in the demand parameters (i.e. the free-flow traffic speed and maximum flow). Highest variability is observed in the supply parameters (i.e. backward wave speed and maximum traffic density). This emphasises the complexity of driving behaviour during congested periods and the possibility of the supply part to take different values under the CTM assumption of a trapezoidal fundamental diagram. Of course, the high ranges and high variability in the resulting

parameters illustrate difficulties in finding a 'global' solution for the problem and justify the use of an ensemble-based optimisation.

TABLE 5.3: [minimum, mean, maximum] Coefficient of Variation (CV in %) of parameter vector values obtained across the time periods with more than 10 *top-solutions* for each CTM modelling scenario

$\boldsymbol{\beta}$	**M60 CTM-basic**	**M1 CTM-basic**	**M25 CTM-vsl**	**M25 CTM-basic**
u_f	[0.1, 0.1, 0.1]	[0.1, 0.3, 0.6]	[0.2, 0.3, 0.4]	[0.2, 0.3, 0.4]
Q_{max}	[4, 7, 11]	[1, 1, 2]	[2, 2, 2]	[2, 2, 2]
w	[3, 12, 18]	[14, 17, 20]	[12, 13, 14]	[12, 14, 16]
ρ_{max}	[1, 6, 10]	[10, 18, 25]	[8, 9, 9]	[11, 11, 11]

While the above shows the results of each parameter separately, it is essential to understand how values of any two parameters covary. As an example, Figure 5.5 shows a matrix of scatter plots of each pair of parameters for the M60/CTM-basic modelling scenario. Such analysis shows that there is a clear nonlinear relationship between the maximum traffic density parameter ρ_{max} and backward wave speed parameter w which is consistent over all modelling scenarios. It is thus essential to account for the depicted relationship when, for example, considering the sensitivity of CTM to perturbation in parameter values.

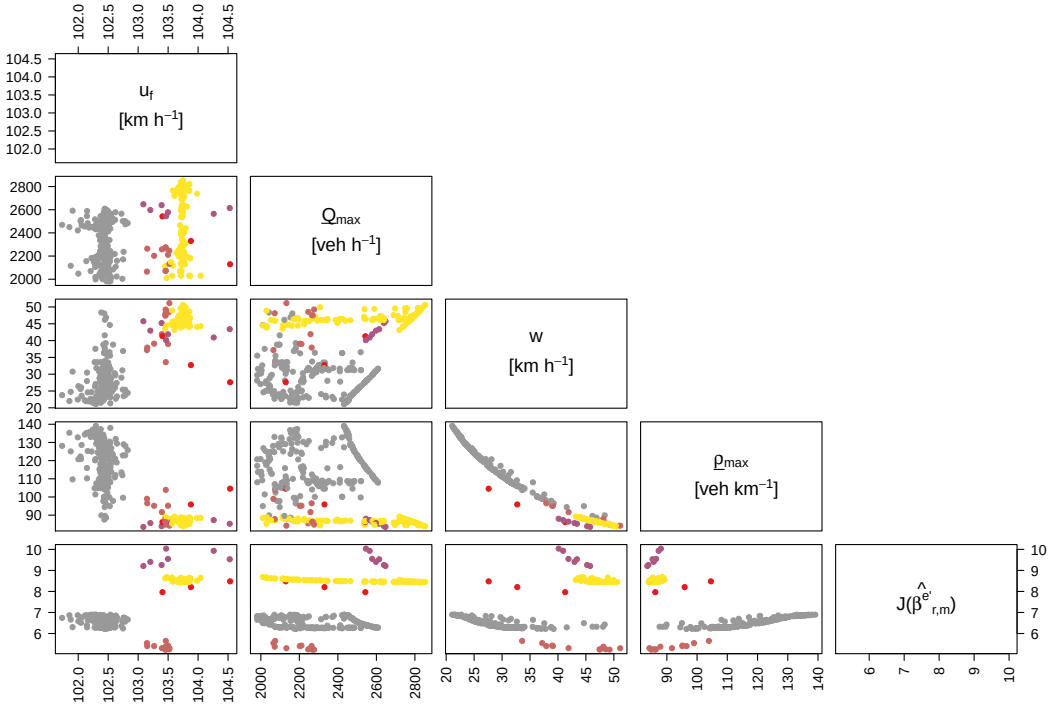


FIGURE 5.5: Matrix of scatter plots between each pair of parameters and objective function values for the *top-solutions* of M60/CTM-basic. The diagonal shows the parameters along with their units. The different coloured points represent the 5 time periods with *top-solutions*.

This relationship is examined in more details in Figure 5.6 which shows the values of both parameters obtained from all four modelling scenarios. The first observation

is that low values of $\underline{\rho}_{max}$ are accommodated by high values of w and vice versa. For instance, the lowest $\underline{\rho}_{max}$ value for the M60/CTM-basic obtained (84 veh km^{-1}) was accommodated by high w value (46 km h^{-1}). Also, the highest $\underline{\rho}_{max}$ value for the M1/CTM-basic obtained (504 veh km^{-1}) was accommodated by very low w value (4 km h^{-1}). While such extreme cases rarely occur in *top-solutions* (for instance the 99th percentile of $\underline{\rho}_{max}$ for the M1/CTM-basic is only 293 as shown in Table 5.2), they cannot be entirely ignored as they replicate certain features of measured traffic conditions given certain *top-solution* selection criteria.

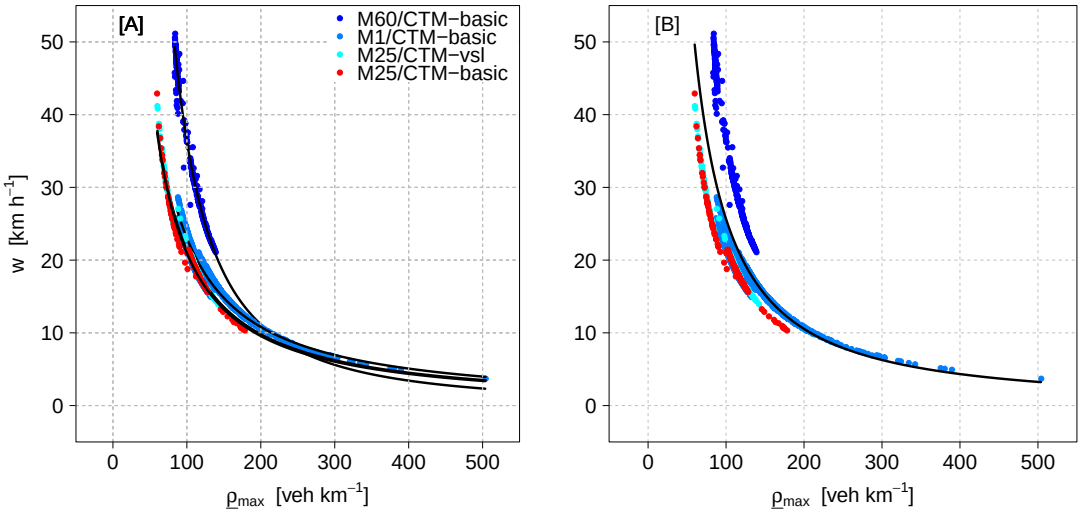


FIGURE 5.6: Scatter plots of maximum traffic density and backward wave speed parameters obtained from *top-solutions* of each CTM modelling scenario. Black coloured lines are fitted [A] on values of each scenario separately and [B] on values of all scenarios.

One possible use of this relationship is reducing the number of parameters to be optimised by setting one parameter (w) to be a function of another parameter ($\underline{\rho}_{max}$). One possible fit is the log-log model of the form:

$$log(w) = a + b \times log(\underline{\rho}_{max}) \quad \text{or} \quad w = e^{a} \times \underline{\rho}_{max}^{b}$$

where a is the intercept and b is the slope of the fitted line on a log-log scale. Figure 5.6A shows four log-log model fits using data points from each modelling scenario separately while Figure 5.6B shows a single log-log model fit using the entire data points. While having a single model for any traffic modelling scenario is more convenient, it generates higher fitting errors (as can be seen in Figure 5.6B) which can be problematic in terms of finding a *top-solution* given such an assumption (three parameters rather than four). This is reflected by the coefficient of determination (R^2) measure which drops from values higher than 0.960 to 0.773; R^2 reflects the proportion of total variation in w which is explained by the log-log model. The intercept, slope, and R^2 obtained for each model are as follows:

- for the M60/CTM-basic: $log(w) = 11.4 - 1.7 \times log(\underline{\rho}_{max})$ - $R^2 = 0.991$;

- for the M1/CTM-basic: $log(w) = 8.2 - 1.1 \times log(\underline{\rho}_{max})$ - $R^2 = 0.967$;
- for the M25/CTM-vsl: $log(w) = 8.2 - 1.1 \times log(\underline{\rho}_{max})$ - $R^2 = 0.985$;
- for the M25/CTM-basic: $log(w) = 8.3 - 1.1 \times log(\underline{\rho}_{max})$ - $R^2 = 0.984$; and
- for all modelling scenarios: $log(w) = 9.2 - 1.3 \times log(\underline{\rho}_{max})$ - $R^2 = 0.773$.

In addition to the clear relationship between these two parameters across all modelling scenarios, relationships between other parameters can also be noted. For instance, high backward wave speed are typically associated with high free-flow speed, although the range of free-flow speed values is very low. These are typically dependent on the modelling scenario and particularly on the range of parameter values obtained across the different time periods. This leads to another observation on Figure 5.5, which is the clustering of obtained parameter values based on the calibrated time period. This is clear for time periods [4] (yellow coloured circles) and [6] (grey coloured circles) for which the highest number of top-solutions was obtained. While the exact reason cannot be specified due to the complexity of traffic data across different time periods and within a time period, possible interpretations are: different fleet mix distributions which is driving the differences between time periods since fleet mix is not accounted for in the model; and different distributions of traffic flow and density at different states leading to different optimisation results. Alternatively (and more specifically), time period [6] is characterised by higher free-flow data points in comparison to time period [4]. Thus the optimisation algorithm led to lower objective function values in comparison to other time periods. This however was accompanied by highest variations in its supply parameters.

5.2.1.2 Model validation

The resulting *top-solutions* for each time period are then validated against all other time periods within the same road network (of a modelling scenario). There exist two ways to presenting validation results. The first answers the question of how *top-solutions* of a particular time period perform on any other time period (hereafter referred to as 'method 1'). The second answers the question of how *top-solutions* of all other time periods perform on a particular time period (hereafter referred to as 'method 2'). Typically, a calibrated time period with a calibrated parameter vector is used, for instance, for the simulation of future scenarios. Method 1 is thus useful to quantify the performing range of calibrated solution on other time periods. However, method 2 allows to better understand how the characteristics of other time periods influence such performing range. Validation results are presented in Figure 5.7 and Figure 5.8 based on these two methods, respectively.

An example of what Figure 5.7 shows is described for the M60/CTM-basic: 3 *top-solutions* have been identified for time period [1]; so, time period [1] on the figure shows the objective function values resulting from their simulation using the input of each of the 5 remaining time periods. These are shown as dark orange coloured circles. This

results in 15 simulations (3×5). The same is applied for all other time periods with *top-solutions*.

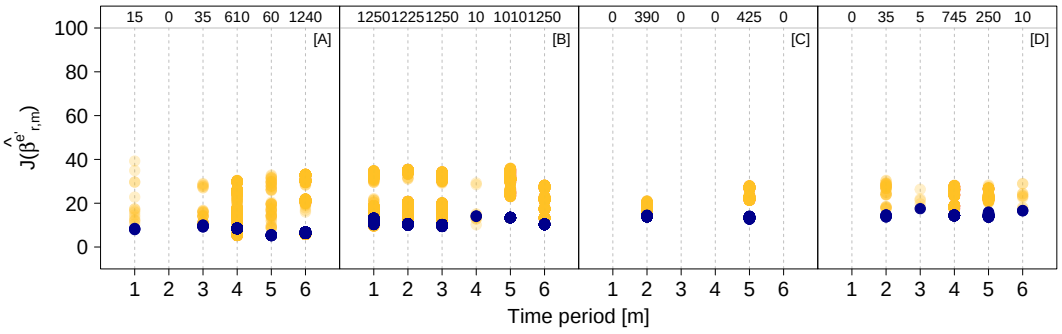


FIGURE 5.7: Objective function values obtained from the validation of *top-solutions*, which are highlighted in dark blue coloured circles, of each CTM modelling scenario (method 1). Total number of validation runs for each time period is highlighted on top.

Similarly, an example of what Figure 5.8 shows is described for the M60/CTM-basic: 389 *top-solutions* have been identified for all time periods except for time period [1]; so, time period [1] on the figure shows the objective function values resulting from their simulation using the input data of time period [1]. Unsurprisingly, results show, almost always, higher values in comparison to those obtained from calibration data.

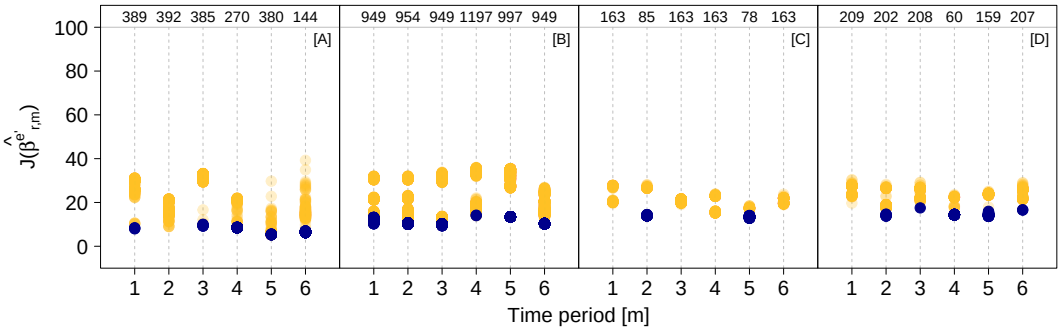


FIGURE 5.8: Objective function values obtained from the validation of *top-solutions* of each CTM modelling scenario (method 2).

This increase in values is examined in details in Figure 5.9. The relative increase (decrease) is calculated as a factor for each validation run (referred to here as validation factor). These are presented for each time period in Figure 5.9A and Figure 5.9B depending on the method, as well as for each modelling scenario (i.e. irrespective of time period) in Figure 5.9C.

As an example, the 3 *top-solutions* have 15 validation values. Each is linked with a calibration value depending on the *top-solution* validated. The factor is the validation value divided by its corresponding calibration one. The 15 factors are plotted for time period [1] of the M60/CTM-basic in Figure 5.9A. A factor higher than 1 means that validation has led to a higher value and vice versa. Similarly, the 389 *top-solutions*

mentioned earlier have 389 validation values for time period [1] and also 389 calibration values depending on which time period they have been calibrated for. These are plotted for time period [1] of the M60/CTM-basic in Figure 5.9B.

Validation factors range from 0.6 to 5.9 across all modelling scenarios with 583 validation runs (out of a total of 9815) leading to a factor lower than 1. A mean factor of 3, 1.9, 1.6, and 1.7 is calculated for the M60/CTM-basic, M1/CTM-basic, M25/CTM-vsl, and M25/CTM-basic, respectively. Here, the two most interesting observations, which follow from re-examining the speed distribution of measurements of Figure 4.7 in Chapter 4, are described below.

1. Measured speed distributions for all 6 time periods of the M25 are the least dissimilar (qualitatively) in comparison to those of the M60 and M1. In addition to the fact that lower number of *top-solutions* are obtained for the M25, this is a valid reason of why the validation factors for the M25/CTM-vsl and M25/CTM-basic are the lowest in range and variation (Figure 5.9C).
2. Measured speed distributions which show highest peaks under free-flow conditions (high speeds) are for time periods [5] and [6] of M60/CTM-basic and time period [3] for M1/CTM-basic. When *top-solutions* for these time periods are validated on other time periods, they produce very high validation factors (Figure 5.9A). When *top-solutions* from other time periods are validated on these time periods, they produce the lowest validation factors, often lower than 1 (Figure 5.9C). This is mostly because free-flow speed parameter is the least uncertain parameter, validating a parameter vector on a data set with lower periods of congestion produces lower errors; but validating parameter vectors obtained from the least congested time periods might produce highly uncertain supply parameters (i.e. produces high errors when validated on relatively more congested time periods).

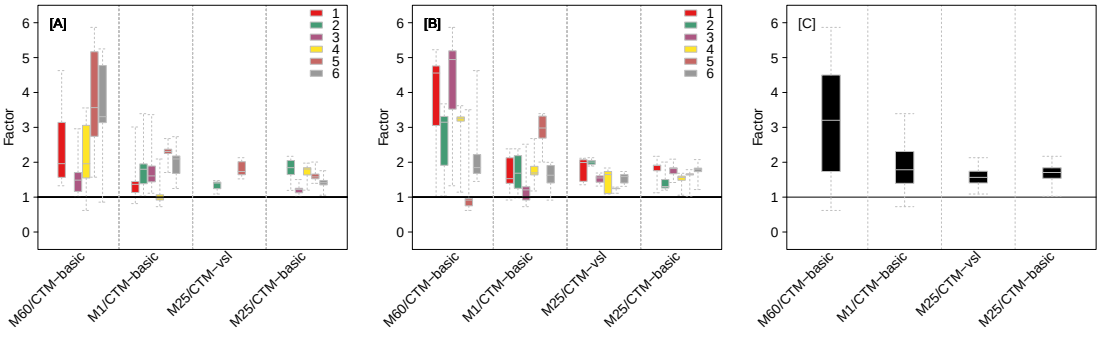


FIGURE 5.9: Validation objective function value divided by the calibration objective function value (factor) of each optimisation run based on [A] each time period of each CTM modelling scenario using method 1; [B] each time period of each CTM modelling scenario using method 2; and [C] irrespective of time periods of each CTM modelling scenario. The coloured boxes show the 25^{th}, 50^{th}, and 75^{th}. The whiskers represent the minimum and maximum factor.

However, in order to better understand the performance of CTM validation outputs in comparison to CTM calibration outputs, analysis of individual errors need to be performed. These are discussed in the results of Phase II, § 5.3.1.

5.2.2 Second-order METANET

5.2.2.1 Model calibration

Phase I results of the four M-METANET (i.e. METANET with a modified fundamental diagram) modelling scenarios are presented here. A justification of why O-METANET (i.e. METANET with original fundamental diagram) was disregarded is expanded upon here using an example for the M60 road network and time periods.

Unlike the CTM, only three out of eight parameters of the M-METANET are fundamental diagram-related. While their initial distributions have been deduced using the same data and procedure for the CTM, those of the other five parameters have been obtained from published literature. Poole and Kotsialos (2012, 2016) have proposed parameter bounds for the METANET model parameter vectors. Uniform distributions, $\sim \mathcal{U}(min, max)$, with the proposed bounds set as ranges, have been used here for the five non-fundamental diagram parameters. Since the proposed optimisation algorithm here does not require parameter bounds, an assessment of validity of those used in other studies can be made.

Figure 5.10 shows the distribution of the three modified fundamental diagram parameters. The free-flow parameter distribution is comparable to that obtained for the CTM, $\sim \mathcal{N}(104, 1.8)$. However, when fitting the original fundamental diagram, it resulted with a mean and variance free-flow speed values of 117 and 11.4, respectively, as expected. The normal distribution approximations is used to draw their initial values as an input for each optimisation run e of the M-METANET modelling scenarios. Assumptions of initial distributions (with the test of normality results) are as follows:

- $u_f \sim \mathcal{N}(104, 1.7)$, p-value $= 0.007$;
- $\alpha \sim \mathcal{N}(0.5, 0.007)$, p-value $= 0.01$;
- $\underline{\rho}_c \sim \mathcal{N}(15, 0.5)$, p-value $= 0.02$;
- $\tau \sim \mathcal{U}(1/3600, 50/3600)$;
- $\eta \sim \mathcal{U}(1, 90)$;
- $\delta \sim \mathcal{U}(0.001, 4)$;
- $\phi \sim \mathcal{U}(0.01, 3)$; and
- $\kappa \sim \mathcal{U}(5, 90)$.

In the absence of lane drop for the M60 and M1 road networks, ϕ parameter does not have to be accounted for/calibrated. This results in only seven parameters to be calibrated for the M60/M-METANET-basic and M1/M-METANET-basic while it is eight parameters for the M25/M-METANET-vsl and M25/M-METANET-basic.

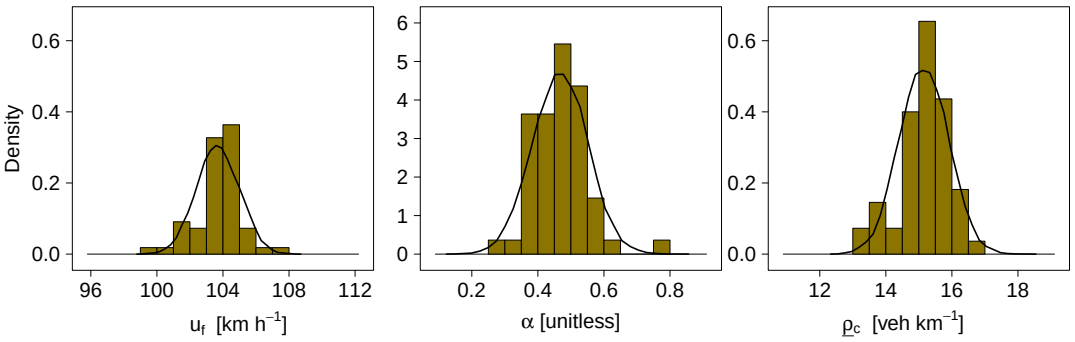


FIGURE 5.10: Histogram of the fitted modified METANET fundamental diagram-related parameters on each of the 55 normal weekdays (from April 2014 to June 2014) traffic flow-density data on the M25 road network overlaid by their normal distribution approximations $\sim \mathcal{N}(\mu, \sigma^2)$.

For each time period and M-METANET modelling scenario, E parameter vectors are randomly drawn to initialise the optimisation runs. Figure 5.11 shows the objective function values of each ensemble of optimisation runs and highlights the *top-solutions*, based on selected criteria. While a minimum traffic density can be specified in the METANET model structure to prevent negative values from occurring for certain parameter vector values (by limiting the traffic density), these were instead penalised. As a result a number of penalised solutions were obtained and are not shown in the figure: 55 for M60/M-METANET-basic, 24 for M1/M-METANET-basic, 84 for M25/M-METANET-vsl, and 85 for M25/M-METANET-basic. Figure 5.12 shows only those for the *top-solutions*.

Similar to CTM, the range of values is dependent on the time period, road network, and model structure; and relatively higher values are obtained for the M25 optimisation runs. The latter strengthens previous speculation that this is a result of the complexity of VSL-operated road networks which is harder to model (using either traffic flow models). Unlike the CTM, Figure 5.11 and Figure 5.12 show that *top-solutions* are identified for each time period and each modelling scenario. This is mostly a result of second-order models being able to reproduce features of traffic data that at least meet the statistical testing criteria. This is possibly a result of using a more complex fundamental diagram in comparison to the simplified trapezium one for CTM; describing the dynamics of traffic speed rather than assuming a steady-state traffic speed-density equation for CTM; and/or taking into account the influence of lane drop and on-ramp merging. Similar to the CTM, Figure 5.11 and Figure 5.12 also show that the number of *top-solutions* per time period is generally lower for the M25/M-METANET-vsl and M25/M-METANET-basic modelling scenarios in comparison to those of the M60/M-METANET-basic and M1/M-METANET-basic.

This again shows that there is a trade-off between uncertainty in model structure and uncertainty in parameter vectors. While the first is reduced for M-METANET, higher uncertainty in its parameter vector is observed in comparison to CTM. While

uncertainty in the M-METANET model structure is reduced for non-VSL road networks, higher uncertainty in its parameter vector is observed in comparison to those of the VSL-operated road network.

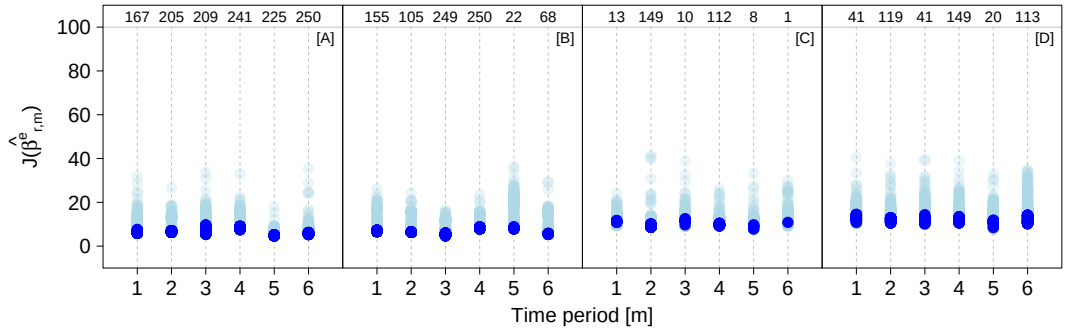


FIGURE 5.11: Minimised objective function values obtained from the calibration runs of [A] M60/M-METANET-basic, [B] M1/M-METANET-basic, [C] M25/M-METANET-vsl, and [D] M25/M-METANET-basic. From now on, the results in figures of the four M-METANET modelling scenarios follow this order. Few calibration runs did not converge, as explained in text and thus are not shown here.

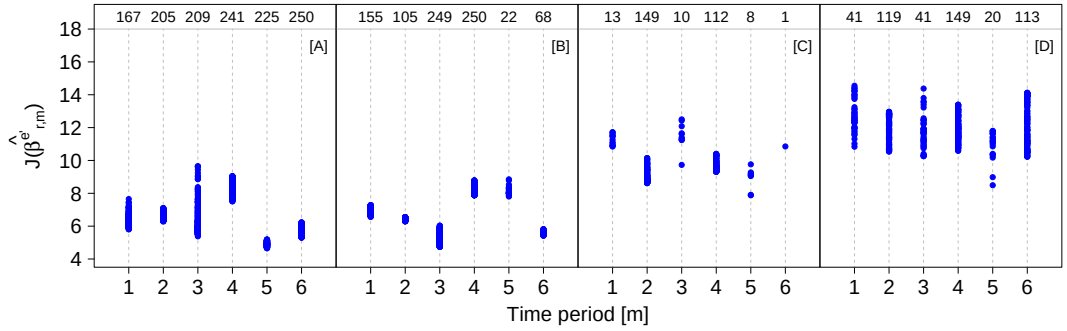


FIGURE 5.12: Minimised objective function values obtained from the *top-solutions* only of each M-METANET modelling scenario.

Table 5.4 shows the number of optimisation runs which have an objective function value less than the threshold and have satisfied the two statistical tests using C1 and/or C2. It shows that there are more optimisation runs which satisfy both C1 and C2 (a total of 670 across all modelling scenarios), in comparison to the CTM (a total of 84 only). This can be a direct result of M-METANET describing the dynamics of both traffic density and speed. Nevertheless, satisfying both conditions is still not frequent as satisfying either conditions (i.e. number of *top-solutions*). The M25/M-METANET-basic resulted in higher number of *top-solutions* (483) compared to the M25/M-METANET-vsl (293). However, unlike the CTM, *top-solutions* are obtained for all 6 time periods under both model structures. This suggests that the M-METANET ability to take into account the dynamics of traffic speed might have accommodated for the deficiencies of the VSL model structure of both CTM and M-METANET. Nevertheless, whether such differences have an influence on individual prediction errors or not is better explored in the next phase.

TABLE 5.4: number of *top-solutions* of each M-METANET modelling scenario satisfying the threshold percentile and statistical tests criteria

Statistical Condition	**M60 M-METANET -basic**	**M1 M-METANET -basic**	**M25 M-METANET -vsl**	**M25 M-METANET -basic**
C1	1001	596	37	150
C2	844	300	278	386
C1&C2	548	47	22	53
Total	1297	849	293	483

Now that *top-solutions* for the M-METANET modelling scenarios are determined, their parameter vectors are examined in more details. An example of how the values of the three M-METANET's parameters: u_f, α, and $\underline{\rho}_c$ translate to the modified fundamental diagram is provided in Figure 5.13 for each time period of the M60/M-METANET-basic. These parameters are only used to calculate the equilibrium traffic speed equation of M-METANET. The same plots are generated for the time periods of the other three modelling scenarios and are shown in Appendix B.

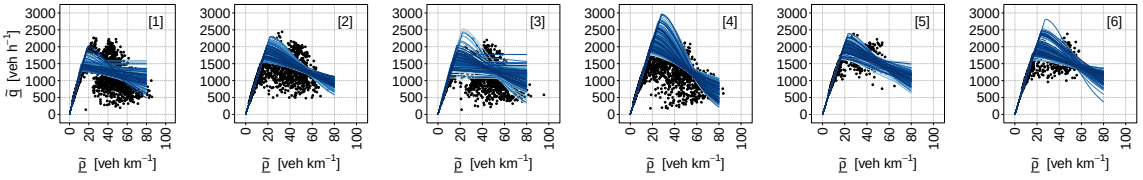


FIGURE 5.13: Traffic flow per lane [veh h^{-1}] and density per lane [veh km^{-1}] scatter plots showing the constructed fundamental diagrams based on *top-solutions* of M60/M-METANET-basic: 167 for time period [1], 205 for time period [2], 209 for time period [3], 241 for time period [4], 225 for time period [5], and 250 for time period [6].

Table 5.5 shows the range of values of each parameter for each M-METANET modelling scenario, irrespective of time period. The range of free-flow speed parameter is larger in comparison to those obtained for the CTM (despite the same physical meaning). This results from the fact that high values outside the expected range are accommodated for by the addition of other four or five parameters to M-METANET. Nevertheless, this issue is diminished when compared with the O-METANET fundamental diagram. Using the original fundamental diagram equation, ensemble-based optimisation was conducted for the M60 time periods, *top-solutions* were identified, and resulting parameter values were compared to those of the M-METANET.

Figure 5.14B shows how very large values of free-flow speed parameters are occasionally present but are accommodated for by a reduction in the α parameter. This issue is partially resolved when the modified fundamental diagram is used, as shown in Figure 5.14A. However, another issue arises: very low values of α are obtained for some scenarios, particularly for time periods with large scatter in the congestion phases. Very low values of α implies that the congested-part of the resulting fundamental diagram is

TABLE 5.5: [minimum, 99th percentile, maximum] of parameter vector values obtained from the *top-solutions* of each M-METANET modelling scenario

$\boldsymbol{\beta}$	**M60 M-METANET -basic**	**M1 M-METANET -basic**	**M25 M-METANET -vsl**	**M25 M-METANET -basic**
u_f	[103, 121, 135]	[106, 114, 118]	[98, 117, 141]	[95, 116, 149]
α	[1×10^{-3}, 2, 2]	[2×10^{-3}, 1, 2]	[1×10^{-3}, 1, 1]	[3×10^{-3}, 1, 1]
ρ_c	[9, 26, 28]	[10, 20, 20]	[8, 17, 18]	[11, 18, 18]
τ^1	[1, 134, 229]	[2, 86, 170]	[10, 65, 138]	[2, 76, 233]
η	[6, 273, 693]	[7, 209, 4023]	[11, 147, 259]	[13, 186, 372]
δ	[1×10^{-7}, 4, 9]	[2×10^{-8}, 2, 4]	[1×10^{-6}, 5, 6]	[9×10^{-8}, 5, 7]
ϕ	NA	NA	[1×10^{-6}, 1, 2]	[6×10^{-7}, 4, 5]
κ	[5×10^{-4}, 212, 622]	[3×10^{-3}, 179, 3611]	[6×10^{-3}, 146, 203]	[1×10^{-5}, 162, 270]

[1] This is $\tau \times 3600$ i.e. reported in seconds for convenience

almost a horizontal straight line, as shown for some cases in Figure 5.13: time period [1] and time period [3]. When one is choosing a single *top-solution* (rather than an ensemble of *top-solutions*), those which give large values of free-flow speed parameters might be disregarded. In this study, however, these are rather kept because they meet statistical testing criteria; and as a result, they highlight issues of the M-METANET model arising from its parameter vector. As a result of the issue observed in Figure 5.14B for the O-METANET modelling scenarios, from now on only the M-METANET modelling results are presented and discussed.

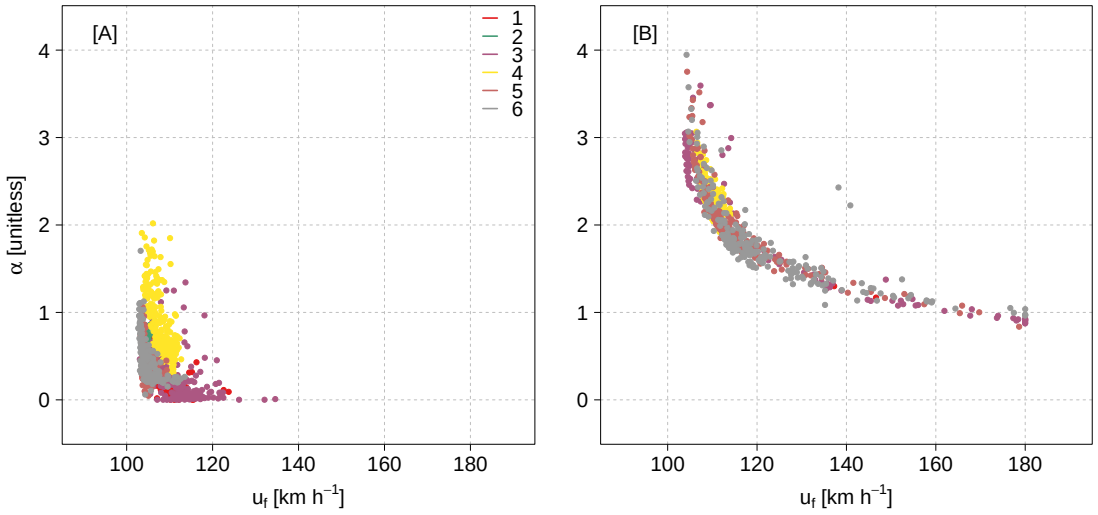


FIGURE 5.14: Scatter plots of α versus u_f parameter values obtained from the *top-solutions* of [A] M60/M-METANET-basic and [B] M60/O-METANET-basic.

The range of values obtained for other (non-fundamental diagram) parameters presented in Table 5.5 always exceeds the bounds proposed by Poole and Kotsialos (2012, 2016) which are used to initialise the optimisation runs. However, in most cases, more than 80% lie within the bounds. Particularly,

- 80%, 95%, 97%, and 92% lie within the proposed range of the parameter τ for each of the M60/M-METANET-basic, M1/M-METANET-basic, M25/M-METANET-vsl, and M25/M-METANET-basic, respectively - from now on, the results of the four M-METANET modelling scenarios follow this order;
- 70%, 78%, 84%, and 72% lie within the proposed range of the parameter η for each modelling scenario;
- 58%, 56%, 96%, and 90% lie within the proposed range of the parameter δ for each modelling scenario;
- 96%, and 89% lie within the proposed range of the parameter ϕ for each of the M25/M-METANET-vsl and M25/M-METANET-basic, respectively; and
- 80%, 87%, 88%, and 84% lie within the proposed range of the parameter κ for each modelling scenario.

There exist few extreme cases, where a very high value of one parameter is compensated by the values of other parameters, as also highlighted in Papageorgiou et al. (1990). This has been particularly the case for one *top-solution* of M1/M-METANET-basic where κ and η values of 3611 and 4023 are obtained. While one can remove such solutions, the results allow to examine model structural issues. Such high values illustrate the issue of the additional parameters of M-METANET, especially in the context of a complex optimisation problem.

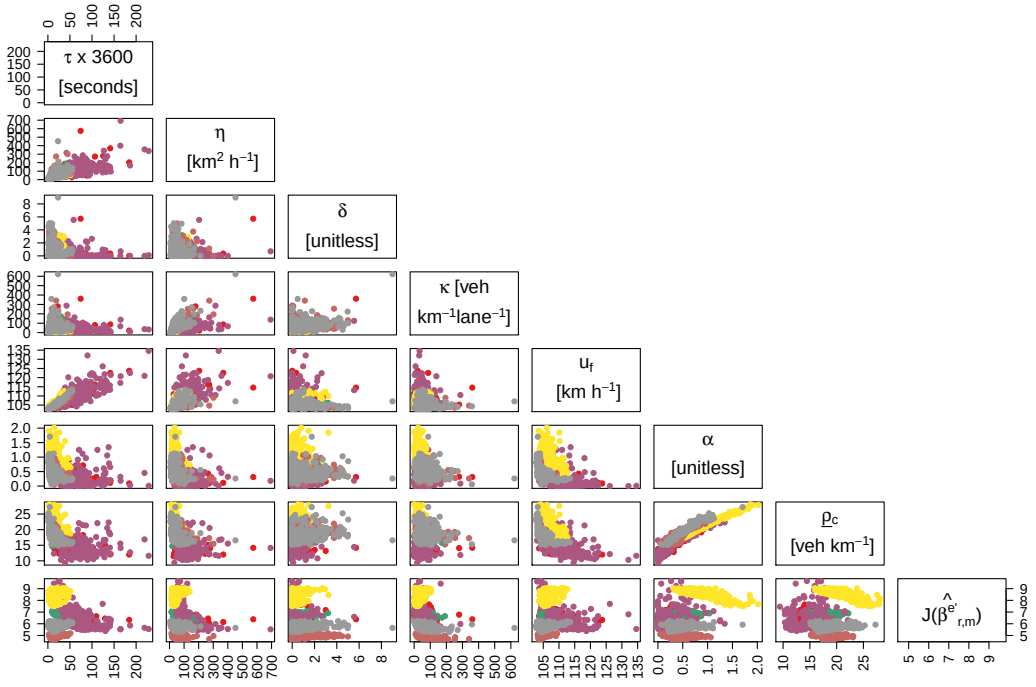


FIGURE 5.15: Matrix of scatterplots between each pair of parameters and objective function values for the *top-solutions* of M60/M-METANET-basic. Extreme κ and η point not shown here but remains in subsequent analysis.

In Figure 5.15 above, an example of the relationship between the values of each pair of parameters is provided for the M60/M-METANET-basic. There is a positive

correlation between: τ and η; η and κ; τ and u_f; η and u_f; and α and $\underline{\rho}_c$. This is also the case for the other modelling scenarios. However, the high scatter does not permit to construct particular models similar to those proposed for CTM (i.e. in Figure 5.6). Figure 5.15 also does not show clear clustering of parameter values across time periods, unlike what has been observed for CTM. One possible interpretation is the additional four (directly unobservable) parameters of METANET. These play a key role in the optimisation process and possibly accommodate for relationships that could otherwise be observed for those of the fundamental diagram.

TABLE 5.6: [minimum, mean, maximum] CV in % of parameter vector values obtained across the time periods with more than 10 *top-solutions* of each M-METANET modelling scenario

$\boldsymbol{\beta}$	**M60 M-METANET -basic**	**M1 M-METANET -basic**	**M25 M-METANET -vsl**	**M25 M-METANET -basic**
u_f	[1, 2, 4]	[0.4, 1, 2]	[2, 3, 5]	[2, 3, 5]
α	[37, 56, 104]	[14, 29, 69]	[19, 26, 32]	[16, 29, 45]
$\underline{\rho}_c$	[8, 12, 18]	[3, 7, 16]	[3, 4, 5]	[3, 4, 6]
τ[1]	[31, 51, 72]	[27, 39, 65]	[23, 30, 38]	[32, 52, 84]
η	[50, 59, 77]	[40, 107, 340]	[41, 51, 69]	[38, 56, 93]
δ	[115, 192, 297]	[142, 208, 286]	[46, 59, 72]	[68, 79, 89]
ϕ	-	-	[63, 72, 84]	[75, 103, 120]
κ	[59, 71, 85]	[58, 139, 448]	[35, 47, 61]	[65, 77, 88]

[1] This is $\tau \times 3600$ i.e. reported in seconds for convenience

Table 5.6 above shows the high CVs obtained for each parameter across the different time periods and modelling scenarios. Lowest CV values are obtained for the demand parameters of the fundamental diagram (u_f and $\underline{\rho}_c$) in comparison to the supply parameter α as a result of the large scatter in the congested time periods; and in comparison to other parameters which do not carry the physical meaning of the fundamental diagram. Similar to CTM, the high ranges, large scatter, and high CV values obtained generally illustrate the difficulties in finding a 'global' solution for the problem and justify the use of an ensemble-based optimisation for the case of M-METANET as well.

5.2.2.2 Model validation

Resulting *top-solutions* for each time period and road network are then validated against all other time periods within the same road network. Figure 5.16 shows the distribution of the relative increase (decrease) in objective function values obtained from the validation of *top-solutions* of each time period validate against all other time periods (method 1 described earlier in § 5.2.1.2); and validation of *top-solutions* of all other time periods against each remaining time period (method 2).

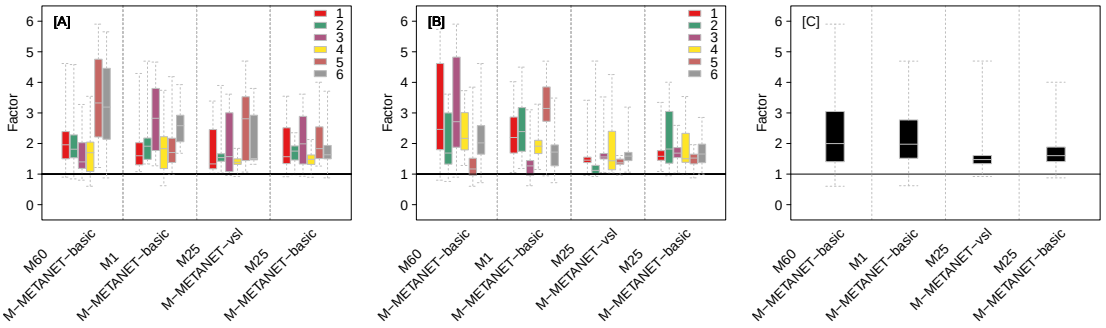


FIGURE 5.16: Validation objective function value divided by the calibration objective function value (factor) of each optimisation run based on [A] method 1; [B] method 2; and [C] irrespective of time periods of each M-METANET modelling scenario.

Results obtained are very similar to those of the CTM validation but also clearer since the number of M-METANET *top-solutions* are generally higher. These are as follows.

1. Range of factors is exactly the same which is between 0.6 and 5.9 across all modelling scenarios. In most cases, factors are higher than 1 indicating lower model performance when the parameter vectors are simulated on other time periods.
2. *Top-solutions* of the least congested time periods have led to highest factors when validated against other more congested time periods. *Top-solutions* of the more congested time periods have led to lowest validation factors (often lower than 1) when validated against the least congested time periods.
3. Mean factors are higher for the M60/M-METANET-basic and M1/M-METANET-basic (2.3 and 2.2, respectively) in comparison to the M25/M-METANET-vsl and M25/M-METANET-basic (1.5 and 1.7, respectively) where speed measurements are least dissimilar.

Having said that, validation results of the M-METANET modelling scenarios are compared with those obtained from the CTM modelling scenarios in more details in Phase II by examining the distributions of individual errors rather than a single measure (i.e. rather than only the objective function values of mean absolute errors in speed).

5.2.3 A note on ensemble-based optimisation parameters

To study the influence of the three ensemble-based optimisation parameters (or inputs) on the results, three scenarios were devised and compared with the initial scenario (here, it is referred to as the 'base-case' scenario). The three scenarios are:

- Scenario 1: reducing the number of optimisation runs to 100 while keeping both the threshold percentile and significance level as the 'base-case' scenario, $E = 100$, $\chi = 50\%$, and $\xi = 0.01$;

- Scenario 2: reducing both the threshold percentile to 10% and the significance level to 0.1 while keeping the number of optimisation runs as the 'base-case' scenario, $E = 500$, $\chi = 10\%$, and $\xi = 0.1$; and
- Scenario 3: reducing all 3 inputs to $E = 100$, $\chi = 10\%$, and $\xi = 0.1$.

The influence of these three is examined using the M60/CTM-basic and M60/M-METANET-basic modelling scenarios, in two ways: the number of *top-solutions* obtained from each scenario is compared with the 'base-case' scenario; and the distribution of each parameter of *top-solutions* is compared with that of the 'base-case' scenario.

Table 5.7 and Table 5.8 below show the number of *top-solutions* based on the different scenarios for the M60/CTM-basic and M60/M-METANET-basic, respectively. Reducing the number of optimisation runs E from 500 to 100 (scenario 1) has a consistent influence for both the CTM and M-METANET. This is a reduction in the number of *top-solutions* by approximately a factor of 5. Since the number of *top-solutions* is high for the M-METANET across all time periods, *top-solutions* are still found for each one. This is not the case for time periods [1-3] using CTM, where *top-solutions* are not found.

TABLE 5.7: number of *top-solutions* obtained from each ensemble-based input scenario for the M60/CTM-basic

Time Period	**'Base-case' Scenario** $E = 500$; $\chi = 50\%$; $\xi = 0.01$	**Scenario 1** 100; 50%; 0.01	**Scenario 2** 500; 10%; 0.1	**Scenario 2** 100; 10%; 0.1
[1]	3	0	0	0
[2]	0	0	0	0
[3]	7	0	0	0
[4]	122	27	0	0
[5]	12	1	0	0
[6]	248	50	50	10
$\sum$	**392**	**78**	**50**	**10**

TABLE 5.8: number of *top-solutions* obtained from each ensemble-based input scenario for the M60/M-METANET-basic

Time Period	**'Base-case' Scenario** $E = 500$; $\chi = 50\%$; $\xi = 0.01$	**Scenario 1** 100; 50%; 0.01	**Scenario 2** 500; 10%; 0.1	**Scenario 2** 100; 10%; 0.1
[1]	167	38	37	7
[2]	205	39	33	8
[3]	209	48	41	10
[4]	241	46	44	8
[5]	225	48	0	0
[6]	250	50	49	10
$\sum$	**1297**	**269**	**204**	**43**

When the threshold percentile χ is restricted to 10% and significance level is increased to 0.1 (i.e. scenarios 2 and 3), *top-solutions* are found for a single time period in the case of CTM but for all time periods in the case of M-METANET. Dropping the threshold percentile to 10% directly drops the remaining solutions in the ensemble to

100 or 10, depending on the value of E. Increasing the significance level also reduces the chance of not rejecting the null hypothesis of the two statistical tests. Here, it is important to mention that while both inputs E and χ do not have any underlying assumptions, this is not the case for the ξ input. Relaxing the significance level (for e.g. 0.01 as in the 'base-case') reduces type-I error of statistical testing at the expense of increased type-II error. This implies that while the risk to rejecting a *top-solution* which captures the distribution and variance of measured traffic variable is reduced, the risk to accepting a *top-solution* which does not capture the distribution and variance of measurements is increased. Here, the reason behind relaxing the significance level to 0.01 is mainly due to difficulty of getting a *top-solution* for the CTM under demanding condition, which prevents further analysis and comparison. At the same time, such result exposes the structural issues of the CTM in comparison to M-METANET. M-METANET was almost always able to find solutions. The flexibility in the M-METANET in terms of additional parameters, modified fundamental diagram, and additional speed dynamics equation allows for finding *top-solutions* even under demanding criteria.

In addition to that, the ranges of parameter values obtained for the three scenarios become much lower for the CTM given scenarios 2 and 3, as shown in Figure 5.17. The distribution of parameter values is maintained in scenario 1. However, the range of parameter values for the M-METANET remains almost the same across the three scenarios, as shown in Figure 5.18. While the fact that finding *top-solutions* across all time periods is advantageous of the M-METANET model structure, it is at the expense of higher uncertainty in its parameterisation, not only in terms of the number of *top-solutions* but also in terms of parameter values.

To conclude, this section shows that results remain almost the same for CTM and M-METANET when reducing the number of optimisation runs only to 100, suggesting that 100 can be enough for the purpose of our analysis, at the expense of reducing the population of prediction error points for analysis in subsequent phases. This section also shows that reducing the significance level influences the results, especially for the

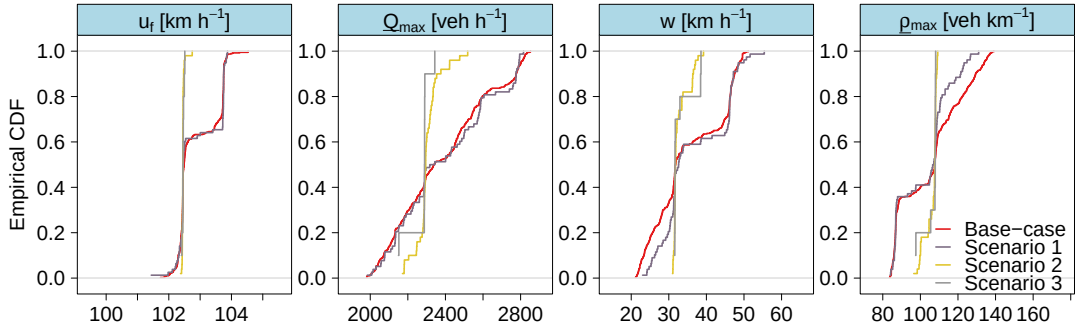


FIGURE 5.17: Cumulative distribution of the four M60/CTM-basic parameter values of *top-solutions* obtained based on the 'base-case' scenario and the three scenarios.

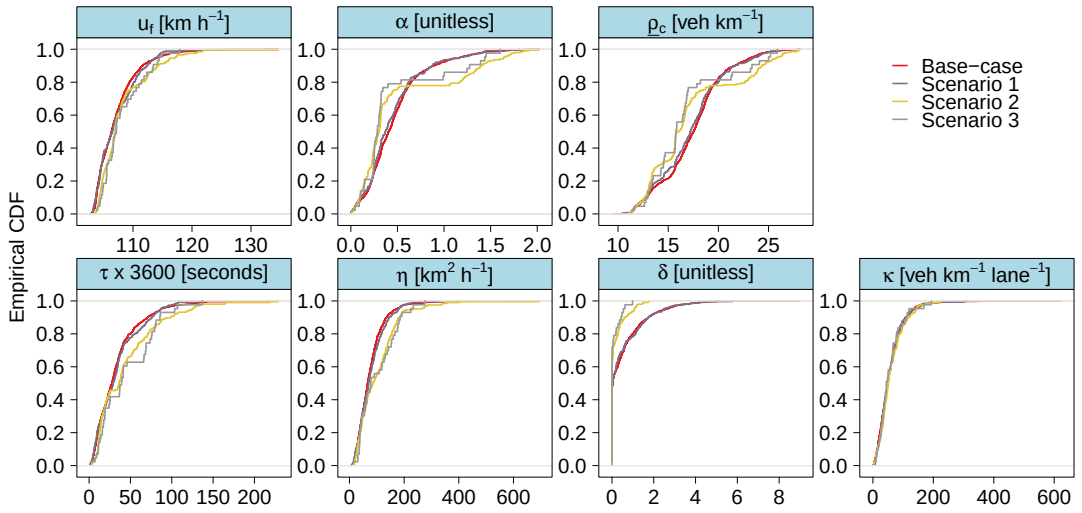


FIGURE 5.18: Cumulative distribution of the seven M60/M-METANET-basic parameter values of *top-solutions* obtained based on the 'base-case' scenario and the three scenarios.

CTM, suggesting that the original input of 0.01 is needed in order to at least get *top-solutions* across different time periods of the modelling scenario, which would allow examining the results of subsequent phases. This is, of course, at the expense of higher risk in examining *top-solutions* that does not capture the distribution and variance of measurements.

5.3 Phase II: Prediction error estimation

The main input of Phase II is the *top-solutions* obtained in the previous phase. Unlike Phase I, there is not any additional parameters (or inputs) that need to be made here. Given the *top-solutions*, the following two sections present and discuss the results of the outputs of Phase II obtained for the four CTM modelling scenarios in § 5.3.1, and the four M-METANET modelling scenarios in § 5.3.2. Calibration and validation results are separately discussed in each section.

5.3.1 First-order CTM

5.3.1.1 Model calibration

The main outputs of Phase II are the two bivariate calibration and validation error distributions for each modelling scenario. These are obtained by simulating *top-solutions* obtained earlier on relevant time periods, extracting and aggregating the calibration or validation outputs to the measurement spatiotemporal resolution, and subtracting measurement values from the aggregated prediction outputs.

The calibration error distributions are shown in Figure 5.19, which shows the bivariate distribution, $\varepsilon_r^c(\bar{\rho}, \bar{u})$, for each CTM modelling scenario. As an example, $[3, 0, 7, 122, 12, 248]$ *top-solutions* (a total of 392) have been identified for each time period of the M60/CTM-basic modelling scenario. The first panel [A] in Figure 5.19 shows the estimated calibration errors from all 392 simulations with a total of 568272 aggregated traffic speed-density error points.

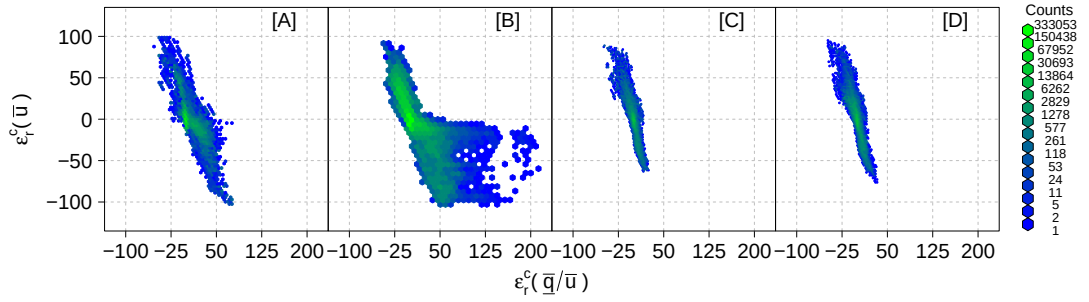


FIGURE 5.19: Bivariate calibration error distributions of aggregated traffic speed and density estimated from the *top-solutions* of each CTM modelling scenario.

The four estimated bivariate error distributions are summarised in Table 5.9. While these have been obtained from the simulation of *top-solutions* only, errors can be of high values. For instance, a maximum over-prediction (under-prediction) of 76 (52) veh km^{-1} and 98 (102) km h^{-1} for the M60/CTM-basic modelling scenario is reached. This is alongside the high standard deviation values of both variables. The standard deviation ranges between 7 and 13 for traffic density and between 13 and 20 for speed across the four modelling scenarios. Such high values, particularly of traffic speed, can be a direct result of the assumed steady-state traffic speed-density relationship, i.e. traffic speed adapting instantaneously to density. In addition to that, bivariate error distributions of M25/CTM-vsl and M25/CTM-basic have similar characteristics in terms of their summary statistics, suggesting that structural errors in using the basic CTM on a VSL-operated road network might have been compensated for by the optimisation problem. It is otherwise possible that M25/CTM-vsl structure does not improve upon the basic CTM structure, given the complexity of VSL-operation along with the fact that *top-solutions* were only found on 2 time periods out of 6. In either cases, the precise reason to such observation is hard to know in the context of this analysis, due to the complex interactions between the different components.

In order to examine how (and if) these bivariate error distributions vary by time period and *top-solutions*, the best perceived way (given the large number of *top-solutions*) is to present the summary statistics of the errors in each traffic variable. Figure 5.20 and Figure 5.21 show the $[0, 10, 25, 50, 75, 90, 100]^{\text{th}}$ percentiles of errors in traffic density and speed, respectively, for each *top-solution* and time period; the 0^{th} percentile and the 100^{th} percentile is the minimum and maximum, respectively. As an example,

TABLE 5.9: summary statistics of calibration errors in aggregated traffic density and speed, $[\varepsilon_r^c(\bar{\rho}), \varepsilon_r^c(\bar{u})]$, for each CTM modelling scenario

Summary Statistics	**M60 CTM-basic**	**M1 CTM-basic**	**M25 CTM-vsl**	**M25 CTM-basic**
Mean	[4, -1]	[1, 3]	[1, -1]	[0, 0]
Std. Dev.	[11, 13]	[13, 19]	[7, 20]	[7, 20]
Covariance	-102	-206	-126	-133
Minimum	[-52, -102]	[-44, -103]	[-49, -63]	[-48, -76]
Maximum	[76, 98]	[213, 91]	[24, 89]	[33, 96]
NR. of points	568272	1537554	357327	396621

time period [1] of the M60/CTM-basic has 3 *top-solutions*. 3 calibration error distributions of each variable are estimated. Each error distribution produces 3 error values of each percentile. Time period [1] of the M60/CTM-basic shows these values using different colours for the different percentiles. The figures show that differences from one *top-solution* to another and from one time period to another are mainly in the extreme values (i.e. minimum and maximum only) of both traffic variables. This is especially the case for time periods with high numbers of *top-solutions*, such as time period [4] of M60/CTM-basic and time periods [1-3] of M1/CTM-basic.

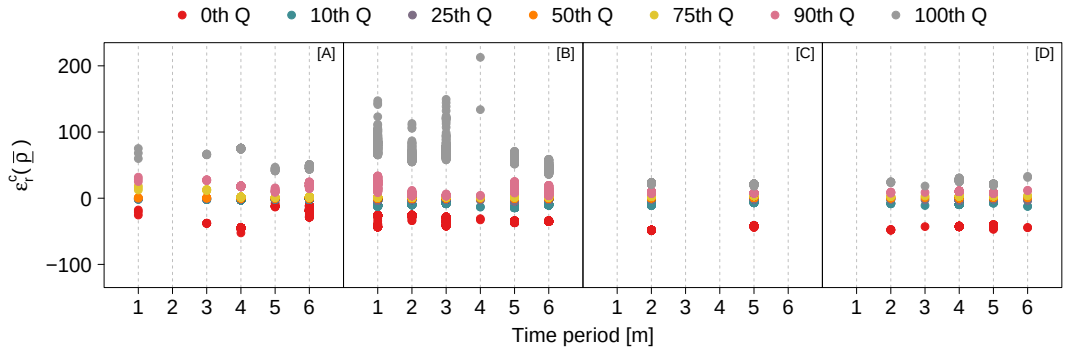


FIGURE 5.20: $[0, 10, 25, 50, 75, 90, 100]^{\text{th}}$ percentiles of calibration errors in aggregated traffic density for each *top-solution* and time period of each CTM modelling scenario.

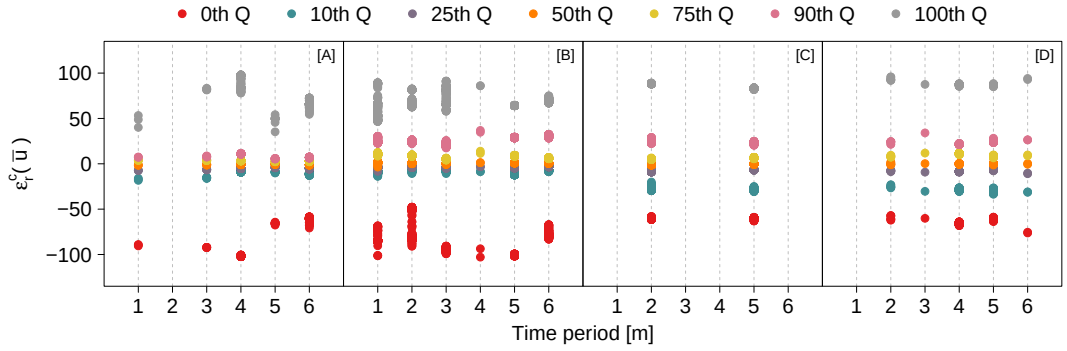


FIGURE 5.21: $[0, 10, 25, 50, 75, 90, 100]^{\text{th}}$ percentiles of calibration errors in aggregated traffic speed for each *top-solution* and time period of each CTM modelling scenario.

Figure 5.19 shows that there is a negative relationship between error distributions in traffic speed and density. This is also reflected in the negative covariances calculated in Table 5.9 for each modelling scenario. Hence, an over-prediction (i.e. positive errors) of traffic speed indicates an under-prediction (i.e. negative errors) of density and vice versa; this is a consequence of the traffic speed-density fundamental diagram. To better understand the strength of the linear relationship between the two variables, correlation coefficients (or Pearson's r) are estimated for each bivariate error distribution arising from each *top-solution* across the four CTM modelling scenarios. Estimated correlation coefficients were in the range of $[-0.82, -0.70]$, $[-0.92, -0.60]$, $[-0.92, -0.84]$, and $[-0.92, -0.84]$ for each of the four CTM modelling scenarios, respectively. This indicates that the distribution of each variable cannot be treated separately. For instance, a traffic speed error point of 100 km h^{-1} should not be sampled along with a density error point of 50 veh km^{-1}. Similar results have been observed for the validation error distributions of CTM as well as the calibration and validation errors of METANET. These are listed here instead.

- a range of $[-0.94, -0.89]$, $[-0.96, -0.58]$, $[-0.92, -0.90]$, and $[-0.93, -0.89]$ based on the validation outputs of each of the four CTM modelling scenarios, respectively.
- a range of $[-0.83, -0.51]$, $[-0.83, -0.56]$, $[-0.89, -0.57]$, and $[-0.91, -0.66]$ based on the calibration outputs of each of the four M-METANET modelling scenarios, respectively.
- a range of $[-0.95, -0.68]$, $[-0.93, -0.76]$, $[-0.93, -0.80]$, and $[-0.93, -0.83]$ based on the validation outputs of each of the four M-METANET modelling scenarios, respectively.

Another previously made assumption is shown here empirically, regarding the use of traffic density instead of flow-based density. The above results clearly use aggregated traffic density predictions. When both predictions and errors of aggregated traffic density are compared with flow-based density (aggregated flow divided by aggregated speed), an almost one-to-one relationship is observed. This is considered the main reason behind choosing one variable to continue with in the analysis.

Figure 5.22 illustrates this point. Figure 5.22A shows the similarity in the prediction outputs of the two variables while Figure 5.22B shows that of the estimated errors of the two variables. Although the aggregated values are not exactly the same, an R^2 of 0.99 has been calculated for each CTM modelling scenario. This is for both the relationship between the prediction outputs as well as the prediction errors. The same has been observed for the validation outputs of CTM as well as the calibration and validation outputs of METANET.

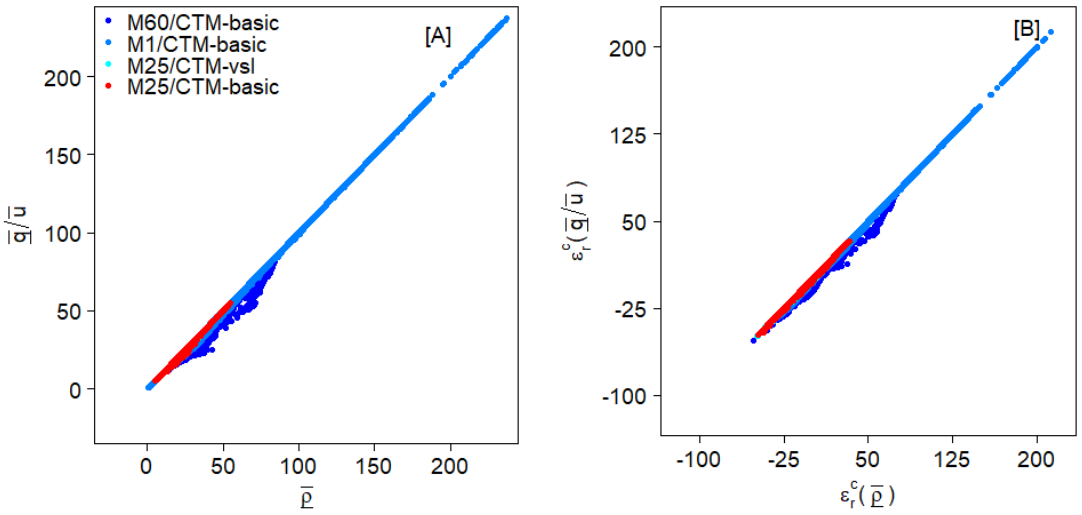


FIGURE 5.22: Calibration prediction [A] outputs and [B] errors of aggregated traffic flow-based density versus density for each CTM modelling scenario.

5.3.1.2 Model validation

By simulating each *top-solution* of each time period using the input data of each of the other 5 time periods, separately, validation errors of the variables of interest are estimated for each simulation run. Figure 5.23 shows the estimated validation errors of all *top-solutions* of each modelling scenario. For instance, the first panel [A] of Figure 5.23 shows those from all the $392 \times 5 = 1960$ simulations, with a total of 3679440 aggregated traffic speed-density error points.

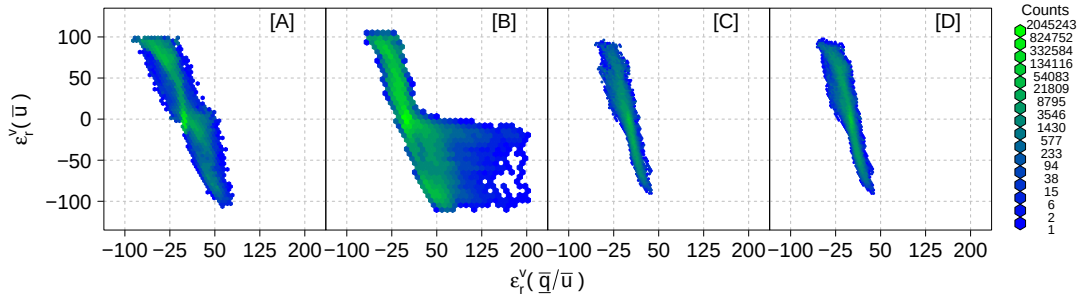


FIGURE 5.23: Bivariate validation error distributions of aggregated traffic speed and density estimated from the *top-solutions* of each CTM modelling scenario.

Summary statistics of the estimated bivariate error distributions are provided in Table 5.10. In comparing the results with those obtained from the calibration errors (Table 5.9), the ranges of validation errors in traffic speed and density is slightly higher than that of the calibration errors by a factor of $[1, 1.1, 1.2, 1.1]$ and $[1.3, 1, 1.3, 1.3]$, respectively, for each modelling scenario. However, validation errors are characterised

by their higher variability. For instance, the standard deviation for the M60/CTM-basic has increased from 11 veh km^{-1} and 13 km h^{-1} to 18 veh km^{-1} and 35 km h^{-1}. The standard deviation of traffic speed and density errors increased by a factor of $[2.7, 1.9, 1.5, 1.6]$ and $[1.6, 1.4, 1.6, 1.6]$, respectively, for each modelling scenario. This illustrates that while the maximum ranges of calibration errors might have been reached, their variability are still lower than the validation ones. Although this is expected since the simulated parameters are not optimal on the other time periods, it can be a direct result of higher bivariate error points. Such differences, however, are compared in the context of their propagation to emission predictions in the next phases.

TABLE 5.10: summary statistics of validation errors in aggregated traffic density and speed, $[\varepsilon_r^v(\bar{\rho}), \varepsilon_r^v(\bar{u})]$, for each CTM modelling scenario

Summary Statistics	**M60 CTM-basic**	**M1 CTM-basic**	**M25 CTM-vsl**	**M25 CTM-basic**
Mean	[-7, 17]	[0, 7]	[0, -1]	[-2, 7]
Std. Dev.	[18, 35]	[18, 36]	[11, 29]	[11, 32]
Covariance	-590	-595	-281	-332
Minimum	[-86, -106]	[-67, -110]	[-55, -90]	[-56, -90]
Maximum	[79, 100]	[202, 105]	[38, 97]	[38, 97]
NR. of points	3679440	7570050	1587915	2097585

5.3.2 Second-order METANET

5.3.2.1 Model calibration

The estimation process of calibration and validation errors based on M-METANET is the same to those of the CTM. Since the arguments for using bivariate analysis and aggregated traffic density have been set earlier, the results of most interest here are the estimated bivariate traffic speed-density calibration and validation errors for each M-METANET modelling scenario.

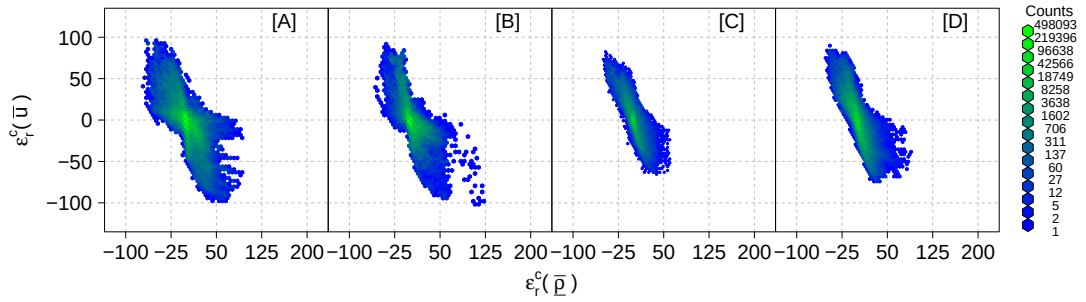


FIGURE 5.24: Bivariate calibration error distributions of aggregated traffic speed and density estimated from the *top-solutions* of each M-METANET modelling scenario.

Figure 5.24 presents the calibration error distributions obtained from the simulation of the *top-solutions* on relevant time periods for each modelling scenario, separately. Table 5.11 shows the range and distribution of the calibration errors obtained. While the ranges of calibration errors are very similar to those obtained for CTM modelling scenarios, their standard deviation are consistently lower, particularly for those of the traffic speed. For instance, a standard deviation of $[11, 11, 14, 17]$ km h^{-1} is obtained for M-METANET modelling scenarios in comparison to $[13, 19, 20, 20]$ km h^{-1} for the CTM ones. This can be a direct implication of the additional traffic speed dynamics equation used in the M-METANET model structure which reduced the variability in speed errors.

TABLE 5.11: summary statistics of calibration errors in aggregated traffic density and speed, $[\varepsilon_r^c(\bar{\rho}), \varepsilon_r^c(\bar{u})]$, for each M-METANET modelling scenario

Summary Statistic	**M60 M-METANET -basic**	**M1 M-METANET -basic**	**M25 M-METANET -vsl**	**M25 M-METANET -basic**
Mean	[1, -1]	[2, 0]	[0, 1]	[1, 0]
Std. Dev.	[8, 11]	[7, 11]	[7, 14]	[8, 17]
Covariance	-65	-52	-90	-114
Minimum	[-71, -98]	[-57, -102]	[-50, -65]	[-52, -74]
Maximum	[93, 96]	[120, 92]	[60, 82]	[89, 90]
NR. of points	2296302	962334	487017	930447

Similar to the CTM, summary statistics of calibration errors in traffic density and speed showing differences from one *top-solution* to another and from one time period to another are plotted in Figure 5.25 and Figure 5.26. The figures also show that most apparent differences are in the extreme values. This of course justifies the importance of using a single error distribution in order to account for such differences, but also to avoid bias in comparing a single arbitrary *top-solution* from CTM with a single arbitrary *top-solution* from M-METANET.

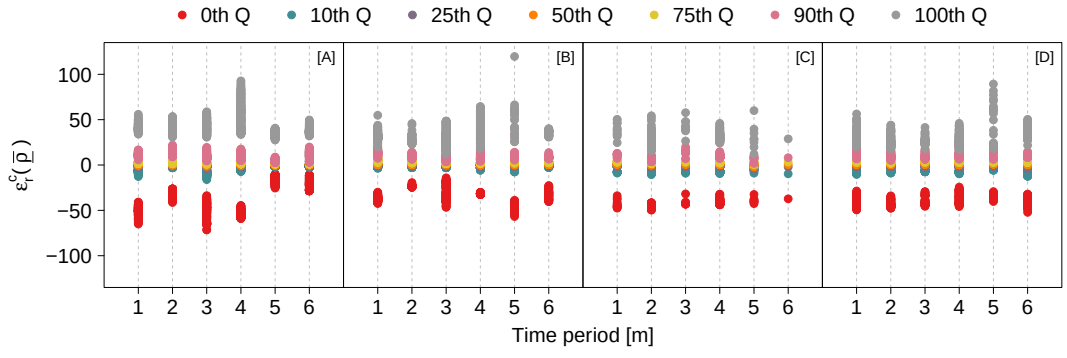


FIGURE 5.25: $[0, 10, 25, 50, 75, 90, 100]^{\text{th}}$ percentiles of calibration errors in aggregated traffic density for each *top-solution* and time period of each M-METANET modelling scenario.

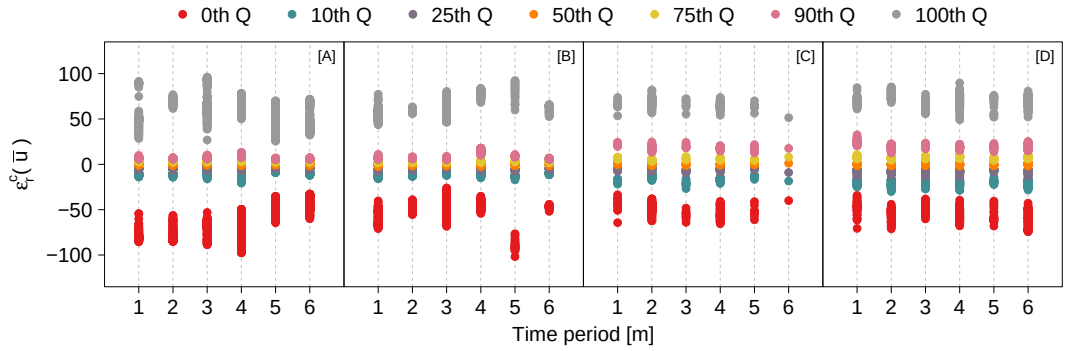


FIGURE 5.26: $[0, 10, 25, 50, 75, 90, 100]^{\text{th}}$ percentiles of calibration errors in aggregated traffic speed for each *top-solution* and time period of each M-METANET modelling scenario.

5.3.2.2 Model validation

As for the estimated M-METANET validation errors, a comparison to those obtained from the M-METANET calibration errors as well as from the CTM validation errors can be made. Figure 5.27 presents the validation error distributions obtained from the simulation of the *top-solutions* on relevant time periods for each modelling scenario, separately. Table 5.12 shows summary statistics of the validation errors obtained.

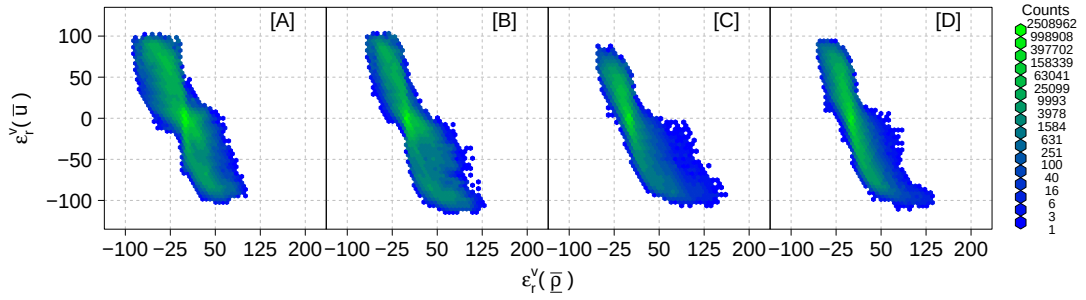


FIGURE 5.27: Bivariate validation error distributions of aggregated traffic speed and density estimated from the *top-solutions* of each M-METANET modelling scenario.

Comparison results are summarised in four points below.

1. The estimated range of M-METANET validation errors in traffic speed is higher than that of calibration errors by a factor of $[1.1, 1.2, 1.3, 1.3]$ for the four modelling scenarios, respectively.
2. The estimated range of M-METANET validation errors in traffic density is higher than that of calibration errors by a factor of $[1.1, 1.1, 1.9, 1.4]$ for the four modelling scenarios, respectively.
3. The standard deviation of M-METANET validation errors in traffic speed is higher than that of the calibration errors by a factor of $[2.4, 2.5, 1.5, 1.6]$ for the four modelling scenarios, respectively.

4. The standard deviation of M-METANET validation errors in traffic density is higher than that of the calibration errors by a factor of $[1.9, 2.1, 1.6, 1.5]$ for the four modelling scenarios, respectively.

Such result can directly be attributed to overfitted models on calibration data sets leading to higher errors values when calibrated parameters are used on validation data sets. However, the number of error points are highest for the validation outputs due to the high number of simulation runs across each modelling scenario. This is also a potential reason for the differences between calibration and validation error distributions. On average, however, there is not a substantial difference between the comparison of validation and calibration results of M-METANET to those obtained from validation and calibration of the CTM ones. Having said that, the variation of these errors across the region of predictions and the influence of such differences on emission predictions are investigated in the following phases.

TABLE 5.12: summary statistics of validation errors in aggregated traffic density and speed, $[\varepsilon_r^v(\bar{\rho}), \varepsilon_r^v(\bar{u})]$, for each M-METANET modelling scenario

Summary Statistic	**M60 M-METANET -basic**	**M1 M-METANET -basic**	**M25 M-METANET -vsl**	**M25 M-METANET -basic**
Mean	[-3, 5]	[1, 3]	[-2, 7]	[0, 5]
Std. Dev.	[15, 26]	[15, 27]	[11, 21]	[12, 28]
Covariance	-354	-360	-207	-314
Minimum	[-86, -102]	[-66, -114]	[-52, -106]	[-55, -111]
Maximum	[100, 104]	[128, 104]	[157, 87]	[136, 95]
NR. of points	11757990	5486670	3009645	4833675

5.4 Phase III: Grid-based Monte Carlo sampling

Given the prediction outputs from which the error distributions have been calculated in Phase II, the only remaining input to Phase III is the number of grids in predicted traffic density and speed, J and K, respectively. The choice of which is a trade-off between having low error data points per grid square and assigning error data points that can under or over-estimate their contribution to the grid square. In this study, a basic sensitivity analysis was performed, based on which appropriate values of J and K were determined. While a value of 15 for both variables was found to be appropriate, the influence of varying the grid-size of J and K on emission predictions is explored in § 5.5.3. Once the gridding criteria is known, the four steps to Phase III directly follow: bivariate gridding of predicted traffic data, evaluating bivariate error distributions, estimating the bivariate kernel density of errors, and sampling of bivariate errors for each grid square which are to be used in Phase IV. The results for each of the four CTM and M-METANET modelling scenarios are presented and discussed in § 5.4.1 and § 5.4.2, respectively.

5.4.1 First-order CTM

Before presenting the results of bivariate gridding, this section starts by empirically justifying the use of a grid-based approach. Figure 5.28A and Figure 5.28B show how errors in traffic density vary across density predictions and how errors in traffic speed vary across speed predictions given the *top-solutions* and time periods of M60/CTM-basic, respectively.

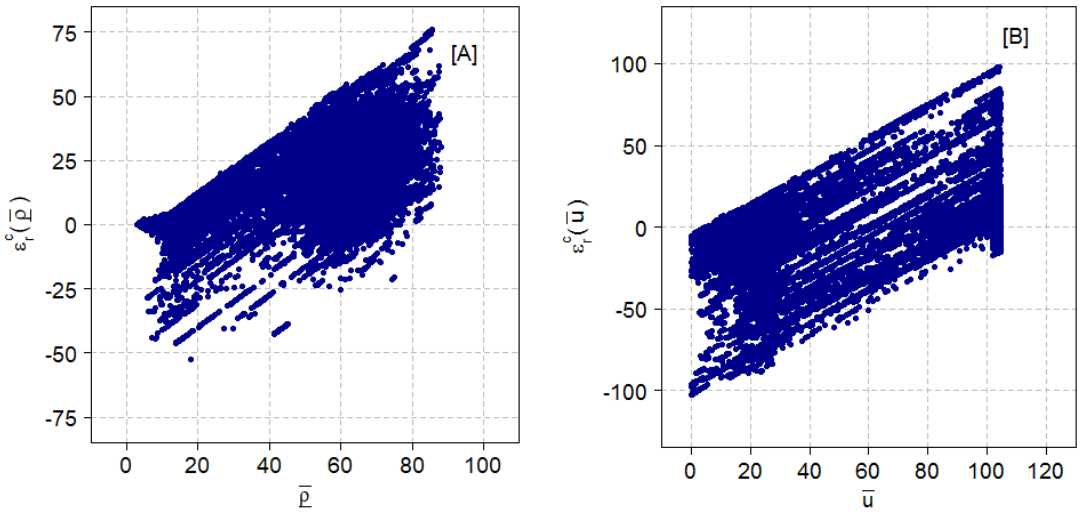


FIGURE 5.28: Scatter plots of predicted [A] traffic density and [B] speed versus their calibration errors obtained from the *top-solutions* and time periods of M60/CTM-basic modelling scenario.

Both figures illustrate that over (under)-predictions are associated with high traffic density (low speed) predictions; this is when the model fails to predict free-flow conditions. Similarly, under (over)-predictions are associated with low traffic density (high speed) predictions; this is when the model fails to predict congested conditions. Of course, these errors also have a joint distribution, as established earlier. Hence, it is inappropriate, for instance, to associate traffic speed-density error point $[-100, 50]$ with speed-density prediction point of $[100, 10]$. If traffic speed of 100 km h^{-1} is under-predicted by 100 km h^{-1}, then it means that the actual measurement is 200 km h^{-1}, a very high measurement. If traffic density of 10 veh km^{-1} is over-predicted by 50 veh km^{-1}, then it implies that the actual measurement is -40 veh km^{-1}, a negative-valued measurement. As such, it is inappropriate to associate the entire bivariate error distributions obtained from Phase II with all traffic speed-density predictions. Alternatively, associating an error distribution for each predicted traffic speed-density point implies very low data occupancy of errors for many prediction points. For instance, traffic speed-density prediction point $[40, 50]$ and $[88, 3]$ of M60/CTM-basic are associated with only 10 and 1 calibration error points, respectively. Therefore a grid-based approach to associating errors with the predicted traffic speed-density data is undertaken.

Figure 5.29 and Figure 5.30 show the gridding of predicted traffic speed and density based on the calibration and validation outputs, respectively. For instance, the first panel [A] of Figure 5.29 represents the calibration outputs of 392 *top-solutions* using the M60/CTM-basic modelling scenario. As noticed, these predictions do not always follow the equilibrium speed-density relationship of the parameter vectors since they are 1-minute aggregated outputs. For this modelling scenario, a total of 33 grid squares is obtained. A total of 33, 48, and 31 grid squares are obtained for M1/CTM-basic, M25/CTM-vsl, and M25/CTM-basic.

Similarly, the first panel [A] of Figure 5.30 represents the validation outputs of 392 *top-solutions* using the M60/CTM-basic modelling scenario. A total of 44 grid squares is obtained. A total of 35, 53, and 42 grid squares are obtained for the other three modelling scenarios, respectively.

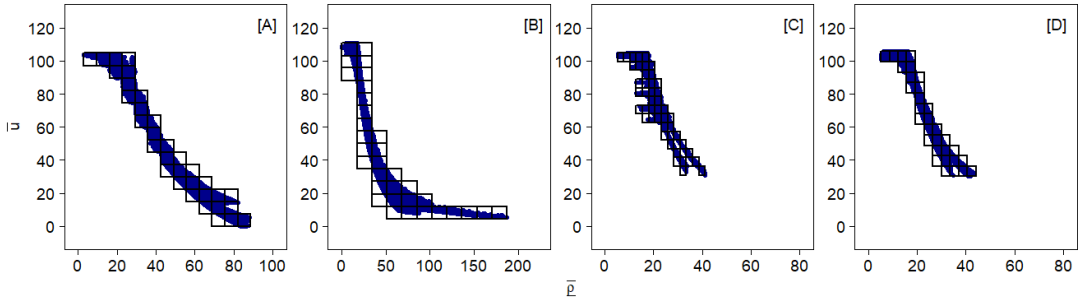


FIGURE 5.29: Aggregated traffic speed-density calibration outputs obtained from the *top-solutions* of each CTM modelling scenario. Black squares show the 15×15 grid squares. Grid squares with data points less than 100 are not considered/removed.

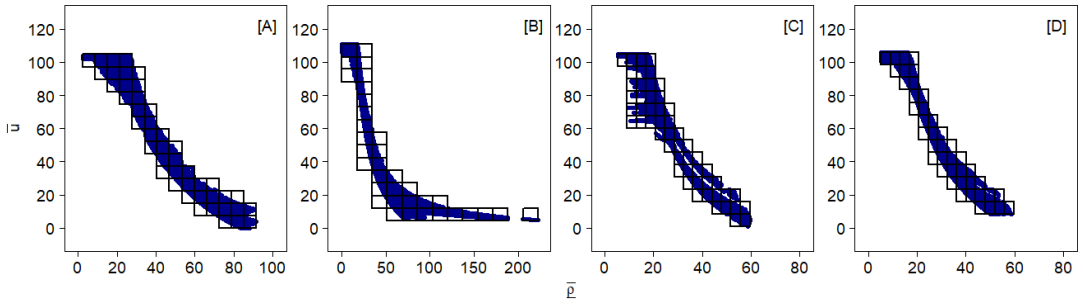


FIGURE 5.30: Aggregated traffic speed-density speed validation outputs obtained from the *top-solutions* of each CTM modelling scenario.

Based on the gridding of prediction points, a bivariate error distribution is evaluated and associated with each grid square. This is a subsetting exercise of two variables (of errors) based on the corresponding values of another two variables (of predictions). In order to better understand the variation of evaluated errors across the grid squares, Figure 5.31 and Figure 5.32 show summary statistics of evaluated calibration and validation errors of the M60/CTM-basic modelling scenario, respectively. For example, the mean and standard deviation of calibration traffic density errors are shown in

Figure 5.31A and Figure 5.31B, respectively. The mean and standard deviation of calibration speed errors are shown in Figure 5.31C and Figure 5.31D, respectively. Those obtained for the M1/CTM-basic, M25/CTM-vsl, and M25/CTM-basic are presented in Appendix C.

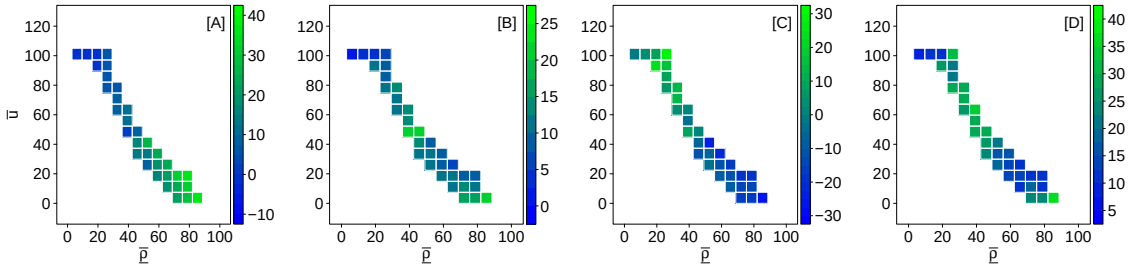


FIGURE 5.31: Summary statistics of calibration error distributions obtained for each grid square of the M60/CTM-basic modelling scenario: [A] mean and [B] standard deviation of traffic density errors and [C] mean and [D] standard deviation of speed errors.

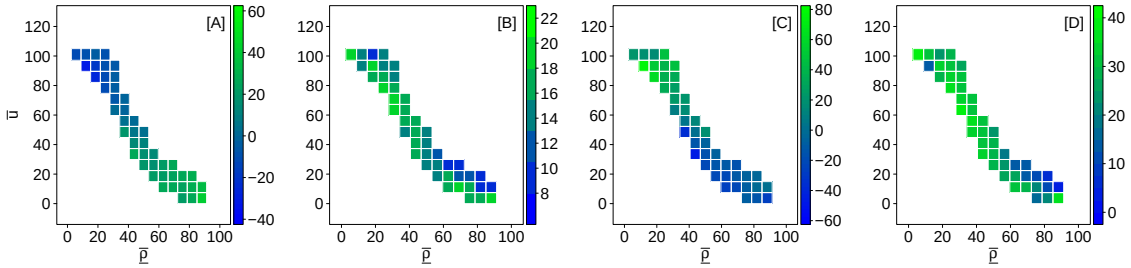


FIGURE 5.32: Summary statistics of validation error distributions obtained for each grid square of the M60/CTM-basic modelling scenario: [A] mean and [B] standard deviation of traffic density errors; and [C] mean and [D] standard deviation of speed errors.

While the actual values vary from one modelling scenario to another or from calibration to validation errors, summary statistics of errors, especially of the mean values, across the grid squares of predictions have similar characteristics. These are:

- mean values of traffic density errors are typically large negatives or blue coloured grids (under-predictions) at free-flow predictions while those for speed errors are large positives or green coloured grids (over-predictions);
- mean values of traffic density errors are typically large positives (over-predictions) at congested-condition predictions while those for speed errors are large negatives (under-predictions);
- mean values of traffic density (speed) errors change gradually from large negatives (positives) to large positives (negatives) across the grid squares of traffic speed-density predictions i.e. from free-flow predictions to congested-flow predictions; and
- standard deviation of both traffic density and speed errors depend on the modelling scenario and varies across the feasible prediction region.

- the range of mean values and standard deviation values per grid square is generally always higher for the validation errors in comparison to the calibration ones; this is consistent with results of the single bivariate error distribution obtained earlier in Phase II.

Bivariate gridding and error evaluation are only the first two steps of Phase III. The last two steps are the bivariate kernel density estimation of errors at each grid square and their random sampling. Given the large number of grid squares of calibration and validation errors across four CTM and four METANET modelling scenarios, an example is provided here, only to illustrate the method empirically.

First, consider grid square $[j = 6, k = 8]$ of calibration outputs of the M60/CTM-basic modelling scenario. This corresponds to $[36, 43]$ veh km^{-1} in traffic density and $[52, 60]$ km h^{-1} in speed, and is shown in Figure 5.33 below. The evaluated bivariate error distribution associated with any point o within the entire $[6, 8]$ grid square, is shown in Figure 5.34A. Traffic density and speed errors range between $[-43, 32]$ and $[-59, 53]$, respectively. The large range of errors is expected since the selected grid square is neither within free-flow nor within congested-flow predictions. The estimated bivariate kernel density, here assuming a bivariate normal kernel (Venables & Ripley, 2002), is shown in Figure 5.34B. The latter is used to randomly sample $s = 1000$ points to be propagated to emission predictions. A random sample for this grid is shown in Figure 5.34C.

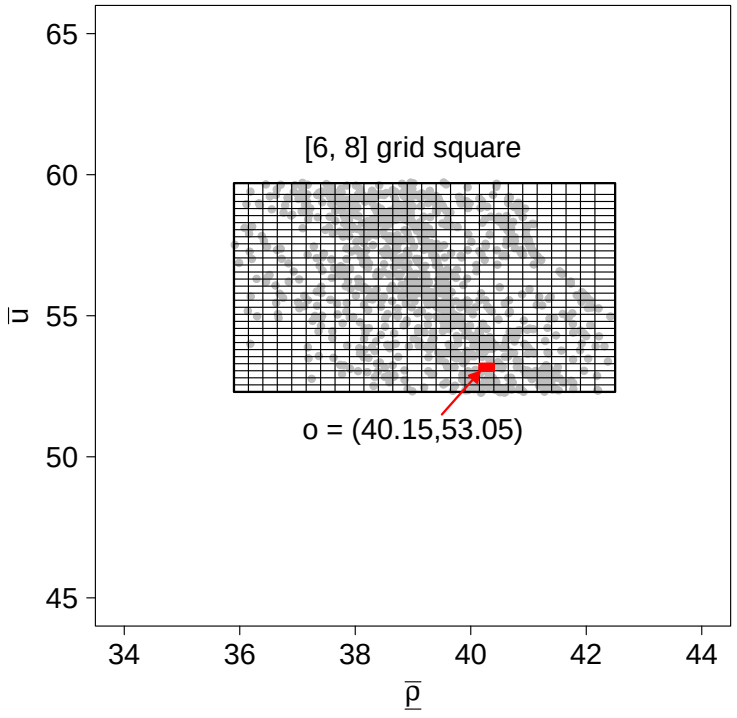


FIGURE 5.33: Example grid square of traffic speed-density calibration outputs of M60/CTM-basic. The grid is divided into finer grids of 1/4 resolution in each direction. Point $o = (40.14, 53.05)$ is used as an example in the next phase.

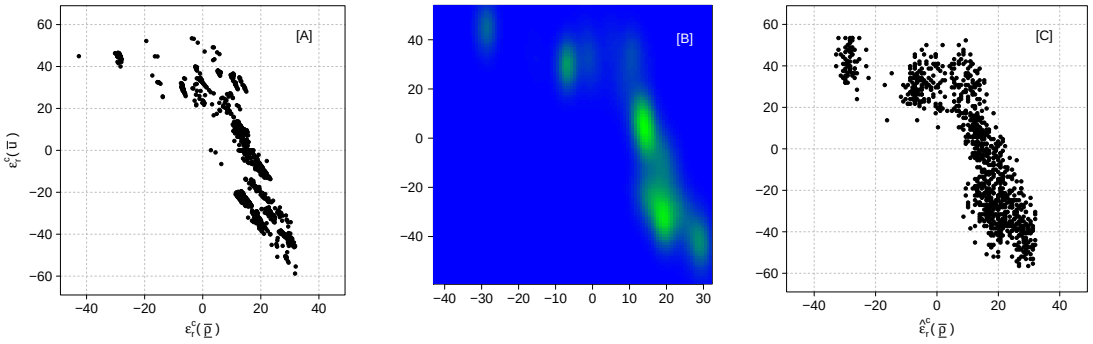


FIGURE 5.34: Example of the last three steps of Phase III: [A] evaluated traffic speed-density error points for grid square $[6, 8]$ of traffic speed-density calibration outputs of M60/CTM-basic; [B] corresponding estimated bivariate kernel density with blue-green indicating low-high density; and [C] random sample of 1000 error points from the estimated bivariate kernel density associated with each point in Figure 5.33.

This process is repeated for each grid square in each modelling scenario: four calibration and four validation CTM modelling scenarios, and four calibration and four validation M-METANET modelling scenarios. The main interest here is the process, and their results in terms of their use as input in the last 'uncertainty propagation phase, § 5.4.

5.4.2 Second-order METANET

Similar analysis for the M-METANET modelling scenarios is provided here, specifically for the first two bivariate gridding and error evaluation steps. Figure 5.35 and Figure 5.36 show the gridding of predicted traffic speed and density based on the calibration and validation outputs, respectively. For instance, the first panel [A] of Figure 5.35 represents the calibration outputs from the simulation of 1297 *top-solutions* of all time periods using the M60/M-METANET-basic modelling scenario. For this modelling scenario, a total of 61 grid squares is obtained. A total of 56, 41, and 41 grid squares are obtained for M1/M-METANET-basic, M25/M-METANET-vsl, and M25/M-METANET-basic, respectively. The first panel [A] of Figure 5.36 represents the validation outputs of the M60/M-METANET-basic. A total of 79 grid squares is obtained. A total of 64, 43, and 43 grid squares are obtained for the other three modelling scenarios, respectively.

In comparison to CTM-based predictions, the number of grid squares are higher i.e. the traffic speed-density prediction region is much larger. This is a result of the speed dynamics equation and the variability around the fundamental diagram induced by the M-METANET model structure. This is in addition to the higher number of *top-solutions* identified for each modelling scenario. The general trends in the mean and standard deviation of error values per grid square is similar to those described for

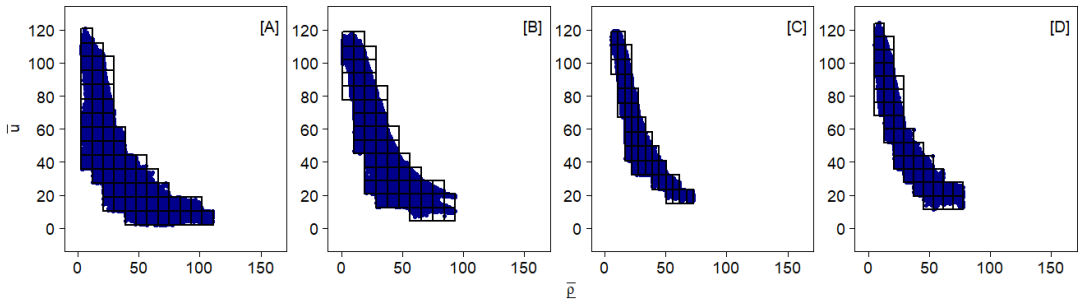


FIGURE 5.35: Aggregated traffic speed-density calibration outputs obtained from the *top-solutions* of each M-METANET modelling scenario.

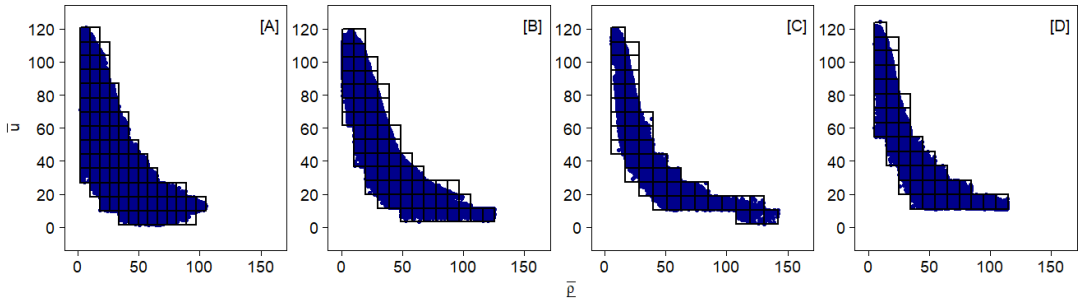


FIGURE 5.36: Aggregated traffic speed-density validation outputs obtained from the *top-solutions* of each M-METANET modelling scenario.

the CTM modelling scenarios. Figure 5.37 and Figure 5.38 show summary statistics of evaluated errors of the M60/M-METANET-basic modelling scenario, respectively. Appendix C shows the detailed trends of calibration and validation errors per grid square for the other three M-METANET modelling scenario.

Because of the differences in the feasible regions and number of grid squares between CTM and METANET, the best perceived way to examining the differences between the two is through the last 'uncertainty propagation' phase

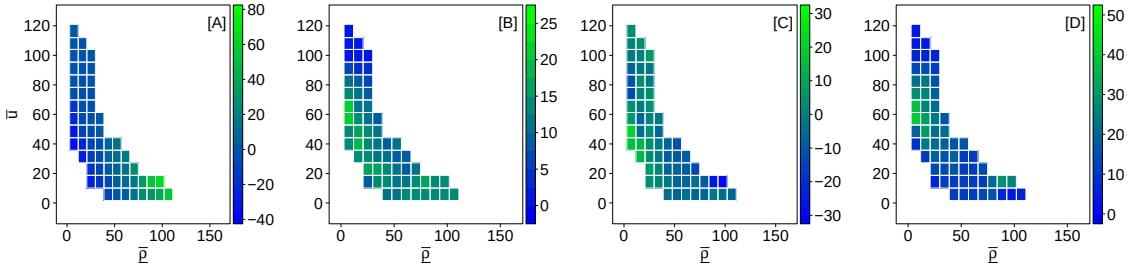


FIGURE 5.37: Summary statistics of calibration error distributions obtained for each grid square of the M60/M-METANET-basic modelling scenario: [A] mean and [B] standard deviation of traffic density errors and [C] mean and [D] standard deviation of traffic speed errors.

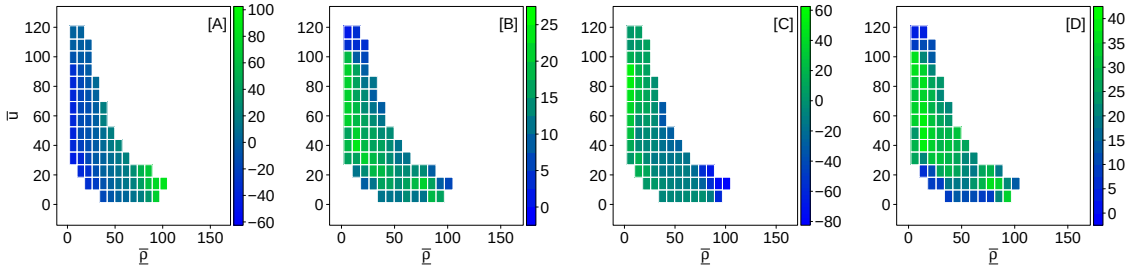


FIGURE 5.38: Summary statistics of validation error distributions obtained for each grid square of the M60/M-METANET-basic modelling scenario: [A] mean and [B] standard deviation of traffic density errors and [C] mean and [D] standard deviation of traffic speed errors.

5.5 Phase IV: Uncertainty propagation

With a random sample of bivariate errors for each grid square, the final phase focusses on their propagation to emission predictions. Here, an example of results is provided for NO_x air pollutant, i.e. $a = NO_x$. Keeping in mind, that any other air pollutant of interest can be used in this phase. The only parameter to this phase is the resolution of points o of each grid square. This mainly depends on the smoothness desired of the final surface maps; it does not affect the results since every point in the grid is assigned the same error distribution. Here, a resolution of 1/4 in each direction (i.e. traffic density and speed) is proposed.

Sections § 5.5.1 and § 5.5.2 present and discuss the results for the four CTM and METANET modelling scenarios, respectively. Section § 5.5.1 starts by providing an example of how errors are propagated for a single grid square, as a continuation to the previously introduced example of Phase III, § 5.4.1. Section § 5.5.3 comments on the influence of varying the grid size on the final outputs of confidence intervals.

5.5.1 CTM-based emission predictions

Given the sampled errors associated with each point o in grid square $[6, 8]$, consider a single point $o = (40.15, 53.05)$ in traffic density and speed, respectively. By adding the 1000-sampled errors to point o, Figure 5.39A shows the distribution of possible corresponding measured traffic speed-density points for the predicted point o i.e. the uncertainty in the predicted point o.

Given that, a point-based emission prediction uses the predicted point o as input. This prediction is shown as the red point in Figure 5.39B. A probabilistic-based emission prediction uses the distribution of possible corresponding measured points as input. This is shown as the distribution in Figure 5.39B which ranges between 0 and 25 grams of NO_x and includes the point-based emission prediction, 16 grams. The estimated 2.5% and 97.5% percentiles are 4 and 18 grams, respectively. The additive

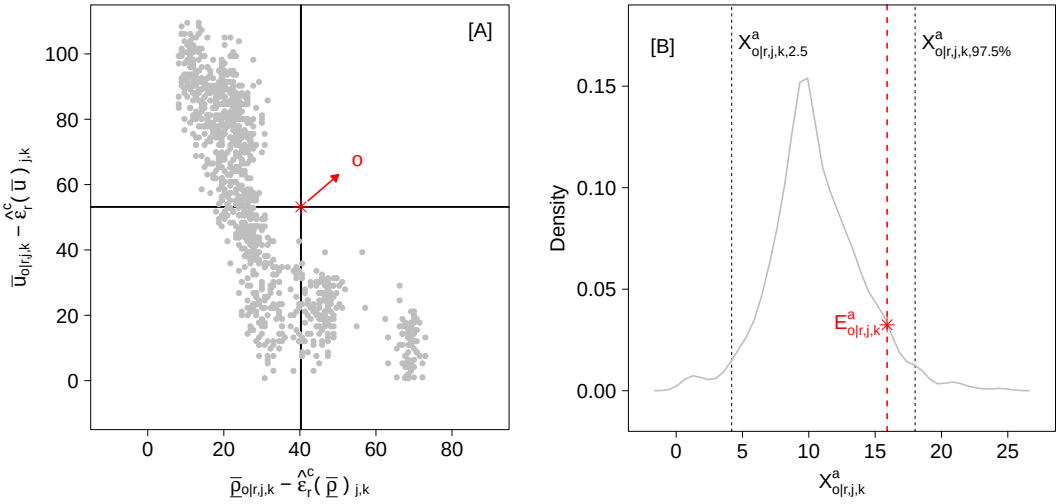


FIGURE 5.39: Example of uncertainty propagation at one prediction point in the feasible calibration output region of M60/CTM-basic: [A] shows the uncertainty in traffic speed-density point o of grid square $[6, 8]$, and [B] shows point-based NO_x emission prediction (red point/line) as well as the distribution reflecting its uncertainty. The 2.5^{th} lower and 97.5^{th} upper bounds are represented by the grey vertical dotted lines.

lower and upper bounds are thus -12 and 2 grams, respectively. This amounts to -75% ($\frac{-12}{16} \times 100$) and 12.5% ($\frac{2}{16} \times 100$) of the point-based prediction, respectively. Predicted NO_x emissions here are in grams produced on a one kilometre-lane during one minute; this is in order to compare between the results from different modelling scenarios.

This is repeated for each point o in the selected grid square and for each grid square in a modelling scenario, in order to obtain a probabilistic-based emission prediction surface throughout the feasible prediction region. These surfaces are presented in Figure 5.40 and Figure 5.41 based on calibration and validation errors, respectively. For example, Figure 5.40A shows the point-based emission prediction for the feasible prediction region of M60/CTM-basic consisting of 33 grid squares [1A] as well as their upper bounds [2A] and lower bounds [3A]. Figure 5.40B, Figure 5.40C, and Figure 5.40D show those for M1/CTM-basic, M25/CTM-vsl, and M25/CTM-basic.

Here it is important to note that COPERT recommends the use of emission factors' functions for average speeds higher than 5 or 10 km h^{-1} (depending on the vehicle category). Because of that, the percentages reported next cut these out of the feasible prediction region. However, even when removing these, there is still a risk of emissions being calculated below these bounds when adding the sampled errors to original prediction points. Another issue observed at the boundaries of the feasible region is the possibility of getting negative values of traffic speed or density when adding the sampled errors (i.e. before calculating emissions). In such rare circumstances, they are

set to be bounded by 1 km h^{-1} or 1 veh km^{-1}. Of course, the lower the number of grid squares, the more these negative values would appear, and vice versa.

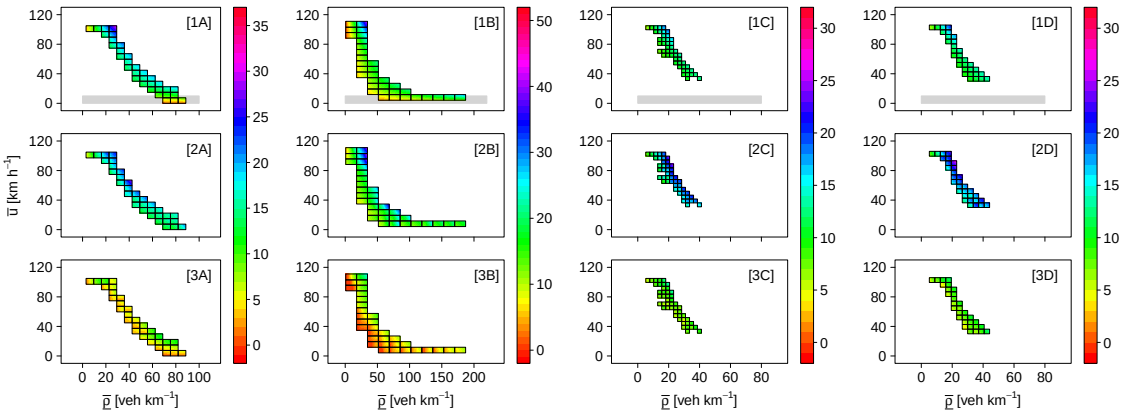


FIGURE 5.40: Coloured contours showing NO_x emission predictions in grams produced in one kilometre space and one minute time as a function of modelled traffic speed and density using COPERT emission factors over the feasible calibration output region of [A] M60/CTM-basic, [B] M1/CTM-basic, [C] M25/CTM-vsl, and [D] M25/CTM-basic. Black lines show the grid squares of each modelling scenario. [1] shows point-based emission predictions. [2] and [3] show the upper and lower bounds of emission predictions at 95% confidence level, respectively. Dark grey rectangles in [1] represent uncertain emission predictions due to COPERT being restricted to > 5 km h^{-1} or > 10 km h^{-1}.

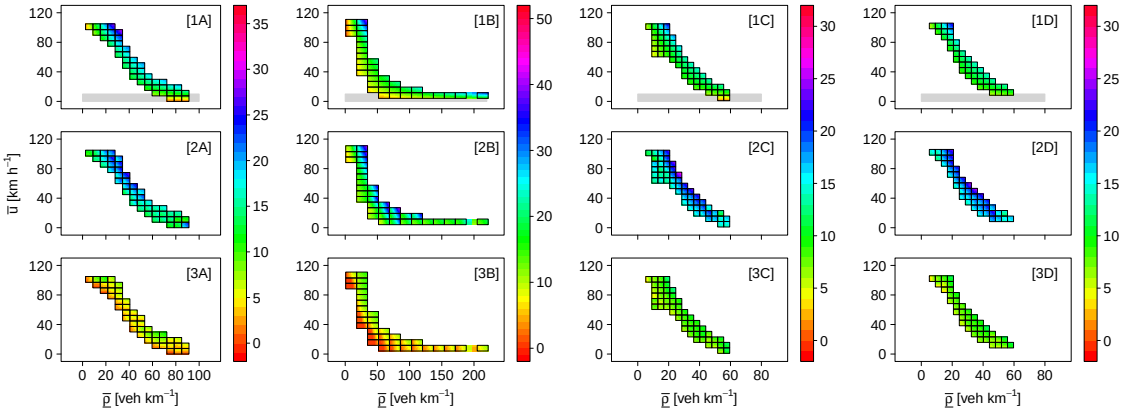


FIGURE 5.41: Coloured contours showing NO_x emission predictions in grams over the feasible validation output region of each CTM modelling scenario.

Before analysing the results of Figure 5.40 and Figure 5.41, it is also essential to reiterate on the main limitation of using the obtained prediction errors in this phase. Prediction errors are considered a result of uncertainty in the traffic flow model outputs. However, uncertainty in COPERT emission predictions using CTM or METANET arise from other sources such as the emission factors (and their parameters) but more importantly from the compatibility of uni-class traffic flow models with multi-class emission models. Deployed traffic flow models in this study use a single fundamental diagram for all traffic modes (e.g. passenger cars versus heavy duty vehicles) and

thus outputs average traffic speed for all vehicles on a road network. This results in uncertainty in traffic speeds (and density) used to predicting emissions as well as uncertainty in the fleet mix proportions used to predict total emissions. While this source can partly be reduced using multi-class traffic flow models (i.e. multi-class CTM or METANET), acquiring multi-class traffic density and speed data in order to optimise the traffic flow models and estimate their prediction errors becomes more challenging.

By calculating the additive upper and lower bounds based on these surfaces, the following can be observed for the four CTM modelling scenarios and their calibration surfaces, noting that all statistics are based on emission predictions in grams produced on one kilometre-lane during one minute:

- the range of additive upper bounds $[-5, 18]$, $[-10, 20]$, $[1, 9]$, and $[1, 10]$ grams for each modelling scenario, respectively;
- the mean of additive upper bounds are 2, 4, 5, and 5 grams;
- upper bounds are on average 8%, 24%, 39%, and 41% higher than their corresponding point-based emission predictions;
- the range of additive lower bounds $[-23, 1]$, $[-19, 1]$, $[-7, 0]$, and $[-9, -1]$ grams;
- the mean of additive lower bounds are -8.8, -7.3, -3.8, and -4.4 grams; and
- lower bounds are on average 60%, 57%, 32%, and 36% lower than their corresponding point-based emission predictions.

The following can be observed for the four CTM modelling scenarios and their validation surfaces:

- the range of additive upper bounds $[-6, 21]$, $[-16, 23]$, $[-1, 13]$, and $[1, 12]$ grams for each modelling scenario, respectively;
- the mean of additive upper bounds are 2, 5, 6, and 6 grams;
- upper bounds are on average 15%, 42%, 57%, and 55% higher than their corresponding point-based emission predictions;
- the range of additive lower bounds $[-20, 1]$, $[-25, 2]$, $[-9, 9]$, and $[-11, 2]$ grams;
- the mean of additive lower bounds are -10, -8, -3, and -4 grams; and
- lower bounds are on average 67%, 57%, 25%, and 32% lower than their corresponding point-based emission predictions.

The above results show the complexity of uncertainty propagation when comparing those across calibration and validation outputs, and across VSL and non-VSL operated road networks. Generally, the results are road network-specific (or even modelling scenario-specific), and are dependent on each grid-specific sampled error points. Such complexity is a result of the potential interactions with the emission model and its embedded emission factors; examples of which were provided in Figure 3.6. Each prediction of emission rely not only on the traffic speed-density point, but also on the proportion of vehicles at each point and their corresponding emission factors.

In addition to that, negative additive upper or positive additive lower bounds show that point-based emission prediction is outside the estimated bounds. This is also a direct result of the non-linearity of the emission model; uncertainty in a traffic speed-density point does not map linearly to uncertainty in emission predictions. The main two conclusions from these results are listed below.

1. Percentage differences of upper and lower bounds with respect to the point-based emission prediction are almost always slightly higher given the validation outputs in comparison to the calibration outputs. One possible reason why it is not 'always' higher despite the fact that standard deviation of validation errors are higher than the calibration ones, is a result of the complex interactions of the results of each grid with the emission model.
2. Differences between M25/CTM-vsl and M25/CTM-basic are negligible, which is expected due to the similarities in their error distributions discussed earlier.

5.5.2 METANET-based emission predictions

The same method is used to calculate the surface maps given the calibration and validation errors of the four M-METANET modelling scenarios. These surfaces are presented in Figure 5.42 and Figure 5.43. For example, Figure 5.42A shows the point-based emission prediction for the feasible prediction region of M60/M-METANET-basic consisting of 61 grid squares [1A] as well as their upper bounds [2A] and lower bounds [3A]. Figure 5.42B, Figure 5.42C, and Figure 5.42D show those for M1/M-METANET-basic, M25/M-METANET-vsl, and M25/M-METANET-basic.

In calculating the additive upper and lower bounds based on these surfaces, the following can be observed for the four M-METANET modelling scenarios and their calibration surfaces:

- the range of additive upper bounds $[-3, 14]$, $[-6, 12]$, $[-1, 9]$, and $[-5, 11]$ grams for each modelling scenario, respectively;
- the mean of additive upper bounds are 5, 3, 3, and 4 grams;
- upper bounds are on average 91%, 35%, 31%, and 38% higher than their corresponding point-based emission predictions;
- the range of additive lower bounds $[-15, 4]$, $[-11, 2]$, $[-11, -2]$, and $[-13, 2]$ grams;
- the mean of additive lower bounds are -4, -4, -4, and -5 grams; and
- lower bounds are on average 43%, 38%, 38%, and 40% lower than their corresponding point-based emission predictions.

The following can be observed for the four M-METANET modelling scenarios and their validation surfaces:

- the range of additive upper bounds $[-4, 17]$, $[-8, 17]$, $[-9, 20]$, and $[-5, 17]$ grams for each modelling scenario, respectively;

- the mean of additive upper bounds are 6, 5, 5, and 5 grams;
- upper bounds are on average 115%, 84%, 51%, and 53% higher than their corresponding point-based emission predictions;
- the range of additive lower bounds $[-15, 2]$, $[-18, 4]$, $[-24, 4]$, and $[-23, 2]$ grams;
- the mean of additive lower bounds are -5, -5, -6, and -6 grams; and
- lower bounds are on average 49%, 48%, 41%, and 47% lower than their corresponding point-based emission predictions.

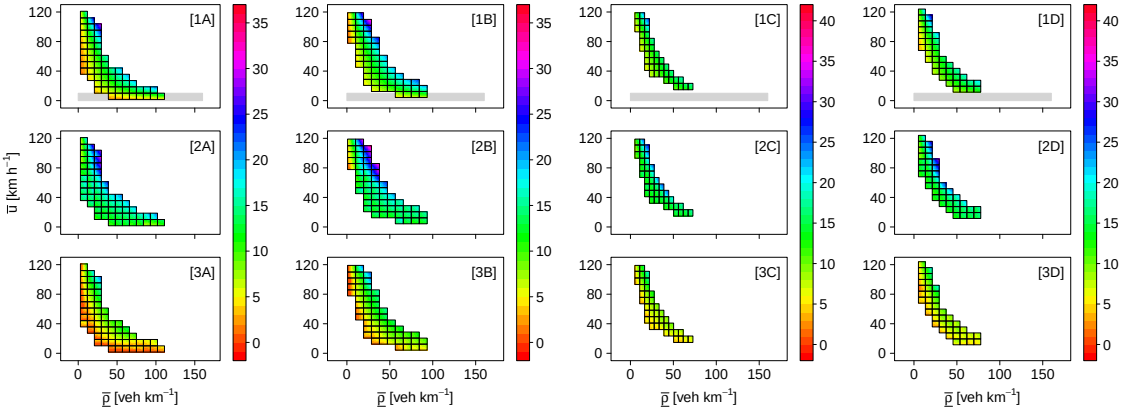


FIGURE 5.42: Coloured contours showing NO_x emission predictions in grams over the feasible calibration output region of each M-METANET modelling scenario.

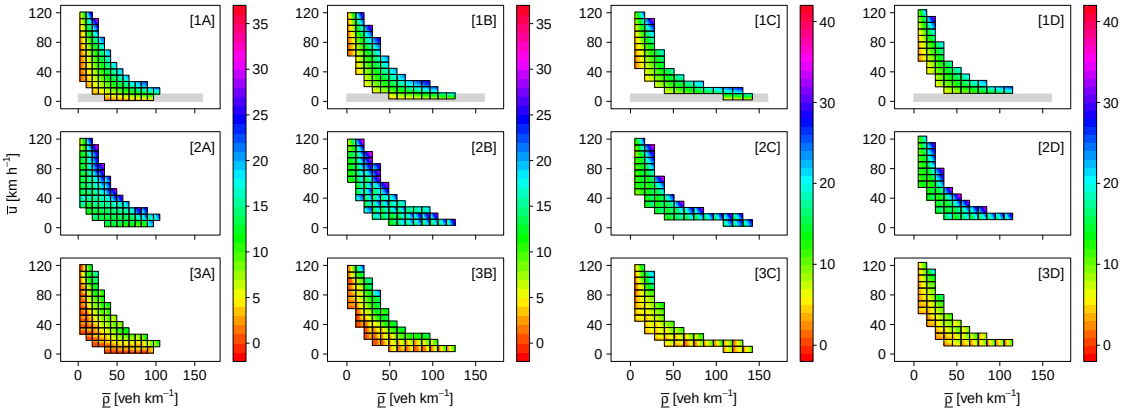


FIGURE 5.43: Coloured contours showing NO_x emission predictions in grams over the feasible validation output region of each M-METANET modelling scenario.

Similar to the CTM results, M-METANET results show that the percentage differences are always slightly higher given the validation errors in comparison to the calibration ones. Also, differences in using M25/M-METANET-vsl and M25/M-METANET-basic are negligible.

In order to check the influences of such differences given the prediction outputs of a single model (i.e. single *top-solution* and time period), an example is provided in the last section to make the final conclusion.

5.5.3 A note on grid-size parameters

To study the influence of the number of grids on the results, two scenarios were devised and compared with the initial scenario of $[J = 15, K = 15]$. These are a lower number of grids with $[J = 10, K = 10]$ and higher number of grids with $[J = 20, K = 20]$. Phase III and Phase IV were repeated given these two scenarios for the M60/CTM-basic and M60/M-METANET-basic only. The final additive upper and lower bounds of the point-based emission predictions were obtained. Figure 5.44 shows the final cumulative distributions of additive upper and lower bounds for the two scenarios in comparison with the initial scenario.

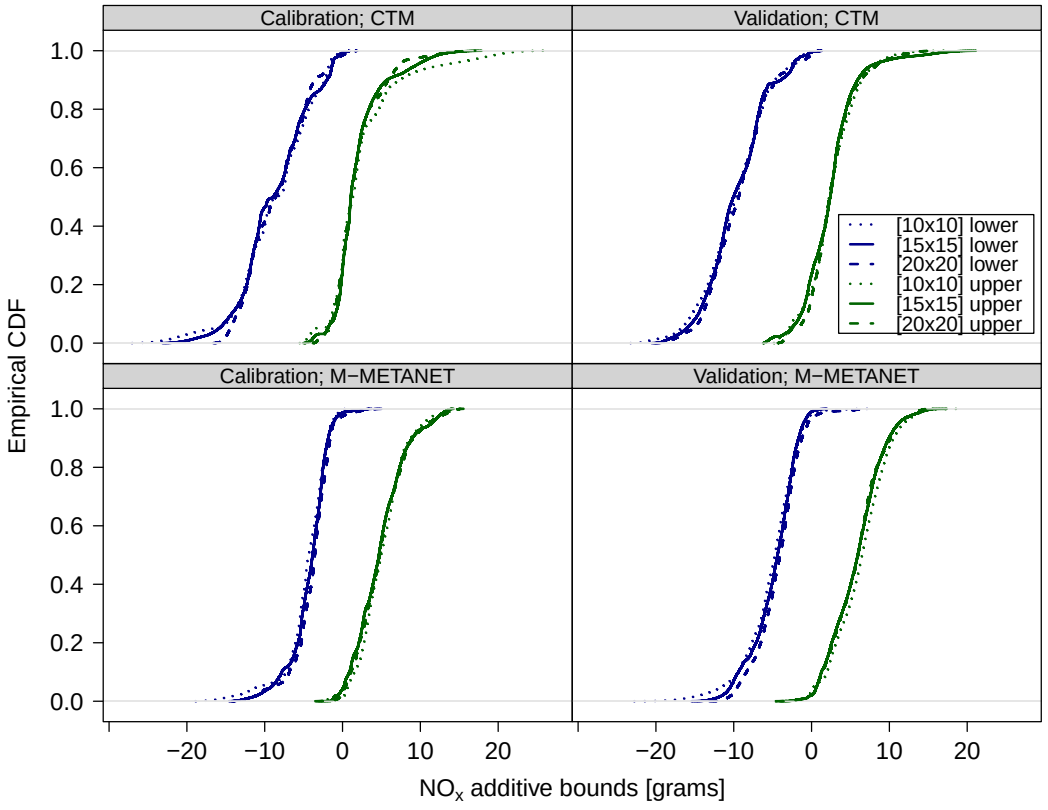


FIGURE 5.44: Cumulative distributions of additive upper and lower bounds of point-based NO_x emission predictions for the M60/CTM-basic and M60/M-METANET-basic modelling scenarios based on three scenarios of the number of grids.

The analysis shows that while the distribution of additive upper bounds on errors does not change significantly based on the three gridding choices, the range can increase for lower grid-number, and reduce with an increased grid-number up to a point at which the range of additive upper bounds ceases to decrease. This is especially the case for the calibration outputs of CTM where the lowest number of points is obtained in comparison to the validation outputs or to those of M-METANET. Generally, as the choice of the grid-number increases, the number of points within each grid square decreases, especially at the boundaries of the feasible region. As the choice of the

grid-number decreases, there is a risk of over-estimating the bounds especially if the population of points used for the grid-based approach is relatively low, as is the case of the calibration outputs of M60/CTM-basic.

5.6 Probabilistic-based emissions: Final example

In principle, the results of the methodological framework can be used to determine probabilistic-based emissions at different levels of spatiotemporal resolution, described as follows.

- The most aggregate level is predicting total emissions on an entire road network and during the entire time period alongside the upper and lower bounds.
- The least aggregate is predicting emissions on each cell and time step alongside their upper and lower bounds. However, since the errors are based on the spatiotemporal resolution of traffic measurements, then the least aggregate is doing so at such spatiotemporal resolution.
- In between these two, one can predict the total emissions either on each cell alongside the upper and lower bounds while keeping the high temporal resolution of predictions, or for the entire time period while keeping the high spatial resolution of predictions.

Given the predictions surfaces of emissions provided earlier and their upper and lower bounds, the following describes major steps to determining probabilistic-based emission predictions. Given a single *top-solution* and its predicted traffic density and speed calibration or validation output,

- For each traffic speed and density per lane point, the corresponding prediction of emissions can be easily extracted alongside its upper and lower bounds.
- Each of the three values are multiplied by their corresponding number of lanes and distance (of two measurement sites). This is made in order to transform grams produced on a one kilometre-lane to grams produced on the relevant section. Values can be divided by 1000 to convert emission predictions to kilograms.

By keeping the spatiotemporal resolution of aggregated traffic speed and density points, this will provide the least aggregate output of probabilistic-based emissions of a single traffic flow model. These, however, can be summed up in space (or in time) to produce time series data (or spatial data), as well as in both space and time to produce total emissions on a road network during an entire time period.

Here an example of total emissions and their upper and lower bounds for a single time period is provided. These are time period [6] for the M60 and time period [2] for M1 and M25; all of which are 3-hour time periods. Their *top-solutions* allow to calculate confidence intervals, given the calibration outputs. The *top-solutions* of all other time periods allow to calculate confidence intervals, given the validation outputs.

Figure 5.45 below shows the results for the four CTM and four M-METANET modelling scenarios. Green coloured points represent the total emission predictions in kilograms during the 3-hours, given each *top-solution*. Red coloured points represent their upper and lower bounds. In order to make the results comparable, the total emissions for each road network were divided by the total length of each road network. However, while both the M60 and M1 are three lanes road networks, those of the M25 are either 4 or 5; the main reason behind the higher total emissions. The following has been observed from the example results.

- additive bound on total emission predictions can reach 28% and 48% higher for calibration and validation CTM models, respectively;
- additive bound on total emission predictions can reach 37% and 55% lower for calibration and validation CTM models, respectively;
- additive bound on total emission predictions can reach 24% and 50% higher for calibration and validation M-METANET models, respectively; and
- additive bound on total emission predictions can reach 33% and 49% lower for calibration and validation M-METANET models, respectively.

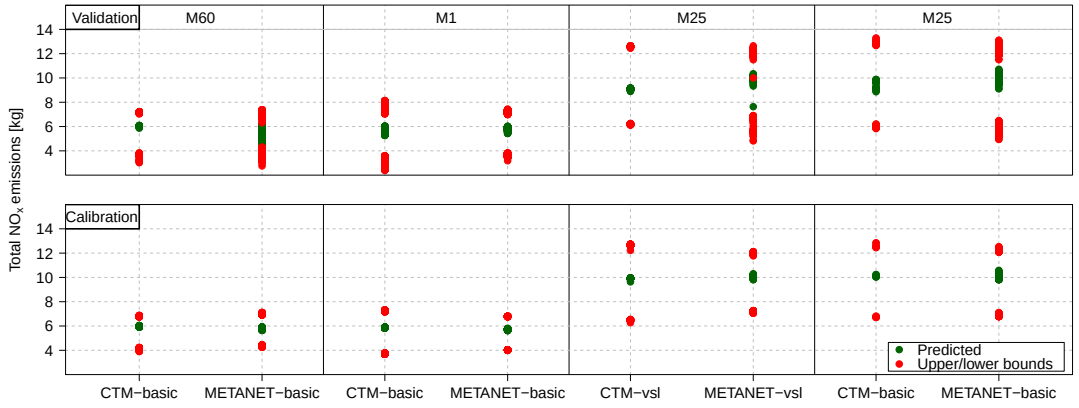


FIGURE 5.45: Total NO_x emission predictions [kg] alongside their upper and lower bounds on each road network during a 3-hour time period, given the calibration and validation CTM and M-METANET outputs of a single time period. Time period [6] is selected for M60, and time period [2] is selected for M1 and M25. Total emissions are normalised by the road network length Table 4.1.

5.7 Summary

This chapter achieves the second objective of this research which is the application of the methodological framework on multiple road networks and different modelling scenarios in order to ensure transferability of results, but also to identify differences or similarities in the probabilistic-based emission prediction results.

The main differences between CTM and M-METANET starts in the results of the ensemble-based optimisation phase. The following were the main conclusions from the application of the first phase.

- The number of *top-solutions* obtained for each modelling scenario reflects the complexity of the optimisation problem.
- Uncertainty in the modelling structure dominates when CTM is applied whereas uncertainty in parameter vector dominates when M-METANET is applied.
- Uncertainty in the modelling structure dominates when CTM and M-METANET are applied on a VSL-operated road network, whereas uncertainty in the parameter vectors dominates when they are applied on a non-VSL road network.
- The validation of *top-solutions* for congested time periods on less congested ones performs better than the validation of those obtained for less congested time periods, reflecting the higher uncertainty in supply parameters in comparison to the demand parameters.

Using the ensemble of *top-solutions* across the different time periods of each road network, estimated calibration prediction errors were observed to always be lower (or at least equal) to the validation prediction errors, given both the CTM and M-METANET modelling scenarios. In addition, there was no substantial difference between the error distributions obtained when using the VSL-model structure on a VSL road network and those obtained when using the basic structure on a VSL road network. This reflects the complexity and the interaction between the impact of VSL on traffic dynamics and the optimisation problem. This of course prevents a better understanding of the impact of VSL in the context of this work.

Such differences or similarities were finally investigated in the context of their influence on emission predictions, using a a grid-based Monte Carlo simulation approach. The methodology was explained empirically in details. Final results of confidence intervals obtained for each traffic speed-density point in the feasible region of each CTM and M-METANET modelling scenarios were presented. The main limitation found in the estimation of such surfaces were at the boundaries of the feasible region. This is where possibly very high (or negative) traffic speed and density points can be obtained upon the addition of errors. This suggests that a varying grid square size across the feasible region might improve upon the developed grid-based phase; a possible future work to refining the developed methodology. Nevertheless, the following can be concluded in terms of the confidence intervals obtained.

- Confidence intervals when using VSL model structure on VSL-operated road networks are similar to those obtained when using the basic model structure (CTM or M-METANET).
- Regardless of the CTM modelling scenario used, additive bounds can on average be 41% (60%) higher (lower) their corresponding point-based emission predictions, given calibration outputs and errors.

- Regardless of the CTM modelling scenario used, additive bounds can on average be 57% (67%) higher (lower) their corresponding point-based emission predictions, given validation outputs and errors.
- Regardless of the M-METANET modelling scenario used, additive bounds can on average be 91% (43%) higher (lower) their corresponding point-based emission predictions, given calibration outputs and errors.
- Regardless of the M-METANET modelling scenario used, additive bounds can on average be 115% (49%) higher (lower) their corresponding point-based emission predictions, given validation outputs and errors.

However, these upper and lower bounds vary from modelling scenario to modelling scenario, depending on the grid-based sampled errors and their interactions with the emission model. Of course, such interactions are quite complex given the non-linearity of vehicle-category specific emission factors' functions embedded in the emission model.

The developed surfaces are finally used to calculate total emission predictions over a particular time period on each of the three road networks, given the *top-solutions* obtained from Phase I. The last section illustrated how these surfaces can be put to use to calculating confidence intervals at different levels of detail, mainly: total emissions across the entire road network and during the entire time periods; total emissions across the entire road network at each time step; and total emissions during the entire time period at each measurement site.

Chapter 6

Conclusion

In Chapter 1, the aim of this research was set to quantify uncertainty in average speed-based emission predictions given the uncertainty in the predictions of macroscopic traffic flow models. To achieve this aim, two main objectives were proposed as follows.

1. To develop a data-driven methodological framework that enables quantifying and propagating uncertainty in the outputs of macroscopic traffic flow models to emission predictions.
2. To apply the developed methodological framework on multiple road network in order to ensure transferability of the framework as well as to allow rigorous quantitative comparison of results amonsgst the utilised macroscopic traffic flow models.

Accomplishing these two objectives has led to four main contributions to current knowledge. These were realised based on the identified gaps of the literature review of Chapter 2. While the outcomes of this research have been discussed at length in the final sections of Chapters 3, 4, and 5, this chapter starts with § 6.1 by providing only an overview of findings with particular focus on the four original contributions and the accomplished objectives. The chapter then discusses in § 6.2 the limitations of this research associated with the development and application of the methodological framework. The chapter finally ends with suggestions of potentially productive future research in § 6.3, which can be built based on the contributions of this research.

6.1 Overview of research findings

The literature review of Chapter 2 has identified four major issues on the topic of predicting emissions using macroscopic traffic flow models. Two of which has led to the first two major contributions of this research and ultimately to accomplishing the first objective. These two issues are summarised below.

- The first issue was identified in § 2.2. While a number of studies on the parameter optimisation problem of macroscopic traffic flow models were found, the selection

criteria of a 'good' solution is often based on subjective judgment or single-valued quantitative assessment. Also, most of these studies appreciate the complexity of the optimisation problem along with the fact that multiple 'good' solutions can exist, and typically vary depending on the time period selected. Yet, a single 'good' solution is often selected and reported, disregarding the landscape of all potential 'good' solutions.

- The second issue was identified in § 2.3, § 2.4, and § 2.5. While uncertainty arising from emission factors (major component of emission models) has been studied, uncertainty propagating from macroscopic traffic flow model outputs (another major component of emission models) has not been investigated before, despite their common use to predicting emissions and the complexity of the sources of uncertainty involved.

The above illustrated the need for a methodology which can quantify uncertainty in emission predictions given the uncertainty inherent in the outputs of macroscopic traffic flow models, but also the need to resolving the issues surrounding the optimisation problem of macroscopic traffic flow models. The latter helps to better quantify the uncertainty in their outputs, which is necessary for their propagation to emission predictions. As a result, a five-phase methodological framework contributing to existing knowledge in two major aspects, listed below, has been devised.

1. A novel data-driven ensemble-based optimisation approach has been proposed, which includes the selection of 'good' solutions (also referred to as *top-solutions*) objectively using a statistical testing method. The approach takes into account the landscape of all possible *top-solutions* i.e. it considers uncertainty around the choice of a solution for a particular time period. The approach also takes into account differences in the landscape of *top-solutions* across different time periods i.e. it considers uncertainty around the choice of a typical day of traffic flow to model. The availability of historical data that matches the requirements of the ensemble-based optimisation approach has indeed facilitated the application of such approach. It also facilitated the estimation and comparison of calibration and validation prediction errors that are considered the total accumulated uncertainty in the outputs of calibration or validation macroscopic traffic flow models. This ultimately led to the first three phases of the methodological framework: the data preparation preliminary phase, the ensemble-based optimisation phase, and the prediction error estimation phase.
2. A novel approach to propagating prediction errors to average speed-based emission predictions has been proposed, which involves both: the use of bivariate analysis of traffic speed-density prediction outputs and prediction errors; and the use of a grid-based Monte Carlo simulation approach to uncertainty propagation. The population of points resulting from the third prediction error estimation phase allows to better understand the relationship between the bivariate traffic speed-density errors and the bivariate traffic speed-density outputs. As a result,

the problem was not restricted to estimating the confidence intervals of emissions resulting from a single traffic flow model simulation, but rather to estimating the confidence intervals (upper and lower bounds) of emissions at each potential traffic speed-density point of a road network. These are referred to as surface maps of emission predictions and their confidence intervals over the feasible prediction region of a road network. These can ultimately be used to construct confidence intervals given the outputs of any single traffic flow model simulation (i.e. based on any *top-solution* and time period). To do so, two final methodological phases were devised: the Grid-based Monte Carlo simulation phase, and the uncertainty propagation phase.

Chapter 3 of this thesis was dedicated to the development of each phase of the methodology. Since it is not restricted to a specific choice of a macroscopic traffic flow model or average speed-based emission model (i.e. was intentionally built as a generic methodology), it allows:

- to examine and compare the influence of first-order and second-order traffic flow models on the confidence intervals of emission predictions using any average speed-based emission model. In this research, particular focus was on CTM and METANET due to their common use in many applications, specifically in predicting emissions.
- to examine the influence of variations (or modifications) of first-order CTM and second-order METANET. In this research, particular focus was on VSL-modified traffic flow models.

Both of the above were found to be gaps in the literature review of § 2.2. While the quantitative comparison of CTM and METANET was undertaken in terms of their parameter optimisation problem, studies have not been found to compare their influence on emission predictions. This is despite the fact that both CTM and METANET have been integrated with, for example, average speed-based emission models. Also, as a result of the widespread implementation of VSL on motorway road networks, studies were found to have modified the initial structures of CTM and METANET models to incorporate the impact of VSL on traffic dynamics; such impacts are typically studied outside the context of traffic flow models, as described in § 3.2. Nevertheless, such modified model structures were not found to be assessed using real traffic data, as noted in § 2.2.2, and not used in the context of studying uncertainty in emission predictions. These two gaps has led to the other two contributions of this research, and ultimately to accomplishing the second objective.

Chapter 4 and Chapter 5 contributed to these two issues through the application of the developed methodological framework on three motorway road networks, one of which operating of VSL. Chapter 4 focussed entirely on the data preparation preliminary phase of the developed methodology, whereas Chapter 5 focussed on developing surface maps of emission predictions and their confidence intervals, given eight different modelling scenarios:

- two CTM-basic and two METANET-basic modelling scenarios on two non-VSL motorway road networks; and
- one CTM-vsl and one METANET-vsl modelling scenarios on a VSL-operated motorway road network. These were compared to CTM-basic and METANET-basic modelling scenarios on the same road network.

The results of Chapter 5 highlighted the complexity and interactions in the sources of uncertainty in macroscopic traffic flow models, primarily in their structure and their parameter vectors. This was illustrated in the outputs of § 5.2 and § 5.3. Chapter 5 also highlighted the complexity and interactions in the propagation of uncertainties and errors from the outputs of macroscopic traffic flow models to emission predictions due to the non-linearity of the emission model and the vehicle category-specific emission factors involved. This was illustrated in the final surface maps produced for each modelling scenario in § 5.4 and § 5.5.

6.2 Research limitations

Despite the above-mentioned contributions and accomplished objectives, limitations did exist in both the development as well as the implementation stages of the four main phases of the methodological framework. These limitations provide a scope for future work in the context of uncertainty in average speed-based emission predictions using macroscopic traffic flow models. These are discussed below separately for each of the four main phases.

Phase I: Five key limitations can be identified in the first phase, three of which arise at the development stage whereas the last two arise at the implementation stage. The first limitation concerning the chosen optimisation algorithm is identified. Phase I has focussed on the use of a relatively fast and simple one, the N-M direct search method. The literature review of § 2.2.2 has shown the different optimisation algorithms that are being used or developed specifically for macroscopic traffic flow models. The potential of these algorithms, for example the kernel-based cross entropy method proposed by Ngoduy and Maher (2012), is worth investigating to study their influence on the resulting uncertainty in parameter vectors, and ultimately on emission predictions. Generally, optimisation algorithms are compared in the context of traffic flow modelling outputs only. Given the objectives of this research, it would be preferable to compare optimisation algorithms in terms of the ensemble-based optimisation framework, and ultimately on emission predictions. It is worth noting, however, that the increased complexity of optimisation algorithms generally results in increased computational requirements; this can be particularly disadvantageous here as the phase is already computationally demanding.

The second limitation concerning the statistical testing approach is noted. Phase I has proposed the use of two statistical tests to objectively select *top-solutions*. However,

in using the two statistical tests, two key assumptions were made. The first is that the comparison between predicted and measured traffic variables disregards the timing or ordering of each. This implies that the distribution of predicted traffic is compared with the entire distribution of measured traffic, instead of applying the statistical tests on the errors themselves (which would be more demanding). The second assumption is that the model outputs are not correlated. In the case of spatiotemporal traffic data, this is often not the case. Ni et al. (2004) proposed a 'batch' technique that can potentially resolve the issue of correlation. However, in the context of an ensemble-based optimisation approach, such technique can be computationally demanding as well. It does, however, offer the potential to improving the process of selecting *top-solutions*, and is thus worth investigating. Here it is worth noting that the main risk in any approach is selecting a *top-solution* that should not be selected. The question that arises is thus: what is the impact of including a 'poor' parameter vector in the *top-solutions*, on the confidence intervals of emission predictions?

The third limitation concerning the 'optimal' number of optimisation runs is identified. The influence of changing the number of optimisation runs has been studied and reported in § 5.2.3. Upon implementation, the number of optimisation runs was assumed fixed. While the ensemble-based approach allows exploration of *top-solutions*, and provides a more rigorous approach to identifying the distribution of possible parameter vectors, the computational requirement limits the ability to find the 'optimal' number of runs.

Upon the implementation of Phase I, a distribution for each parameter had to be set out in order to draw initial parameters required for the optimisation algorithm. While the distributions of those related to the fundamental diagram were obtained using multiple days of real traffic data,

- ranges of few (unobservable) METANET-related parameters were obtained from previous studies, which are not always within the corresponding parameter ranges of the final *top-solutions*;
- initial parameters were drawn from approximate normal distributions rather than directly from the obtained empirical distributions; and
- initial parameters were drawn independently from each other, while the study has ultimately shown correlation between different model parameters.

The influence of changing the ranges of unobservable parameters on the final *top-solutions* can be explored in further studies. Also, sampling parameters directly from the joint empirical distributions can be a better and more consistent alternative to initialise the ensembles of optimisation runs.

In addition to that, also upon the implementation of Phase I, *top-solutions* for the CTM modelling scenarios were not found for all time periods. In the context of this work, it was not possible to pinpoint the exact reason of why this is the case. The fact that METANET modelling scenarios were able to always find solutions suggests

structural issues lie behind failing to find *top-solutions* for those few time periods using CTM or similarly enhanced structural properties lie behind METANET's performance. Investigating the characteristics of a time period where *top-solutions* cannot be found and the potential consequences in terms of using CTM to real-world applications (i.e. not only for emission predictions), are two questions that can further be investigated to provide more insights on such observations and to enable suggesting new techniques to resolving them.

Phase II: In the second phase, three key limitations arise. The first is due to the resolution of measurement data i.e. it can be viewed as arising from the data preparation preliminary phase. It thus has an influence not only on this phase but also on the first phase. Although measurement data fitting the requirements of macroscopic traffic flow model simulation (or calibration and validation) is abundant, most measurement data does not completely match the resolution of macroscopic traffic flow models. This necessitated (in this study but more generally in any other study - the reason behind aggregation being a key step of Phase II): the interpolation of measurement data to match the input requirements to model simulation; the aggregation of model outputs to the resolution of measurements to minimise the objective function value; and the aggregation of model outputs to estimate the calibration and validation prediction errors of Phase II. This implied that errors at the spatiotemporal resolution of traffic flow models are not entirely understood. First, an optimal parameter vector obtained based on aggregate model outputs might not be the optimal parameter vector if the initial resolution is considered, possibly leading to underestimation of errors at the aggregate level. Second, the magnitude of prediction errors at the initial resolution (if these were used for the optimisation problem) might be lower or higher, depending on the structural performance of the macroscopic traffic flow models at its resolution (which might not necessarily be worse). While the methodological framework itself does not depend on the resolution of the measurements, nor on the resolution of the final prediction errors, the influence of higher measurement resolution on the confidence intervals obtained at the end would be worth investigating.

Another limitation at the development stage concerning what constitute uncertainty in a future scenario is noted. Phase II involved estimating validation errors in an attempt to understand potential prediction errors of future scenarios. However, uncertainty in modelling future scenarios is much more complex as noted in Chapter 3. It is thus worth investigating how uncertainty in the input to traffic flow models of future scenarios can add up to the prediction errors if they were also considered. This can be a step closer to better understanding confidence intervals of emission predictions for future scenarios. The main challenge here is devising a new methodological framework to quantifying uncertainty in model input of future scenarios. More challenging to investigate is the uncertainty in model structure in future scenarios. Changes in infrastructure (road network), driving behaviour (drivers), and vehicle characteristics (or technology) can all potentially necessitate new model structures that can describe the

impact of such changes on traffic dynamics. Perhaps one way to tackle such problem is by looking into how the present system has changed in comparison to the past in order to infer future uncertainties; although future changes at the three levels of road network, drivers, and technology are different to those of the present and/or past.

In addition to the above, the estimation of both calibration and validation errors during the implementation stage involved using the *top-solutions* of six time periods of a road network. Alongside the challenge of obtaining the optimal number of optimisation runs, there is the challenge of obtaining the optimal number of time periods which allows capturing the uncertainty associated with modelling traffic flow on a typical day. Again, while this does not influence the methodology and the steps involved, it is rather an implementation problem, reflecting the computational demand of the ensemble-based approach.

Phase III: Main limitations of the third phase arise at the development stage. The first is related to the selected sampling method. Monte Carlo is one simple approach to random sampling, and it has been chosen in order to first illustrate the method for uncertainty propagation. However, this work can be extended to more complex sampling approaches of the bivariate error distributions. Examples of these are the Markov-Chain Monte Carlo sampling and Latin Hypercube Sampling.

In terms of the gridding approach, the influence of changing the size of the sampling grids has been studied and reported in § 5.5.3. However, in any scenario, there remains the assumption that errors are equally important across the entire grid square. Refinements to this approach could be investigated, with their impact on the surfaces of confidence intervals monitored. Examples of such refinements include the use of moving average methods, or weighted contribution of errors across the grid. The complexity of the problem implies that there might not be a single solution to assigning errors to respective prediction points. The aim of such investigation is rather to check the sensitivity of the results to the different approaches.

The last limitation of Phase III is that prediction errors are not distinguished by space and time. This is also related to the idea of correlation in model outputs described earlier (on the assumptions of statistical testing in Phase I). In the final example provided in § 5.6, traffic speed-density predictions and their associated upper and lower bounds involve the assumption that predictions (and their errors) are not correlated. This assumption may be challenged since, for instance, congestion builds up and dissipates in space and time. As a result, high errors at a particular measurement site and measurement time step can also imply high errors at the next measurement site or next time step. A potential remedy would be to refine the gridding approach to be dependent on the location and time of prediction, not only on where the prediction is within the feasible region. Although this would not necessarily affect the existing results, it would help to reveal any dependence of errors in space and time. Depending

on the results, one can decide whether the gridding approach needs to take this into account or not.

Phase IV: The main limitation of the last phase is that probabilistic-based emission predictions only take into account uncertainty in the traffic flow modelling outputs, which was set from the start as the scope of this research. So, it arises at the development stage. Chapter 2 identified studies which have focussed on quantifying uncertainty in emission factors. Given the complexity of interaction between traffic flow model outputs and the emission model (or embedded emission factors), the interaction between uncertainty in traffic flow model outputs and uncertainty in emission factors is also worth investigating. This would lead to a better understanding of the confidence intervals in the final emission predictions. The main challenge here is incorporating different uncertainty quantification methods for the different sources, while taking into consideration the difficulties of obtaining emission measurement data in comparison to traffic measurement data. Chapter 3 also highlighted other potential sources of uncertainty in the modelling chain studied, in addition to those of the traffic flow model outputs. In § 5.5, uncertainty due to incomplete 'compatibility' of the uni-class traffic flow models and the multi-class emission model was focussed upon. This source of uncertainty can be accommodated for using multi-class traffic flow models instead, such as multi-class CTM or multi-class METANET. The main challenge behind using multi-class traffic flow models, however, is obtaining the necessary traffic data (mainly traffic speed and density) which allows the optimisation of traffic flow models (as described in Phase I) and the estimation of prediction errors (as described in Phase II). Of course, the use of more complex traffic flow models can also influence uncertainty in the parameters; this has been one of the main conclusions of Phase I, where uncertainty in model structure is shown to be related to that in model parameters.

6.3 Future research

The above has highlighted the main limitations of the methodological framework and suggested refinements that could potentially improve upon the final results, or at least provide a better understanding of the problem being studied in this research. However, these refinements are focussed on the propagation of uncertainty from macroscopic traffic flow models to average speed-based emission predictions. In this section, further potential research is explored in the broader areas of: the 'traffic flow - emission' modelling chain; and the five-step 'traffic assignment - traffic flow - emission - air quality - human exposure' modelling chain introduced in Chapter 1.

Chapter 2 of this thesis showed that emissions are not only predicted based on average speeds, nor are they predicted using the outputs of macroscopic traffic flow models. Various integration approaches have been developed and are used to informing policies. The availability of historical and adequate data to quantify uncertainty in prediction

errors of macroscopic traffic flow models was one main reason (but not the only reason) to selecting these as the basis for the methodological framework. The question that arise is how can the methodological framework be used in other integration approaches? The answer to such question depends on the complexity of the integration approach in comparison to that of this research. For instance, macroscopic traffic flow models were shown to be indirectly used with driving mode emission models or even instantaneous emission models. In such integration approaches, the first two phases (I and II) can arguably remain the same. Potential future work remains on how to use the prediction outputs and their errors so that they could be propagated efficiently to driving mode models or instantaneous emission models (for example). This would involve propagating uncertainty to the intermediary step (for example to the derived driving modes) and then to emission predictions. Achieving this could help reveal differences between first-order and second-order traffic flow models which might not have been captured through the use of average speed-based emission models. In the more complex integration of microscopic traffic flow models with instantaneous emission models, modifications starts from the data preparation phase. With the recent emergence of individual vehicle data, there is a potential to developing similar methodological framework to estimate prediction errors and to propagate them to instantaneous emission models. However, the nature of individual vehicle data is quite different than aggregate traffic data which makes it difficult (or even impractical) to estimate prediction errors of each modelled vehicle trajectory in a relatively straightforward manner. One reason is that vehicle data cannot be assigned individually to modelled data. As such, the problem becomes more challenging; while the use of microscopic approaches might become easier as a result of data availability, the study of uncertainty becomes more crucial.

Alternatively, the 'traffic flow - emission' modelling chain was positioned in Chapter 1 within a much more complex five-step 'traffic assignment - traffic flow - emission - air quality - human exposure' modelling chain. As a result, it is argued here that further research needs to be neither restricted to the choice of macroscopic traffic flow models and average speed-based emission models, nor to the sub-modelling chain of 'traffic flow - emission' modelling. Many of the studies reviewed in Chapter 2 have used this sub-modelling chain to further model air quality and human exposure. As such, potential future work is incorporating the resulting confidence interval surfaces in the prediction of, for example, air quality levels.

This also leads to another (and perhaps more interesting) research area which is the automatic incorporation of uncertainty in real-time (or on-line) applications. The proposed framework here focusses on the use of historical data and develops a methodology accordingly. Currently, technological advancements and the demand of real-time information provide research opportunities (or challenges) to answering how uncertainty can be incorporated in real-time systems in order to improve upon the reported information.

A final note on potential future work is motivated by the conclusion in Evans (2012) which says: "...it is also worth highlighting an alternative viewpoint put forward by Beck (1987), who suggests that rather than asking what the future will be, given parameters now, we instead ask what parameters would be necessary now, to create a reasonable future...". In this work, a 'bottom-up' approach has been suggested into quantifying uncertainty in emission modelling systems. Following this point of view, an important question is, how these can be used in a 'top-down' manner to devise policies that can create a notably better environment to live in? For instance, if policy and decision makers decided to demand certain emissions levels (from transport) with certain confidence, then what are the maximum traffic levels (including how congestion behaves with or without traffic management operation) that would then meet these emission levels, and with what confidence?

Appendix A

MIDAS-based R Functions

Both the simulation and optimisation of CTM-basic, CTM-vsl, METANET-basic, and METANET-vsl traffic flow models given MIDAS traffic data have been implemented in R (R Core Team, 2015) through developing the functions of Figure A.1. A total of 12 functions are built for the extraction and pre-processing of traffic data, three functions are built to initialise the simulation or optimisation of CTM or METANET (i.e. regardless of the model to be used), and two main functions were built for the simulation or optimisation of CTM and METANET. Each of these take a number of arguments, providing the user appropriate flexibility. A description of each of these functions is provided below in terms of their purpose and main inputs.

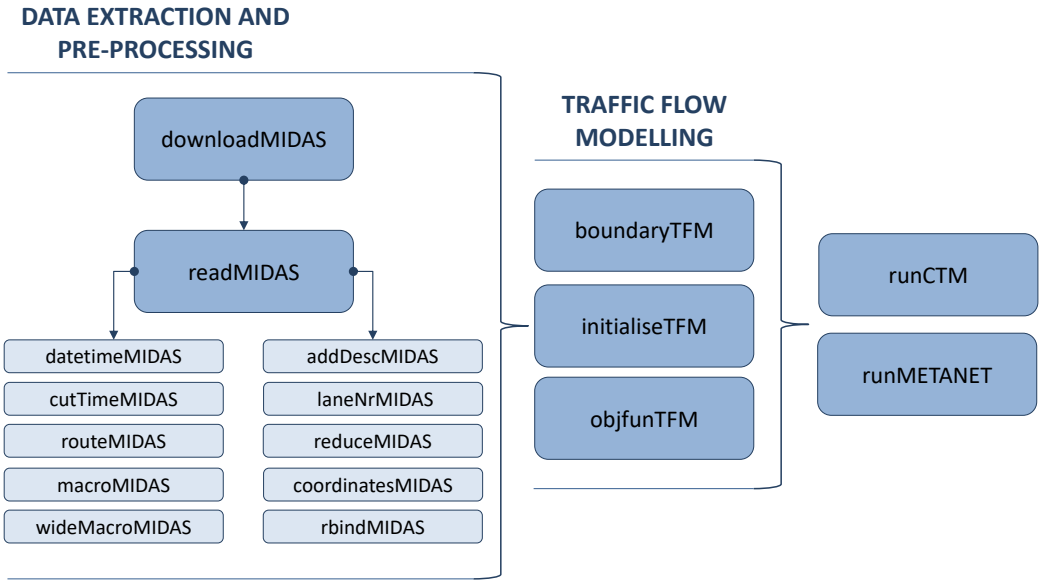


FIGURE A.1: Flow-diagram of MIDAS-based R functions developed for the implementation of CTM and METANET traffic flow models.

1. **downloadMIDAS** downloads traffic data file, unzips it, and decodes it to .csv format using MIDAS decoder. The function first asks for the MIDAS user name and password. If available, the function additionally requires for the destination directory, RCC, and day of traffic data needed.
2. **readMIDAS** reads .csv MIDAS traffic file and outputs a 'midas' object which is a list of two elements. The first is a data frame of raw traffic data, while the second is a description of the read traffic data which includes its date and weekday. Once the object is created, a number of functions can be applied on it. Functions of interest to traffic flow modelling are listed and described briefly below.
    - **datetimeMIDAS** transforms the date and time columns of raw traffic data to appropriate 'date-time' column which makes it easier to manipulate. The main input is a 'midas' object with separate day and time columns.
    - **cutTimeMIDAS** uses the date-time column in order to subset the traffic data based on a start and end time. The main input is a 'midas' object and a specific start and end times in hours.
    - **routeMIDAS** uses the geographic address column in order to subset a specific route to be studied i.e. all measurement sites between a start point and an end point. The main input is a 'midas' object, motorway name, and optionally start and end points along the route; if these are not specified, traffic data for the entire motorway is extracted. Of course, the shorter the selected route, the better in terms of checking data, for instance in terms of potential duplicates. The user can also specify whether mainline, on-ramps, and/or off-ramps need to be extracted.
    - **macroMIDAS** calculates macroscopic traffic variables of average flow, speed, and density using raw traffic data per lane. The main input is a 'midas' object with raw traffic data.
    - **wideMacroMIDAS** reshapes 'midas' data frame element from long format to wide format i.e. to columns per measurement site rather than columns per traffic variable. The main input is a 'midas' object alongside the traffic variables to be reshaped and the main output is a list of data frames for each traffic variable.
    - **addDescMIDAS** extracts details from the geographic address of each measurement site in a 'midas' object. Details include the road, its code, its location (mainline, on-ramp, or off-ramp), and its direction (clockwise or anti-clockwise). The main input is a 'midas' object with geographic addresses as one of its columns.
    - **laneNrMIDAS** tabulates the number of lanes of each measurement site in the a 'midas' object. The main input is a 'midas' object with raw traffic data.
    - **reduceMIDAS** removes columns where all entries are NA - these correspond to measurement sites with less than seven lanes. The main input is a

'midas' object.

- **coordinatesMIDAS** adds XY coordinates of measurements sites in a 'midas' object. Coordinates for each geographic address is separately provided by Highways England. The main input is a 'midas' object with geographic addresses as one of its columns.
- **rbindMIDAS** row binds multiple 'midas' objects to a single one. The main input is a list of 'midas' objects.

3. **boundaryTFM** extracts necessary inputs (initial and boundary conditions) for the macroscopic traffic flow modelling of a specific route. The main input is a 'midas' object with macroscopic traffic data (rather than per lane) on the mainline, on-ramps, and/or off-ramps. If off-ramps exist, traffic flow rates are also calculated and prepared.
4. **initialiseTFM** initialises the matrices of traffic variables to be modelled. This is mainly used in conjunction with 'runCTM' or 'runMETANET'. The main input is the initial and boundary conditions outputs of 'boundaryTFM', as well as modelling related inputs such as time step and total number of cells which allows preparing the matrices with appropriate size (number of rows and columns).
5. **objfunTFM** calculates the objective function value to be optimised. This is mainly used in conjunction with 'runCTM' or 'runMETANET'. The main input is the matrices of observed and modelled traffic variables. Other inputs include specifying the objective function. This function currently supports the use of either Root Squared Mean Error (RMSE) or Mean Absolute Error (MAE).
6. **runCTM** simulates CTM given a specific set of parameters or optimises any or all of the CTM parameters. The main input is the road network and time period characteristics, observed traffic data necessary for the calculation of objective function, and most importantly whether simulation or optimisation is required. If variable speed limits are provided as inputs, CTM-vsl is simulated or optimised, otherwise CTM-basic is used. In the case of simulation i.e. the parameters are specified, the main output of the function is a list of data frames of modelled traffic variables, as well as the performance measure value (similar to the objective function used) given the observed traffic data. In the case of optimisation i.e. one or more parameter is not specified, the output is a function by itself to be used as input in an optimisation algorithm function chosen by the user; it mainly calculates the objective function at each iteration of the optimisation algorithm.
7. **runMETANET** simulates METANET given a specific set of parameters or optimises any or all of the METANET parameters. The inputs and outputs are similar to those of the 'runCTM' function. Similarly, if variable speed limits are provided as inputs, METANET-vsl is simulated or optimised, otherwise METANET-basic is used.

Appendix B

Ensemble-based Optimisation

The following shows the fundamental diagrams based on *top-solutions* obtained for each time period of the three CTM and M-METANET modelling scenarios. Those for the M60/CTM-basic and M60/M-METANET-basic are provided as an example in § 5.2.1.1 and § 5.2.2.1, respectively.

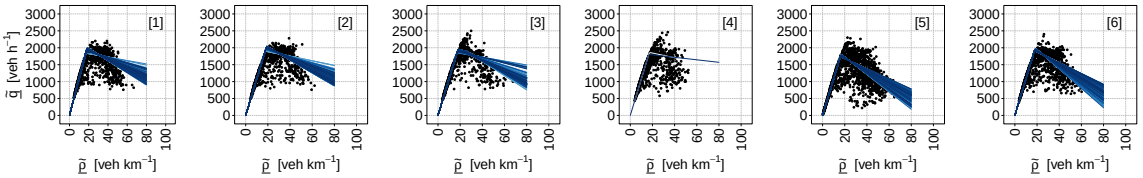


FIGURE B.1: Measured traffic flow per lane [veh h^{-1}] and density per lane [veh km^{-1}] scatter plots (black dots) overlaid by fundamental diagrams (shades of blue) based on *top-solutions* of M1/CTM-basic: 250 for time period [1], 245 for time period [2], 250 for time period [3], 2 for time period [4], 202 for time period [5], and 250 for time period [6].

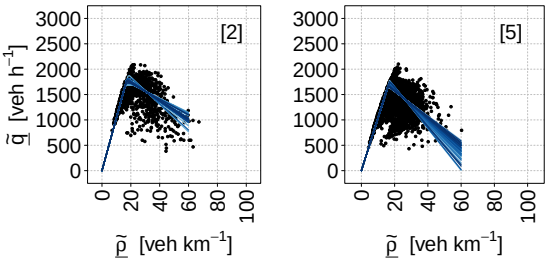


FIGURE B.2: Overlaid fundamental diagrams based on the parameter vectors of *top-solutions* of M25/CTM-vsl: 78 for time period [2] and 85 for time period [5].

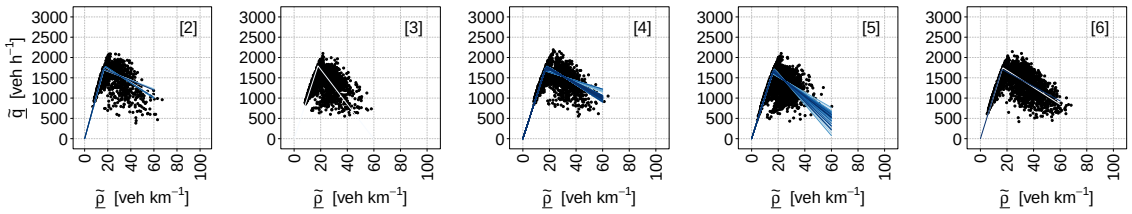


FIGURE B.3: Overlaid fundamental diagrams based on the parameter vectors of *top-solutions* of M25/CTM-basic: 7 for time period [2], 1 for time period [3], 149 for time period [4], 50 for time period [5], and 2 for time period [6].

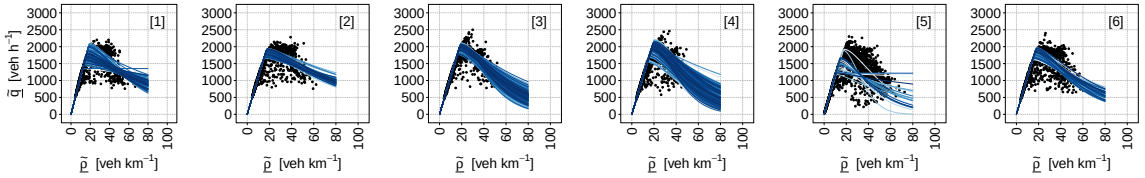


FIGURE B.4: Overlaid fundamental diagrams based on the parameter vectors of *top-solutions* of M1/M-METANET-basic: 155 for time period [1], 105 for time period [2], 249 for time period [3], 250 for time period [4], 22 for time period [5], and 68 for time period [6].

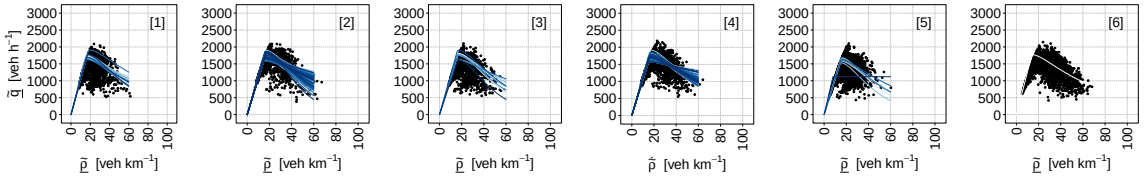


FIGURE B.5: Overlaid fundamental diagrams based on the parameter vectors of *top-solutions* of M25/M-METANET-vsl: 13 for time period [1], 149 for time period [2], 10 for time period [3], 112 for time period [4], 8 for time period [5], and 1 for time period [6].

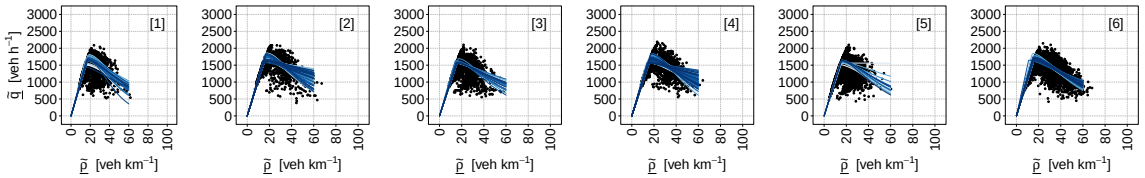


FIGURE B.6: Overlaid fundamental diagrams based on the parameter vectors of *top-solutions* of M25/M-METANET-basic: 41 for time period [1], 119 for time period [2], 41 for time period [3], 149 for time period [4], 20 for time period [5], and 113 for time period [6].

Appendix C

Grid-based Monte Carlo Sampling

The following shows summary statistics (in terms of mean and standard deviations) of evaluated calibration and validation errors of the three CTM and M-METANET modelling scenarios. Those for the M60/CTM-basic and M60/M-METANET-basic are provided as an example in § 5.4.1 and § 5.4.2, respectively.

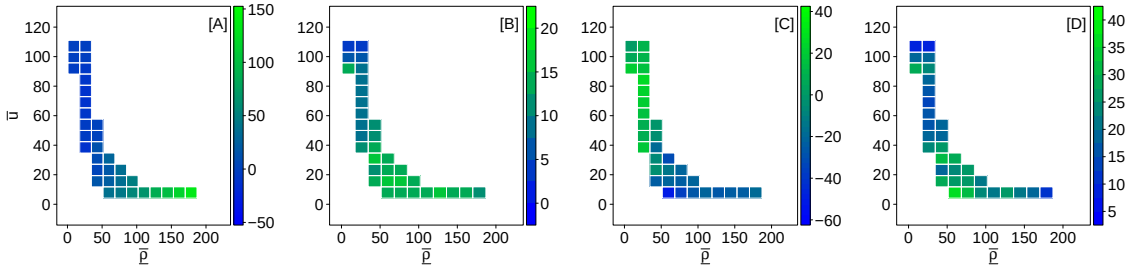


FIGURE C.1: Summary statistics of calibration error distributions obtained for each grid square of the M1/CTM-basic modelling scenario: [A] mean and [B] standard deviation of traffic density errors and [C] mean and [D] standard deviation of speed errors.

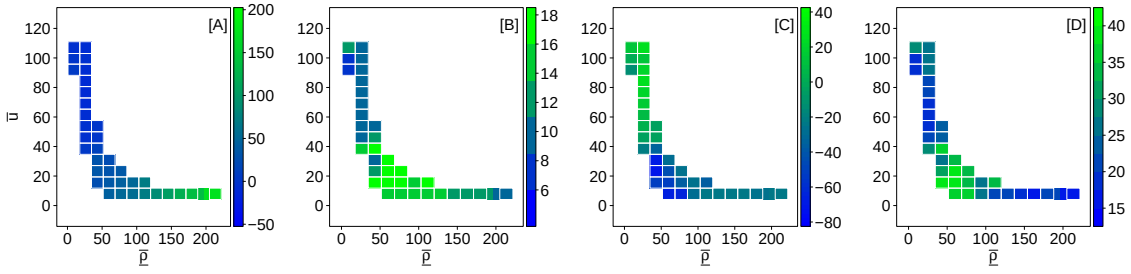


FIGURE C.2: Summary statistics of validation error distributions obtained for each grid square of the M1/CTM-basic modelling scenario.

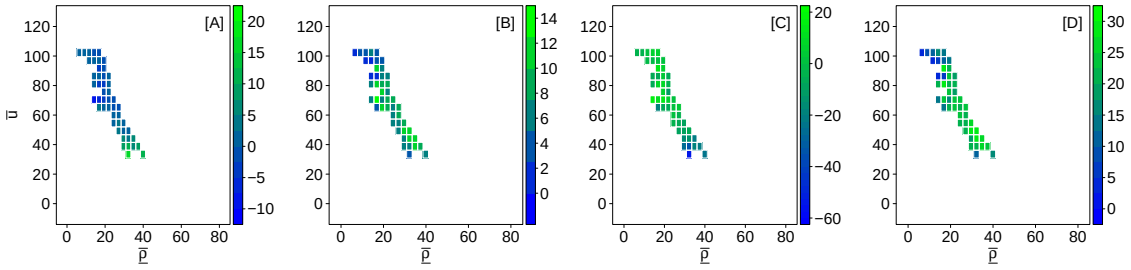


FIGURE C.3: Summary statistics of calibration error distributions obtained for each grid square of the M25/CTM-vsl modelling scenario.

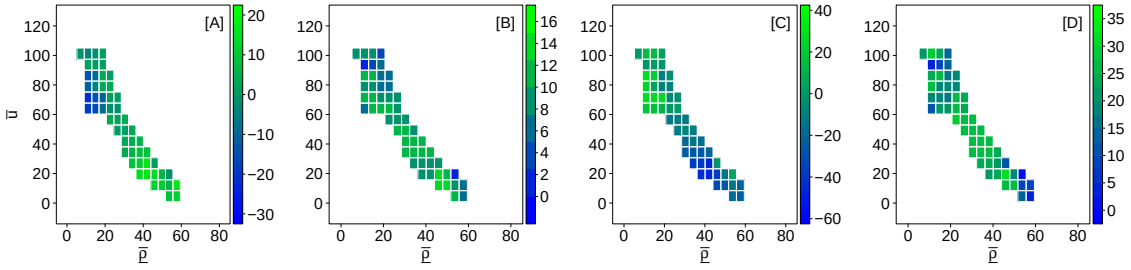


FIGURE C.4: Summary statistics of validation error distributions obtained for each grid square of the M25/CTM-vsl modelling scenario.

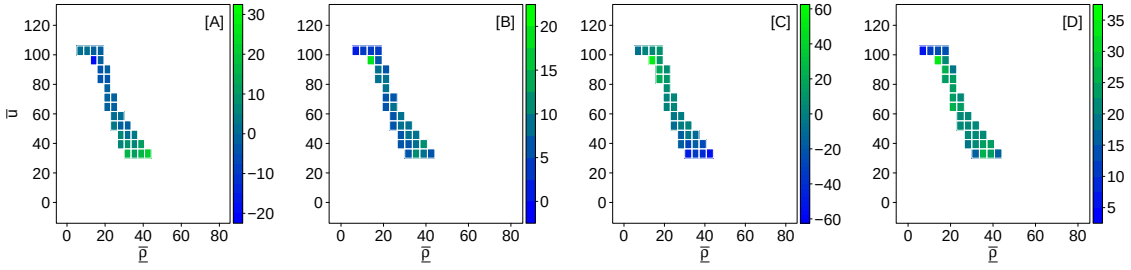


FIGURE C.5: Summary statistics of calibration error distributions obtained for each grid square of the M25/CTM-basic modelling scenario.

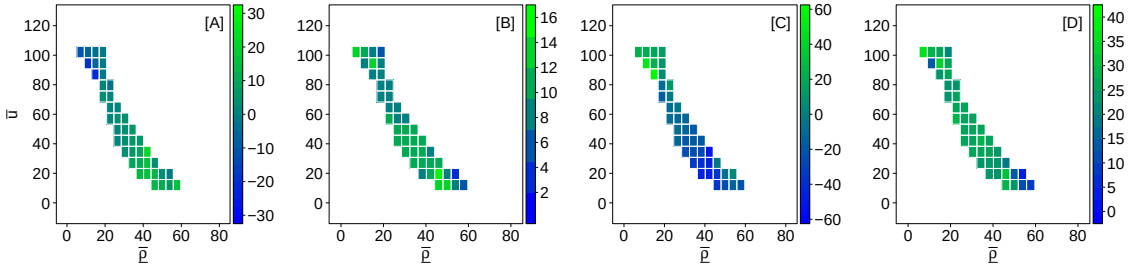


FIGURE C.6: Summary statistics of validation error distributions obtained for each grid square of the M25/CTM-basic modelling scenario.

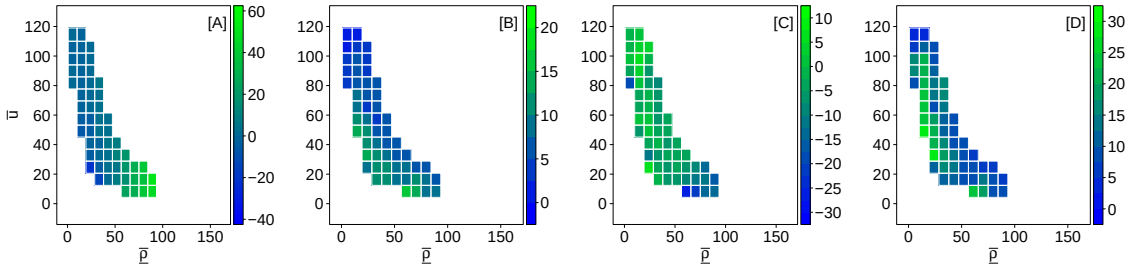


FIGURE C.7: Summary statistics of calibration error distributions obtained for each grid square of the M1/M-METANET-basic modelling scenario: [A] mean and [B] standard deviation of traffic density errors and [C] mean and [D] standard deviation of speed errors.

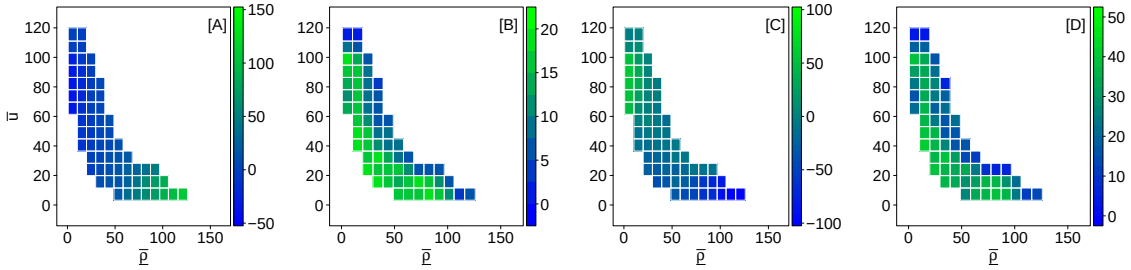


FIGURE C.8: Summary statistics of validation error distributions obtained for each grid square of the M1/M-METANET-basic modelling scenario.

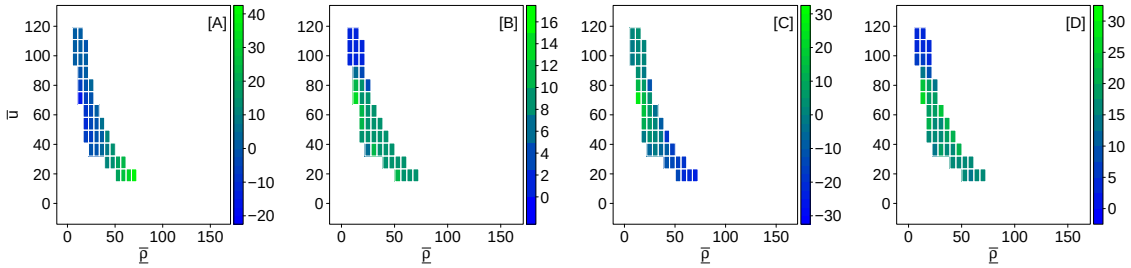


FIGURE C.9: Summary statistics of calibration error distributions obtained for each grid square of the M25/M-METANET-vsl modelling scenario.

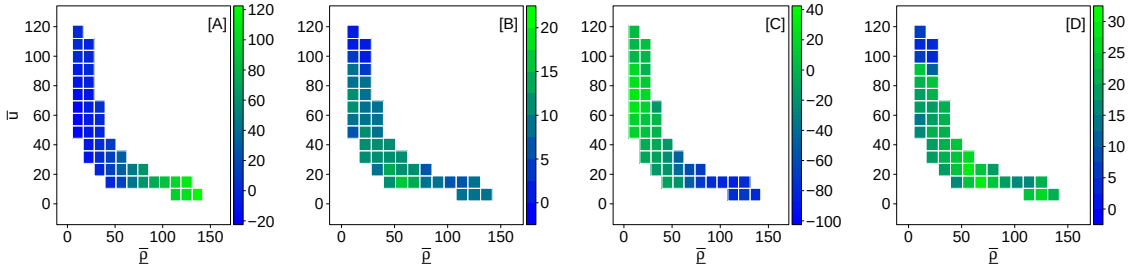


FIGURE C.10: Summary statistics of validation error distributions obtained for each grid square of the M25/M-METANET-vsl modelling scenario.

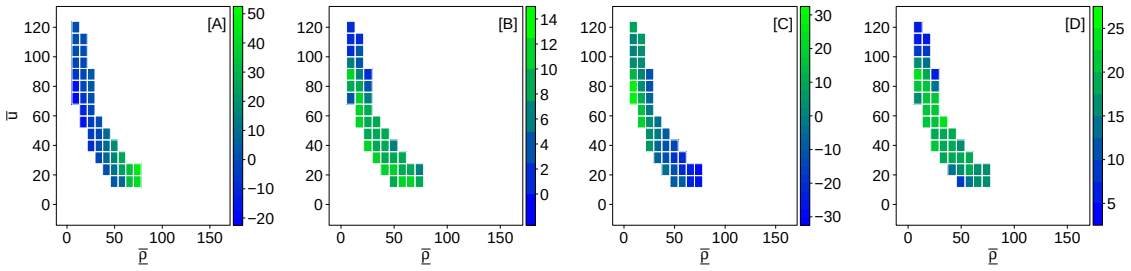


FIGURE C.11: Summary statistics of calibration error distributions obtained for each grid square of the M25/M-METANET-basic modelling scenario.

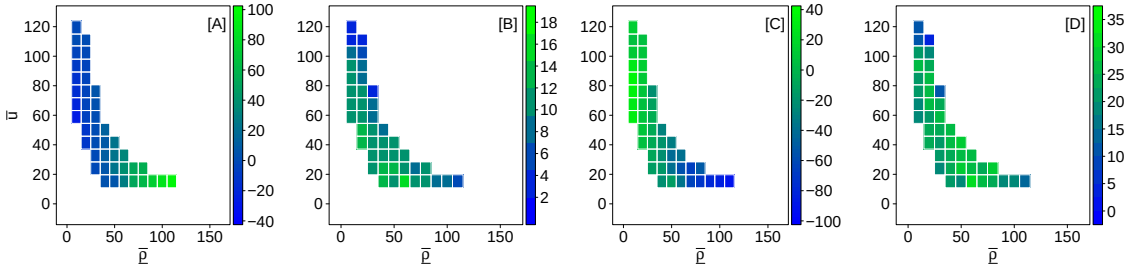


FIGURE C.12: Summary statistics of validation error distributions obtained for each grid square of the M25/M-METANET-basic modelling scenario.

References

Alessandri, A., Di Febbraro, A., Ferrara, A., & Punta, E. (1999). Nonlinear optimization for freeway control using variable-speed signaling. *IEEE Transactions on Vehicular Technology*, *48*(6), 2042–2052.

Aw, A., & Rascle, M. (2000). Resurrection of "second order" models of traffic flow. *SIAM Journal on Applied Mathematics*, *60*(3), 916–938.

Aziz, H. M., & Ukkusuri, S. V. (2012). Integration of environmental objectives in a system optimal dynamic traffic assignment model. *Computer-Aided Civil and Infrastructure Engineering*, *27*(7), 494–511.

Bai, S., Chiu, Y. E., & Niemeier, D. A. (2007). A comparative analysis of using trip-based versus link-based traffic data for regional mobile source emissions estimation. *Atmospheric Environment*, *41*(35), 7512–7523.

Bandeira, J. M., Coelho, M. C., Sá, M. E., Tavares, R., & Borrego, C. (2011). Impact of land use on urban mobility patterns, emissions and air quality in a Portuguese medium-sized city. *Science of the Total Environment*, *409*(6), 1154–1163.

Barlow, T., Latham, S., McCrae, I., & Boulter, P. (2009). A reference book of driving cycles for use in the measurement of road vehicle emissions. *TRL Published Project Report*.

Barth, M., An, F., Norbeck, J., & Ross, M. (1996). Modal emissions modeling: A physical approach. *Transportation Research Record: Journal of the Transportation Research Board*(1520), 81–88.

Bauer, D. F. (1972). Constructing confidence sets using rank statistics. *Journal of the American Statistical Association*, *67*(339), 687–690.

Beck, M. B. (1987). Water quality modeling: a review of the analysis of uncertainty. *Water Resources Research*, *23*(8), 1393–1442.

Bellomo, N., & Dogbe, C. (2011). On the modeling of traffic and crowds: A survey of models, speculations, and perspectives. *SIAM Review*, *53*(3), 409–463.

Bellomo, N., Marasco, A., & Romano, A. (2002). From the modelling of driver's behavior to hydrodynamic models and problems of traffic flow. *Nonlinear Analysis: Real World Applications*, *3*(3), 339–363.

Ben-Tal, A., Do Chung, B., Mandala, S. R., & Yao, T. (2011). Robust optimization for emergency logistics planning: Risk mitigation in humanitarian relief supply chains. *Transportation Research Part B: Methodological*, *45*(8), 1177–1189.

Bihorel, S., & Baudin, M. (2015). neldermead: R port of the scilab neldermead module [Computer software manual]. Retrieved from http://CRAN.R-project.org/package=neldermead (R package version 1.0-10)

Bishop, G. A., Starkey, J. R., Ihlenfeldt, A., Williams, W. J., & Stedman, D. H. (1989). IR long-path photometry: A remote sensing tool for automobile emissions. *Analytical Chemistry*, *61*(10), 671A–677A.

Borrego, C., Tchepel, O., Costa, A. M., Amorim, J. H., & Miranda, A. I. (2003). Emission and dispersion modelling of Lisbon air quality at local scale. *Atmospheric Environment*, *37*(37), 5197–5205.

Borrego, C., Tchepel, O., Salmim, L., Amorim, J., Costa, A. M., & Janko, J. (2004). Integrated modeling of road traffic emissions: application to Lisbon air quality management. *Cybernetics and Systems: An International Journal*, *35*(5-6), 535–548.

Boulter, P. G., McCrae, I. S., & Barlow, T. J. (2007). *A review of instantaneous emission models for road vehicles.* TRL Limited.

Boundary-Line™. (2016). *[SHAPE geospatial data] EDINA Digimap Ordnance Survey (OS) Service.* Retrieved 2016-10-31, from http://digimap.edina.ac.uk/

Box, M. J. (1965). A new method of constrained optimization and a comparison with other methods. *The Computer Journal*, *8*(1), 42–52.

Brackstone, M., & McDonald, M. (1999). Car-following: A historical review. *Transportation Research Part F: Traffic Psychology and Behaviour*, *2*(4), 181–196.

Carey, M., & Bowers, M. (2012). A review of properties of flow–density functions. *Transport Reviews*, *32*(1), 49–73.

Carlson, R. C., Papamichail, I., Papageorgiou, M., & Messmer, A. (2010). Optimal motorway traffic flow control involving variable speed limits and ramp metering. *Transportation Science*, *44*(2), 238–253.

Cernuschi, S., Giugliano, M., Cemin, A., & Giovannini, I. (1995). Modal analysis of vehicle emission factors facteurs d'émission modaux des véhicules. *Science of the Total Environment*, *169*(1), 175–183.

Chiou, Y. C., Chiou, Y. S., & Hsieh, C. W. (2013). An integrated emission and dispersion model under mixed traffic conditions. *Journal of the Eastern Asia Society for Transportation Studies*, *10*, 1786–1796.

Chiou, Y. C., & Hsieh, C. W. (2012). Mixed traffic cell transmission models: Development and validation. *Journal of the Chinese Institute of Transportation*, *24*, 245–276.

Chiou, Y. C., Huang, Y. F., & Lin, P. C. (2012). Optimal variable speed-limit control under abnormal traffic conditions. *Journal of the Chinese Institute of Engineers*, *35*(3), 299–308.

Clavreul, J., Guyonnet, D., & Christensen, T. H. (2012). Quantifying uncertainty in LCA-modelling of waste management systems. *Waste Management*, *32*(12), 2482–2495.

Conover, W. J., Johnson, M. E., & Johnson, M. M. (1981). A comparative study of tests for homogeneity of variances, with applications to the outer continental shelf bidding data. *Technometrics*, *23*(4), 351–361.

Corsmeier, U., Imhof, D., Kohler, M., Kühlwein, J., Kurtenbach, R., Petrea, M., Rosenbohm, E., Vogel, B., & Vogt, U. (2005). Comparison of measured and model-calculated real-world traffic emissions. *Atmospheric Environment*, *39*(31), 5760–5775.

Cremer, M., & Ludwig, J. (1986). A fast simulation model for traffic flow on the basis of boolean operations. *Mathematics and Computers in Simulation*, *28*(4), 297–303.

Cremer, M., & Papageorgiou, M. (1981). Parameter identification for a traffic flow model. *Automatica*, *17*(6), 837–843.

Csikós, A., Varga, I., & Hangos, K. M. (2014). Motorway control system for the reduction of pollutant concentrations in local rural areas. In *22nd Mediterranean Conference of Control and Automation (MED)* (pp. 966–971).

Csikós, A., Varga, I., & Hangos, K. M. (2015). Modeling of the dispersion of motorway traffic emission for control purposes. *Transportation Research Part C: Emerging Technologies*, *58*, 598–616.

Cullen, A. C., & Frey, H. C. (1999). *Probabilistic techniques in exposure assessment: A handbook for dealing with variability and uncertainty in models and inputs.* Springer Science & Business Media.

Daamen, W., Buisson, C., & Hoogendoorn, S. P. (2014). *Traffic simulation and data: Validation methods and applications.* CRC Press.

Daganzo, C. F. (1994). The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory. *Transportation Research Part B: Methodological*, *28*(4), 269–287.

Daganzo, C. F. (1995a). The cell transmission model, part II: Network traffic. *Transportation Research Part B: Methodological*, *29*(2), 79–93.

Daganzo, C. F. (1995b). Requiem for second-order fluid approximations of traffic flow. *Transportation Research Part B: Methodological*, *29*(4), 277–286.

Darbha, S., Rajagopal, K. R., & Tyagi, V. (2008). A review of mathematical models for the flow of traffic and some recent results. *Nonlinear Analysis: Theory, Methods and Applications*, *69*(3), 950–970.

de Dios Ortúzar, J., & Willumsen, L. (2011). *Modelling transport.* John Wiley & Sons.

Del Castillo, J. M., Pintado, P., & Benitez, F. G. (1994). The reaction time of drivers and the stability of traffic flow. *Transportation Research Part B: Methodological*, *28*(1), 35–60.

Deo, P., De Schutter, B., & Hegyi, A. (2009). Model predictive control for multi-class traffic flows. *IFAC Proceedings Volumes*, *42*(15), 25–30.

Department for Environment, Food and Rural Affairs. (2011). *UK and EU Air Quality Policy Context.* Retrieved 2016-09-01, from http://uk-air.defra.gov.uk/air-pollution/uk-eu-policy-context

Department for Environment Food and Rural Affairs. (2016). *National Atmospheric Emissions Inventory - Emission factors for transport.* Retrieved 2016-09-01, from http://naei.defra.gov.uk/data/ef-transport

Department for Transport. (2013). *Road transport forecasts 2013.* Retrieved 2016-09-01, from https://www.gov.uk/government/publications/road-transport-forecasts-2013/

Department for Transport. (2016a). *Road safety data.* Retrieved 2016-09-01, from https://data.gov.uk/dataset/road-accidents-safety-data

Department for Transport. (2016b). *Traffic counts.* Retrieved 2016-09-01, from http://www.dft.gov.uk/traffic-counts/

Dervisoglu, G., Gomes, G., Kwon, J., Horowitz, R., & Varaiya, P. (2009). Automatic calibration of the fundamental diagram and empirical observations on capacity. In *Transportation Research Board 88th Annual Meeting* (Vol. 15).

El-Fadel, M., & Hashisho, Z. (2000). Vehicular emissions and air quality assessment in roadway tunnels: The Salim Slam tunnel. *Transportation Research Part D: Transport and Environment*, *5*(5), 355–372.

Esteves-Booth, A., Muneer, T., Kubie, J., & Kirby, H. (2002). A review of vehicular emission models and driving cycles. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, *216*(8), 777–797.

European Commission. (2016). *Air quality - existing legislation.* Retrieved 2016-09-01, from http://ec.europa.eu/environment/air/quality/legislation/existing_leg.htm

European Environment Agency. (2016). EMEP/EEA air pollutant emission inventory guidebook - 2016. *European Environment Agency Report*, *21*.

Evans, A. (2012). Uncertainty and error. In *Agent-based models of geographical systems* (pp. 309–346). Springer.

FEAT. (2017). *Fuel Efficiency Automobile Test FEAT Data Center.* Retrieved 2017-01-01, from http://www.feat.biochem.du.edu/

Feldman, O., & Maher, M. (2002). Optimisation of traffic signals using a cell transmission model. In *34th Annual Universities' Transport Study Group Conference, Napier University, Edinburgh.*

Fontes, T., Pereira, S. R., Fernandes, P., Bandeira, J. M., & Coelho, M. C. (2015). How to combine different microsimulation tools to assess the environmental impacts of road traffic? Lessons and directions. *Transportation Research Part D: Transport and Environment*, *34*, 293–306.

Franco, V., Kousoulidou, M., Muntean, M., Ntziachristos, L., Hausberger, S., & Dilara, P. (2013). Road vehicle emission factors development: A review. *Atmospheric Environment*, *70*, 84–97.

Frejo, J. R. D., Camacho, E. F., & Horowitz, R. (2012). A parameter identification algorithm for the METANET model with a limited number of loop detectors. In *51st IEEE Conference on Decision and Control (CDC)* (pp. 6983–6988).

Frey, H. C., Unal, A., Chen, J., & Li, S. (2003). Modeling mobile source emissions based upon in-use and second-by-second data: Development of conceptual approaches for EPA's new MOVES model. In *Proceedings, Annual Meeting of the Air and Waste Management Association, Pittsburgh.*

Frey, H. C., Unal, A., Rouphail, N. M., & Colyar, J. D. (2002). Use of on-board tailpipe emissions measurements for development of mobile source emission factors. In *Proceedings of US Environmental Protection Agency Emission Inventory Conference* (pp. 1–13).

Frey, H. C., Unal, A., Rouphail, N. M., & Colyar, J. D. (2003). On-road measurement of vehicle tailpipe emissions using a portable instrument. *Journal of the Air and Waste Management Association*, *53*(8), 992–1002.

Frey, H. C., & Zheng, J. (2002). Probabilistic analysis of driving cycle-based highway vehicle emission factors. *Environmental Science and Technology*, *36*(23), 5184–5191.

Ghalanos, A., & Theussl, S. (2015). Rsolnp: General non-linear optimization using Augmented Lagrange Multiplier method [Computer software manual]. (R package version 1.16.)

Gipps, P. G. (1981). A behavioural car-following model for computer simulation. *Transportation Research Part B: Methodological*, *15*(2), 105–111.

Gipps, P. G. (1986). A model for the structure of lane-changing decisions. *Transportation Research Part B: Methodological*, *20*(5), 403–414.

Gkiotsalitis, K., & Chow, A. (2014). Significance of fundamental diagrams to first-order macroscopic traffic modelling. *International Journal of Transportation*, *2*(2), 15–32.

Golaz, J.-C., Larson, V. E., Hansen, J. A., Schanen, D. P., & Griffin, B. M. (2007). Elucidating model inadequacies in a cloud parameterization by use of an ensemble-based calibration framework. *Monthly Weather Review*, *135*(12), 4077–4096.

Gomes, G., & Horowitz, R. (2006). Optimal freeway ramp metering using the asymmetric cell transmission model. *Transportation Research Part C: Emerging Technologies*, *14*(4), 244–262.

Gomes, G., Horowitz, R., Kurzhanskiy, A. A., Varaiya, P., & Kwon, J. (2008). Behavior of the cell transmission model and effectiveness of ramp metering. *Transportation Research Part C: Emerging Technologies*, *16*(4), 485–513.

Google Inc. (2016). *Google Earth (Version 6.2.9200.0) [Software].* Retrieved 2016-25-08, from https://www.google.co.uk/intl/en_uk/earth/explore/products/

Greenshields, B. D., Channing, W. S., & Miller, H. H. (1935). A study of traffic capacity. In *Highway Research Board Proceedings* (Vol. 1935).

Greenshields, B. D., Thompson, J. T., Dickinson, H. C., & Swinton, R. S. (1934). The photographic method of studying traffic behavior. In *Highway Research Board Proceedings* (Vol. 13).

Groot, N., De Schutter, B., & Hellendoorn, H. (2013). Integrated model predictive traffic and emission control using a piecewise-affine approach. *IEEE Transactions*

on Intelligent Transportation Systems, *14*(2), 587–598.

Groot, N., De Schutter, B., Zegeye, S., & Hellendoorn, J. (2011a). Model-based predictive traffic control: A piecewise-affine approach based on METANET. *IFAC Proceedings Volumes*, *44*(1), 10709–10714.

Groot, N., De Schutter, B., Zegeye, S. K., & Hellendoorn, H. (2011b). Model-based traffic and emission control using pwa modelsa mixed-logical dynamic approach. In *14th International IEEE Conference on Intelligent Transportation Systems (ITSC)* (pp. 2142–2147).

Guenther, P. L., Stedman, D. H., Bishop, G. A., Hannigan, J. W., Bean, J. H., & Quine, R. W. (1991). *Remote sensing of automobile exhaust.* American Petroleum Institute.

Gulliver, J., & Briggs, D. J. (2005). Time–space modeling of journey-time exposure to traffic-related air pollution using GIS. *Environmental research*, *97*(1), 10–25.

Guo, H., Zhang, Q., Shi, Y., & Wang, D. (2007). On-road remote sensing measurements and fuel-based motor vehicle emission inventory in Hangzhou, China. *Atmospheric Environment*, *41*(14), 3095–3107.

Hadiuzzaman, M., & Qiu, T. Z. (2013). Cell transmission model based variable speed limit control for freeways. *Canadian Journal of Civil Engineering*, *40*(1), 46–56.

Han, K., Liu, H., Gayah, V. V., Friesz, T. L., & Yao, T. (2016). A robust optimization approach for dynamic traffic signal control with emission considerations. *Transportation Research Part C: Emerging Technologies*, *70*, 3–26.

HBEFA. (2015). *HandBook Emission FActors for road transport (HBEFA).* Retrieved 2016-09-01, from http://www.hbefa.net/e/

Hegyi, A. (2004). *Model predictive control for integrating traffic control measures.* TU Delft, Delft University of Technology.

Hegyi, A., De Schutter, B., & Hellendoorn, J. (2005a). Model predictive control for optimal coordination of ramp metering and variable speed limits. *Transportation Research Part C: Emerging Technologies*, *13*(3), 185–209.

Hegyi, A., De Schutter, B., & Hellendoorn, J. (2005b). Optimal coordination of variable speed limits to suppress shock waves. *IEEE Transactions on Intelligent Transportation Systems*, *6*(1), 102–112.

Helbing, D. (1996). Gas-kinetic derivation of navier-stokes-like traffic equations. *Physical Review E*, *53*(3), 2366.

Helbing, D. (2001). Traffic and related self-driven many-particle systems. *Reviews of Modern Pphysics*, *73*(4), 1067.

Helly, W. (1959). Simulation of bottlenecks in single-lane traffic flow. In *Proceedings of the Symposium on Theory of Traffic Flow* (pp. 207–238).

Highways England. (2005). *Detector loops for motorways.* (Design Manual for Roads and Bridges, Volume 9, Section 3, Part 1)

Highways England. (2016). *Traffic England a service from Highways England.* Retrieved 2016-09-01, from http://www.trafficengland.com/services-info/

Hoogendoorn, S. P., & Bovy, P. H. L. (2001). State-of-the-art of vehicular traffic flow modelling. In *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering* (Vol. 215, pp. 283–303). Sage Publications.

Hoogendoorn, S. P., & Knoop, V. (2013). Traffic flow theory and modelling. In B. van Wee, J. A. Annema, & D. Banister (Eds.), *The Transport System and Transport Policy: An Introduction* (pp. 125–159). Edward Elgar Publishing Limited.

Hurdle, V., & Son, B. (2000). Road test of a freeway model. *Transportation Research Part A: Policy and Practice*, *34*(7), 537–564.

John, C., Friedrich, R., Staehelin, J., Schläpfer, K., & Stahel, W. A. (1999). Comparison of emission factors for road traffic from a tunnel study (Gubrist tunnel, Switzerland) and from emission modeling. *Atmospheric Environment*, *33*(20), 3367–3376.

Johnson, L., Jamriska, M., Morawska, L., & Ferreira, L. (2000). Vehicle emissions in Australia: From monitoring to modelling. *WIT Transactions on The Built Environment*, *49*.

Kerner, B. S., Konhauser, P., & Schilke, M. (1996). A new approach to problems of traffic flow theory. In *Proceedings of the 13th International Symposiumon Transportation and Traffic Theory, Lyon, France, 24-26 July 1996.*

Kesting, A., Treiber, M., Schönhof, M., & Helbing, D. (2008). Adaptive cruise control design for active congestion avoidance. *Transportation Research Part C: Emerging Technologies*, *16*(6), 668–683.

Kimms, A., & Maassen, K.-C. (2011). Extended cell-transmission-based evacuation planning in urban areas. *Pesquisa Operacional*, *31*(3), 405–441.

Kini, M., & Frey, H. (1997). *Probabilistic evaluation of mobile source air pollution: Volume 1, probabilistic modeling of exhaust emissions from light duty gasoline vehicles* (Tech. Rep.).

Kioutsioukis, I., Kouridis, C., Gkatzoflias, D., Dilara, P., & Ntziachristos, L. (2010). Uncertainty and sensitivity analysis of national road transport inventories compiled with COPERT 4. *Procedia-Social and Behavioral Sciences*, *2*(6), 7690–7691.

Kioutsioukis, I., & Tarantola, S. (2003). Uncertainties in emission inventory modelling. *Deliverable 44 for the ARTEMIS Project, EUR 20697 EN*.

Kotsialos, A., Papageorgiou, M., Diakaki, C., Pavlis, Y., & Middelham, F. (2002). Traffic flow modeling of large-scale motorway networks using the macroscopic modeling tool METANET. *IEEE Transactions on Intelligent Transportation Systems*, *3*(4), 282–292.

Kühlwein, J., & Friedrich, R. (2000). Uncertainties of modelling emissions from road transport. *Atmospheric Environment*, *34*(27), 4603–4610.

Kuhne, R. (1991). Traffic patterns in unstable traffic flow on freeways. In *International Symposium on Highway Capacity.*

Lebacque, J. P. (1997). A finite acceleration scheme for first order macroscopic traffic flow models. In M. Papageorigiou & A. Pouliezos (Eds.), *Proceedings of the 8th IFAC Symposium on Transportation Systems. Chania, Greece* (Vol. 2, p. 815-820).

Lebacque, J. P. (2002). A two phase extension of the LWR model based on the boundedness of traffic acceleration. In *Transportation and Traffic Theory in the 21st Century: Proceedings of the 15th International Symposium on Transportation and Traffic Theory, Adelaide, Australia, 16-18 July 2002* (pp. 697–718).

Lebacque, J. P. (2003). Two-phase bounded-acceleration traffic flow model: Analytical solutions and applications. *Transportation Research Record: Journal of the Transportation Research Board*(1852), 220–230.

Leutzbach, W. (1988). *Introduction to the theory of traffic flow* (Vol. 47). Springer.

Li, J., Chen, Q.-Y., Wang, H., & Ni, D. (2012). Analysis of LWR model with fundamental diagram subject to uncertainties. *Transportmetrica*, *8*(6), 387–405.

Li, M. (2008). A generic characterization of equilibrium speed-flow curves. *Transportation Science*, *42*(2), 220–235.

Li, Z., Liu, P., Xu, C., & Wang, W. (2015). Optimal mainline variable speed limit control to improve safety on large-scale freeway segments. *Computer-Aided Civil and Infrastructure Engineering*.

Lighthill, M. J., & Whitham, G. B. (1955a). On kinematic waves. I. Flood movement in long rivers. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* (Vol. 229, pp. 281–316).

Lighthill, M. J., & Whitham, G. B. (1955b). On kinematic waves. II. A theory of traffic flow on long crowded roads. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* (Vol. 229, pp. 317–345).

Lin, J., & Ge, Y. E. (2006). Impacts of traffic heterogeneity on roadside air pollution concentration. *Transportation Research Part D: Transport and Environment*, *11*(2), 166–170.

Lin, W., & Ahanotu, D. (1995). Validating the basic cell transmission model on a single freeway link. *PATH Technical Note; 95-3*.

Liu, S., De Schutter, B., & Hellendoorn, H. (2013). Multi-class traffic flow and emission control for freeway networks. In *16th International IEEE Conference on Intelligent Transportation Systems (ITSC)* (pp. 2223–2228).

Liu, S., De Schutter, B., & Hellendoorn, J. (2014). Integrated traffic flow and emission control based on FASTLANE and the multi-class VT-macro model. In *2014 European Control Conference (ECC)* (pp. 2908–2913).

Lo, H. K. (1999). A novel traffic signal control formulation. *Transportation Research Part A: Policy and Practice*, *33*(6), 433–448.

Loecher, M., & Ropkins, K. (2015). RgoogleMaps and loa: Unleashing R graphics power on map tiles. *Journal of Statistical Software*, *63*(4), 1–18. Retrieved from http://www.jstatsoft.org/v63/i04/

Maerivoet, S., & De Moor, B. (2005). Transportation planning and traffic flow models. *arXiv preprint physics/0507127*.

Mayer, H. (1999). Air pollution in cities. *Atmospheric Environment*, *33*(24), 4029–4037.

McCubbin, D. R., & Delucchi, M. A. (1999). The health costs of motor-vehicle-related air pollution. *Journal of Transport Economics and Policy*, 253–286.

Messner, A., & Papageorgiou, M. (1990). METANET: A macroscopic simulation program for motorway networks. *Traffic Engineering and Control*, *31*(8-9), 466–470.

Morgan, M. G., Henrion, M., & Small, M. (1992). *Uncertainty: A guide to dealing with uncertainty in quantitative risk and policy analysis*. Cambridge University Press.

Mott MacDonald. (2016). *MIDAS data.* Retrieved 2016-09-01, from https://www.midas-data.org.uk/

Mumovic, D., Crowther, J. M., & Stevanovic, Z. (2006). Integrated air quality modelling for a designated air quality management area in Glasgow. *Building and Environment*, *41*(12), 1703–1712.

Muñoz, L., Sun, X., Sun, D., Gomes, G., & Horowitz, R. (2004). Methodological calibration of the cell transmission model. In *Proceedings of the 2004 American Control Conference.* (Vol. 1, pp. 798–803).

Namdeo, A., Mitchell, G., & Dixon, R. (2002). TEMMS: An integrated package for modelling and mapping urban traffic emissions and air quality. *Environmental Modelling and Software*, *17*(2), 177–188.

Nejadkoorki, F., Nicholson, K., Lake, I., & Davies, T. (2008). An approach for modelling CO_2 emissions from road traffic in urban areas. *Science of the Total Environment*, *406*(1), 269–278.

Nelder, J. A., & Mead, R. (1965). A simplex method for function minimization. *The computer journal*, *7*(4), 308–313.

Newell, G. F. (1961). Nonlinear effects in the dynamics of car-following. *Operations Research*, *9*(2), 209–229.

Newell, G. F. (1993a). A simplified theory of kinematic waves in highway traffic, part I: General theory. *Transportation Research Part B: Methodological*, *27*(4), 281–287.

Newell, G. F. (1993b). A simplified theory of kinematic waves in highway traffic, part III: Multi-destination flows. *Transportation Research Part B: Methodological*, *27*(4), 305–313.

Newell, G. F. (1993c). A simplified theory of kinematic waves in highway traffic, part II: Queueing at freeway bottlenecks. *Transportation Research Part B: Methodological*, *27*(4), 289–303.

Newell, G. F. (2002). A simplified car-following theory: A lower order model. *Transportation Research Part B: Methodological*, *36*(3), 195–205.

Ngoduy, D. (2011). Multiclass first-order traffic model using stochastic fundamental diagrams. *Transportmetrica*, *7*(2), 111–125.

Ngoduy, D., Hoogendoorn, S., & Van Zuylen, H. (2004). Comparison of numerical schemes for macroscopic traffic flow models. *Transportation Research Record: Journal of the Transportation Research Board*(1876), 52–61.

Ngoduy, D., & Maher, M. J. (2012). Calibration of second order traffic models using continuous cross entropy method. *Transportation Research Part C: Emerging Technologies*, *24*, 102–121.

Nguyen, S., & Dupuis, C. (1984). An efficient method for computing traffic equilibria in networks with asymmetric transportation costs. *Transportation Science*, *18*(2), 185–202.

Ni, D., Leonard, J., Guin, A., & Williams, B. (2004). Systematic approach for validating traffic simulation models. *Transportation Research Record: Journal of the Transportation Research Board*(1876), 20–31.

Ntziachristos, L., Gkatzoflias, D., Kouridis, C., & Samaras, Z. (2009). COPERT: A European road transport emission inventory model. In *Information technologies in environmental engineering* (pp. 491–504). Springer.

Nyhan, M., Sobolevsky, S., Kang, C., Robinson, P., Corti, A., Szell, M., Streets, D., Lu, Z., Britter, R., Barrett, S. R. H., & Ratti, C. (2016). Predicting vehicular emissions in high spatial resolution using pervasively measured transportation data and microscopic emissions model. *Atmospheric Environment*, *140*, 352–363.

Orosz, G., Wilson, R. E., & Stépán, G. (2010). *Traffic jams: Dynamics and control.* The Royal Society.

Osorio, C., & Nanduri, K. (2015). Urban transportation emissions mitigation: Coupling high-resolution vehicular emissions and traffic models for traffic signal optimization. *Transportation Research Part B: Methodological*, *81*, 520–538.

Ossen, S., & Hoogendoorn, S. (2008). Validity of trajectory-based calibration approach of car-following models in presence of measurement errors. *Transportation Research Record: Journal of the Transportation Research Board*(2088), 117–125.

Palmer, T. N. (2000). Predicting uncertainty in forecasts of weather and climate. *Reports on Progress in Physics*, *63*(2), 71.

Papageorgiou, M. (1998). Some remarks on macroscopic traffic flow modelling. *Transportation Research Part A: Policy and Practice*, *32*(5), 323–329.

Papageorgiou, M., Blosseville, J. M., & Hadj-Salem, H. (1989). Macroscopic modelling of traffic flow on the Boulevard Périphérique in Paris. *Transportation Research Part B: Methodological*, *23*(1), 29–47.

Papageorgiou, M., Blosseville, J. M., & Hadj-Salem, H. (1990). Modelling and real-time control of traffic flow on the southern part of Boulevard Peripherique in Paris: Part I: Modelling. *Transportation Research Part A: General*, *24*(5), 345–359.

Papageorgiou, M., Kosmatopoulos, E., & Papamichail, I. (2008). Effects of variable speed limits on motorway traffic flow. *Transportation Research Record: Journal*

of the Transportation Research Board(2047), 37–48.

Papamichail, I., Kampitaki, K., Papageorgiou, M., & Messmer, A. (2008). Integrated ramp metering and variable speed limit control of motorway traffic flow. *IFAC Proceedings Volumes*, *41*(2), 14084–14089.

Payne, H. J. (1971). Models of freeway traffic and control. *Mathematical Models of Public Systems*.

Payne, H. J. (1979). FREFLO: A macroscopic simulation model of freeway traffic. *Transportation Research Record*(722).

Phillips, W. F. (1979). A kinetic model for traffic flow with continuum implications. *Transportation Planning and Technology*, *5*(3), 131–138.

Pipes, L. A. (1953). An operational analysis of traffic dynamics. *Journal of Applied Physics*, *24*(3), 274–281.

Poole, A., & Kotsialos, A. (2012). METANET model validation using a Genetic Algorithm. *IFAC Proceedings Volumes*, *45*(24), 7–12.

Poole, A., & Kotsialos, A. (2016). Swarm intelligence algorithms for macroscopic traffic flow model validation with automatic assignment of fundamental diagrams. *Applied Soft Computing*, *38*, 134–150.

Powell, M. (1998). Direct search algorithms for optimization calculations. *Acta numerica*, 287–336.

Prigogine, I., & Andrews, F. C. (1960). A Boltzmann-like approach for traffic flow. *Operations Research*, *8*(6), 789–797.

R Core Team. (2015). R: A Language and Environment for Statistical Computing [Computer software manual]. Vienna, Austria. Retrieved from http://www.R-project.org/

Rahman, M., Chowdhury, M., Xie, Y., & He, Y. (2013). Review of microscopic lane-changing models and future research opportunities. *IEEE Transactions on Intelligent Transportation Systems*, *14*(4), 1942–1956.

Rakha, H., & Zhang, W. (2005). Estimating traffic stream space mean speed and reliability from dual-and single-loop detectors. *Transportation Research Record: Journal of the Transportation Research Board*(1925), 38–47.

Refsgaard, J. C., van der Sluijs, J. P., Højberg, A. L., & Vanrolleghem, P. A. (2007). Uncertainty in the environmental modelling process–a framework and guidance. *Environmental modelling and software*, *22*(11), 1543–1556.

Rehimi, F., & Landolsi, J. (2013). The impact of traffic dynamic and wind angle on vehicular emission dispersion. *Transportation Research Part D: Transport and Environment*, *21*, 1–6.

Richards, P. I. (1956). Shock waves on the highway. *Operations Research*, *4*(1), 42–51.

Ropkins, K., Beebe, J., Li, H., Daham, B., Tate, J., Bell, M., & Andrews, G. (2009). Real-world vehicle exhaust emissions monitoring: Review and critical discussion. *Critical Reviews in Environmental Science and Technology*, *39*(2), 79–152.

Royston, J. P. (1982a). Algorithm AS 181: The W test for normality. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, *31*(2), 176–180.

Royston, J. P. (1982b). An extension of Shapiro and Wilk's W test for normality to large samples. *Applied Statistics*, 115–124.

Samaranayake, S., Glaser, S., Holstius, D., Monteil, J., Tracton, K., Seto, E., & Bayen, A. (2014). Real-time estimation of pollution emissions and dispersion from highway traffic. *Computer-Aided Civil and Infrastructure Engineering*, *29*(7), 546–558.

Samaras, C., Tsokolis, D., Toffolo, S., Garcia-Castro, A., Vock, C., Ntziachristos, L., & Samaras, Z. (2014). Limits of applicability of COPERT model to short links and congested conditions. In *20th International Transport and Air Pollution Conference.*

Sanwal, K. K., Petty, K., Walrand, J., & Fawaz, Y. (1996). An extended macroscopic model for traffic flow. *Transportation Research Part B: Methodological*, *30*(1), 1–9.

Sarkar, D. (2008). *Lattice: Multivariate data visualization with R.* New York: Springer. Retrieved from http://lmdvr.r-forge.r-project.org (ISBN 978-0-387-75968-5)

Seetharaman, P., Errampalli, M., Senathipati, V., Shukla, A., & Gangopadhyay, S. (2011). Modeling time mean speed and space mean speed for heterogeneous traffic conditions. *Transportation Planning and Technology*, *34*(8), 823–838.

Shang, H.-Y., & Huang, H.-J. (2007). A cell transmission model and its application in optimizing the location of Variable Message Signs. In *3rd International Conference on Natural Computation (ICNC)* (Vol. 4, pp. 75–79).

Shorshani, M. F., André, M., Bonhomme, C., & Seigneur, C. (2015). Modelling chain for the effect of road traffic on air and water quality: Techniques, current status and future prospects. *Environmental Modelling and Software*, *64*, 102–123.

Smit, R., & McBroom, J. (2009). Development of a new high resolution traffic emissions and fuel consumption model for Australia and New Zealand - Data quality considerations. *Air Quality and Climate Change*, *43*(2), 13.

Smit, R., Ntziachristos, L., & Boulter, P. (2010). Validation of road vehicle and traffic emission models - A review and meta-analysis. *Atmospheric Environment*, *44*(25), 2943–2953.

Spiliopoulou, A., Kontorinaki, M., Papageorgiou, M., & Kopelias, P. (2014). Macroscopic traffic flow model validation at congested freeway off-ramp areas. *Transportation Research Part C: Emerging Technologies*, *41*, 18–29.

Spiliopoulou, A., Papamichail, I., Papageorgiou, M., Tyrinopoulos, I., & Chrysoulakis, J. (2015). Macroscopic traffic flow model calibration using different optimization algorithms. *Transportation Research Procedia*, *6*, 144–157.

St-Denis, M. J., Cicero-Fernandez, P., Winer, A. M., Butler, J. W., & Jesion, G. (1994). Effects of in-use driving conditions and vehicle/engine operating parameters on "Off-Cycle" events: Comparison with Federal Test Procedure conditions. *Air and Waste*, *44*(1), 31–38.

Stedman, D. H. (1989). Automobile carbon monoxide emission. *Environmental Science and Technology*, *23*(2), 147–149.

Strategi®. (2015). *[SHAPE geospatial data] EDINA Digimap Ordnance Survey (OS) Service.* Retrieved 2016-10-31, from http://digimap.edina.ac.uk/

Sumalee, A., Zhong, R., Pan, T., & Szeto, W. (2011). Stochastic cell transmission model (SCTM): A stochastic dynamic traffic model for traffic state surveillance and assignment. *Transportation Research Part B: Methodological*, *45*(3), 507–533.

Sun, X., Muñoz, L., & Horowitz, R. (2003). Highway traffic state estimation using improved mixture Kalman filters for effective ramp metering control. In *Proceedings of the 42nd IEEE Conference on Decision and Control* (Vol. 6, pp. 6333–6338).

Treiber, M., Hennecke, A., & Helbing, D. (1999). Derivation, properties, and simulation of a gas-kinetic-based, nonlocal traffic model. *Physical Review E*, *59*(1), 239.

Treiber, M., & Kesting, A. (2012). Validation of traffic flow models with respect to the spatiotemporal evolution of congested traffic patterns. *Transportation Research Part C: Emerging Technologies*, *21*(1), 31–41.

Treiber, M., & Kesting, A. (2013). *Traffic flow dynamics: Data, models and simulation.* Springer-Verlag Berlin Heidelberg.

Tyagi, V., Darbha, S., & Rajagopal, K. R. (2009). A review of the mathematical models for traffic flow. *International Journal of Advances in Engineering Sciences and Applied Mathematics*, *1*(1), 53–68.

US EPA. (2015). *Notice of violation (18 September 2015) sent by EPA to Volkswagen Group of America, Inc.*

Van den Berg, M., Hegyi, A., De Schutter, B., & Hellendoorn, J. (2003). A macroscopic traffic flow model for integrated control of freeway and urban traffic networks. In *Proceedings of the 42nd IEEE Conference on Decision and Control* (Vol. 3, pp. 2774–2779).

van den Berg, M., Hegyi, A., De Schutter, B., & Hellendoorn, J. (2007). Integrated traffic control for mixed urban and freeway networks: A model predictive control approach. *European Journal of Transport and Infrastructure Research*, *7*(3), 223–250.

Van Lint, J. (2004). *Reliable travel time prediction for freeways: bridging artificial neural networks and traffic flow theory.* TU Delft, Delft University of Technology.

Van Lint, J., Hoogendoorn, S., & Schreuder, M. (2008). Fastlane: New multiclass first-order traffic flow model. *Transportation Research Record: Journal of the Transportation Research Board*(2088), 177–187.

Van Lint, J., & Van der Zijpp, N. (2003). Improving a travel-time estimation algorithm by using dual loop detectors. *Transportation Research Record: Journal of the Transportation Research Board*(1855), 41–48.

van Wageningen-Kessels, F. (2013). *Multi-class continuum traffic flow models: Analysis and simulation methods* (Unpublished doctoral dissertation). TRAIL.

van Wageningen-Kessels, F., Van Lint, H., Vuik, K., & Hoogendoorn, S. (2015). Genealogy of traffic flow models. *EURO Journal on Transportation and Logistics*, *4*(4), 445–473.

Venables, W. N., & Ripley, B. D. (2002). *Modern Applied Statistics with S* (Fourth ed.). New York: Springer. Retrieved from http://www.stats.ox.ac.uk/pub/MASS4 (ISBN 0-387-95457-0)

Walker, W. E., Harremoës, P., Rotmans, J., van der Sluijs, J. P., van Asselt, M. B., Janssen, P., & Krayer von Krauss, M. P. (2003). Defining uncertainty: A conceptual basis for uncertainty management in model-based decision support. *Integrated Assessment*, *4*(1), 5–17.

Willsky, A., Chow, E., Gershwin, S., Greene, C., Houpt, P., & Kurkjian, A. (1980). Dynamic model-based techniques for the detection of incidents on freeways. *IEEE Transactions on Automatic Control*, *25*(3), 347–360.

Wilson, R. E. (2008). Mechanisms for spatio-temporal pattern formation in highway traffic models. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, *366*(1872), 2017–2032.

Wismans, L., Van Berkum, E., & Bliemer, M. (2011). Modelling externalities using dynamic traffic assignment models: A review. *Transport Reviews*, *31*(4), 521–545.

Wismans, L. J. J. (2012). *Towards sustainable dynamic traffic management.* University of Twente.

Wong, G. C. K., & Wong, S. C. (2002). A multi-class traffic flow model–an extension of lwr model with heterogeneous drivers. *Transportation Research Part A: Policy and Practice*, *36*(9), 827–841.

Yang, Q., & Koutsopoulos, H. N. (1996). A microscopic traffic simulator for evaluation of dynamic traffic management systems. *Transportation Research Part C: Emerging Technologies*, *4*(3), 113–129.

Ye, Y. (1987). *Interior algorithms for linear, quadratic, and linearly constrained nonlinear programming* (Unpublished doctoral dissertation). Department of ESS, Stanford University.

Yu, L., Jia, S., & Shi, Q. (2009). Research on transportation-related emissions: Current status and future directions. *Journal of the Air and Waste Management Association*, *59*(2), 183–195.

Zallinger, M., Tate, J., & Hausberger, S. (2008). An instantaneous emission model for the passenger car fleet. In *16th International Transport and Air Pollution Congress.*

Zappa, M., Jaun, S., Germann, U., Walser, A., & Fundel, F. (2011). Superposition of three sources of uncertainties in operational flood forecasting chains. *Atmospheric Research*, *100*(2), 246–262.

Zegeye, S. K., De Schutter, B., Hellendoorn, H., & Breunesse, E. (2009). Model-based traffic control for balanced reduction of fuel consumption, emissions, and travel time. *IFAC Proceedings Volumes*, *42*(15), 149–154.

Zegeye, S. K., De Schutter, B., Hellendoorn, J., & Breunesse, E. A. (2010). Variable speed limits for area-wide reduction of emissions. In *13th International IEEE Conference on Intelligent Transportation Systems (ITSC)* (pp. 507–512).

Zegeye, S. K., De Schutter, B., Hellendoorn, J., & Breunesse, E. A. (2011). Nonlinear MPC for the improvement of dispersion of freeway traffic emissions. *IFAC Proceedings Volumes*, *44*(1), 10703–10708.

Zegeye, S. K., De Schutter, B., Hellendoorn, J., Breunesse, E. A., & Hegyi, A. (2013). Integrated macroscopic traffic flow, emission, and fuel consumption model for control purposes. *Transportation Research Part C: Emerging Technologies*, *31*, 158–171.

Zhang, H. (1998). A theory of nonequilibrium traffic flow. *Transportation Research Part B: Methodological*, *32*(7), 485–498.

Zhang, H. (2001). New perspectives on continuum traffic flow models. *Networks and Spatial Economics*, *1*(1-2), 9–33.

Zhang, L., Yin, Y., & Chen, S. (2013). Robust signal timing optimization with environmental concerns. *Transportation Research Part C: Emerging Technologies*, *29*, 55–71.

Zhong, R., Chen, C., Chow, A. H., Pan, T., Yuan, F., & He, Z. (2015). Automatic calibration of fundamental diagram for first-order macroscopic freeway traffic models. *Journal of Advanced Transportation*.

Zhou, X., Tanvir, S., Lei, H., Taylor, J., Liu, B., Rouphail, N. M., & Frey, H. C. (2015). Integrating a simplified emission estimation model and mesoscopic dynamic traffic simulator to efficiently evaluate emission impacts of traffic management strategies. *Transportation Research Part D: Transport and Environment*, *37*, 123–136.

Zhu, F., Lo, H. K., & Lin, H. Z. (2013). Delay and emissions modelling for signalised intersections. *Transportmetrica B: Transport Dynamics*, *1*(2), 111–135.